Assessment of the Florida College and Career Readiness Initiative: 2018 Final Technical Report

Christine Mokher, Florida State University/CNA Education

Daniel Leeds, CNA Education

Julie Harris, CNA Education

May 2018

CNA ANALYSIS & SOLUTIONS

Acknowledgments

This study was made possible by the collaboration and hard work of many individuals beyond the authors. We would like to thank other current and former members of the research team, including Louis Jacobson, James Rosenbaum, Robert LaLonde, John Hughes, Thomas Geraghty, Juliana Pearson, and Michael Flory. We would also like to thank the members of our Technical Working Group for comments on this report and/or previous related reports: David Figlio, Stephen Raudenbush, and Jeffrey Smith.

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305E120010 to the CNA Corporation. The report represents the best opinion of CNA at the time of issue and does not represent views of the Institute or the U.S. Department of Education.

Distribution

Distribution unlimited. Specific authority contracting number: R305E120010.

Other requests for this document shall be referred to inquiries@cna.org.

Abstract

Florida enacted legislation in 2008 for a statewide program known as the Florida College and Career Readiness Initiative (FCCRI), which was intended to reduce the need for postsecondary remediation. The FCCRI consisted of testing grade 11 students to determine their college readiness and offering math and English college readiness and success (CRS) courses in grade 12 for students who did not test college-ready the year before. The theory of action for the FCCRI is that providing college readiness testing and CRS courses in high school may raise students' awareness of their academic deficiencies and motivate them to further develop college-level skills in their senior year. We found considerable variation across districts and schools in the initiative's implementation and level of compliance with state requirements for participation.

We estimated program impacts using two different methods. First, we used a regression discontinuity design to compare outcomes for students scoring just above and below test score cutoffs for assignment to the FCCRI. Among the two cohorts of students required to participate in the FCCRI, we found little to no impact on short-term outcomes including high school graduation, college enrollment, and enrolling in or passing non-developmental courses. There is little evidence for improved enrollment or pass rates in for-credit coursework among the highest- and lowest-performing targeted students. However, enrollment and pass rates in transition and degree credit courses were similar for students on the margins of assignment to college readiness courses in either subject, indicating that students just below college-ready were able to "catch up" by the time they enrolled in college. There were also few differences between marginal targeted and non-targeted students on longer-term outcomes including persistence, transfer, non-developmental enrollment and pass, and degree completion rates.

Second, we examined the impact of offering the FCCRI to students from a wider range of academic performance levels by using regression analysis to compare student outcomes for targeted students in schools before and after the schools implemented the FCCRI. As with the regression discontinuity analyses, we found little to no effect of the FCCRI on most short-term student outcomes. We did find that the treatment group was more likely to both take and pass nondevelopmental courses in math and English. Although the average effects were small, the magnitudes of these effects were quite large for some portions of the achievement distribution, with impacts of up to 10.7 percentage points for the treatment group.

Finally, we estimated the cost for the FCCRI's ongoing implementation at $57 per targeted student in 2014/15, with about 63 percent of per-student program costs incurred at the school level ($36), 33 percent at the district level ($19), and 3 percent at the state level ($2). We also found that while net costs exceeded benefits for the FCCRI as a whole, program impacts and costs varied both within and between districts.

This page intentionally left blank.

Executive Summary

There is nationwide concern about the number of students who leave high school unprepared for college-level coursework and the lack of student awareness of their level of preparation. It is estimated that in 2010, 86 percent of community college students believed that they were academically prepared for college, yet 67 percent tested into developmental coursework (Center for Community College Student Engagement, 2016). Results were similar in Florida, where 63.6 percent of first-time degree-seeking students at two-year state colleges did not meet the college-ready entry-level scores in at least one subject on the placement test and were required to enroll in developmental education (Florida College System, 2012a).

Florida attempted to address these concerns by implementing the Florida College and Career Readiness Initiative (FCCRI), a statewide initiative that provided college placement testing to grade 11 students who were mid-performing on the grade 10 Florida Comprehensive Achievement Test (FCAT), and college readiness and success courses (CRS) in grade 12 to students who scored below the threshold for college-ready on the placement test in math and reading. Student participation in both the college readiness testing and the courses was voluntary when the initiative began in 2008/09. Legislative changes in 2011/12 required college readiness testing in grade 11 for mid-performing students on the grade 10 state assessment and CRS course participation in grade 12 for students who do not test college-ready. Additional legislative changes in 2015 made participation in the FCCRI voluntary once again. The theory of action for the FCCRI is that providing college readiness testing and CRS courses in high school may raise students' awareness of their academic deficiencies and motivate them to further develop college-level skills in their senior year. The CRS courses were designed to develop these skills by improving the alignment between the content taught in grade 12 courses and first year college courses in math and English. In English many students would otherwise be taking standard English IV, while in math many students would otherwise be taking courses such as Pre-Calculus, Financial Applications, Trigonometry, or Math Analysis.

Data and analysis

The primary data source consists of student-level records from the Florida K–20 Education Data Warehouse, which follows all Florida public school students from grade 10 through postsecondary education as long as they remain in Florida and attend a public high school, college or university. These data are supplemented with school-level variables from the National Center for Education Statistics'

Elementary/Secondary Information System and reports produced by the Florida Department of Education (FLDOE).

This study used two types of research designs to examine the FCCRI's effects on student outcomes. First, we used a regression discontinuity design to compare outcomes for students scoring just above and below test score cutoffs for assignment to the FCCRI. For this analysis, we focused on students who were in grade 11 in 2011/12 and 2012/13, which are the only two cohorts in which targeted students were required to take both the Postsecondary Education Readiness Test (PERT) in grade 11 and to enroll in a CRS course in grade 12 if they scored below college-ready. This methodology was used to assess the FCCRI's impact on both short-term and long-term outcomes (chapter 4), and whether it provided a net benefit (chapter 6).

Second, as the regression discontinuity analysis can indicate the FCCRI's effectiveness only for students near the treatment cutoffs, we used regression analysis with a before-after design to shed light on the FCCRI's overall impact for students from a wider range of academic performance levels. We used variation in school-level FCCRI compliance rates as an exogenous source of assignment to treatment to obtain an analytical sample in which assignment to treatment is conditionally independent from the outcomes of interest. We considered high compliance schools to have implemented the FCCRI and low-compliance schools to have not; from this division, we obtained a treated group and a comparison group, enabling us to compare student outcomes before and after schools implemented the FCCRI. We calculated the school-level compliance rate as the proportion of grade 12 students who were targeted by the FCAT, did not score college-ready on the placement test, and went on to enroll in a CRS course. We defined the treatment group as students who were targeted by the FCAT and attended a high school in grade 12 with at least a 50 percent compliance rate (high-compliance schools). The comparison group was limited to students who were targeted by the FCAT and attended a high school in grade 12 with less than a 5 percent compliance rate (low-compliance schools). We also limited the analytical sample to schools that were categorized as both low- and high-compliance at some point during the initiative to control for differences in schools across compliance rates. We then used regression analysis to estimate the intent-to-treat effect of the FCCRI for students targeted for the FCCRI based on their FCAT scores.

This report also describes the initiative's implementation, using administrative records on PERT test-taking and CRS course participation rates; site visits to high schools, district offices, and state colleges; and interviews and focus groups with educators. Lastly, the cost-benefit analysis draws on administrative records and interviews with state-, district-, and school-level administrators about resources allocated to support the FCCRI. The ingredients method was used to identify and price inputs for program implementation. This information on costs was compared with findings on the benefits of the FCCRI from the RD analysis.

Key findings

When the FCCRI began in 2008/09, student participation in both the college readiness testing and the CRS courses was voluntary. In 2011/12, participation in both components became mandatory for targeted students. The mandatory FCCRI continued through 2014/15. This report includes findings for the three cohorts of students in the voluntary FCCRI (cohorts V1, V2, and V3) and the subsequent two cohorts under the mandatory FCCRI (cohorts M1 and M2).

FCCRI implementation

- During the first year of the voluntary FCCRI, only about half of all high schools offered college readiness testing and fewer than one-third offered CRS courses even though all schools were supposed to offer both interventions. The percentage of schools offering each component rose slightly during each subsequent year of the voluntary FCCRI. There was a large increase in participation during the two years of the mandatory FCCRI, with over 90 percent of schools offering both college readiness testing and CRS courses.
- Even though FCCRI participation was supposed to be required for targeted students under the mandatory FCCRI, schools did not always adhere to student eligibility criteria in the state policy.
    - The PERT was taken by 58 to 73 percent of students (depending on cohort and subject area) in the mandatory FCCRI who were targeted to take the test based on their FCAT scores. Between 58 and 65 percent of students who scored below college-ready on the PERT enrolled in a CRS course in grade 12.
    - Some students took the PERT and enrolled in the CRS courses even though they were not required to participate under the mandatory FCCRI. Their participation rates ranged from 5 percent to 25 percent, depending on FCAT level and cohort.
- The content and rigor of the CRS courses varied substantially across schools. In some schools the CRS courses were very similar to the courses taken by students prior to the FCCRI. Yet in other schools the CRS courses placed more emphasis on developing academic skills for college and testing college-ready.
- In both math and English, targeted students who scored just above college-ready were more likely to enroll in honors and college credit–bearing courses than students who scored just below college-ready. However, we also found a small increase (3.5 percentage points) in the likelihood of taking no math in the senior year for students who scored just above college-ready on the PERT in grade 11.

Confirmatory impact analysis using regression discontinuity

- Data on the two mandatory FCCRI cohorts fit the What Works Clearinghouse's eligibility standards for regression discontinuity analysis. At a minimum, all estimates contained in this study meet What Works Clearinghouse (WWC, 2017) RDD standards with reservations.

- Short-term outcomes include the probabilities of receiving a high school diploma or equivalent, seamless college enrollment (i.e. in the fall after high school graduation), and enrolling in or passing for-credit coursework in math or English. For the latter two outcomes, we restricted our sample to students who enrolled in the fall after on-time high school graduation.

- For cohort M1, effects near FCAT cutoffs for PERT targeting were generally insignificant and/or small in magnitude.

    - Students targeted for any stage of the FCCRI in either subject were no more likely to graduate high school; or seamlessly enroll in college.

    - Being targeted for the FCCRI had no visible impact on enrolling in or passing a non-remedial math course. In English, students targeted for PERT testing at the low FCAT margin were 4.3 percentage points more likely to take a non-developmental English class but no more likely to pass one.

- For cohort M1, the FCCRI may have been helped students scoring near the college readiness benchmark in grade 11 prepare for college.

    - Students who barely scored college-ready were more likely to enroll in honors or other advanced courses in grade 12 than those targeted for CRS courses.

    - Students targeted for CRS courses performed comparably to those who were already college-ready in grade 11, which suggests the FCCRI may have helped them "catch up" by the time they enrolled in college. Based on performance in grade 11, all students who were below college-ready had scores corresponding to developmental courses, while all students who were above college-ready were exempt from developmental education. Thus, if the students' performance remained the same between grade 11 and college enrollment, we would expect those scoring just below college-ready to be much more likely to be placed into developmental education courses. Yet in both subjects, there was little to no difference in the likelihood of passing nondevelopmental courses.

- For cohort M2, results also tended to be insignificant and/or small in magnitude. One difference for this cohort was that students at the college readiness cutoffs on the grade 11 PERT in either subject were less likely to enroll in non-developmental courses but no less likely to pass.

- We also examined the impact of the FCCRI two years after on-time high school graduation for cohorts M1 and M2 and three years after on-time high school graduation for cohort M1.
  - Long-term outcomes included persistence in a postsecondary institution, transfer from a two-year institution to a four-year institution, and degree receipt. For year two outcomes for cohort M1, we also included the probability of enrolling in or passing a non-developmental course
  - Estimates for both mandatory cohorts were generally statistically insignificant and/or small in magnitude, suggesting little long-term impact on student outcomes for students near the cutoffs for assignment to the FCCRI.

Exploratory impact analysis using regression analysis

- The exploratory analyses used regression analysis with a before-after design to examine the same outcomes from the confirmatory impact analyses, while using a sample of students from a wider range of pretreatment achievement levels.
- Most of the results were similar to those of the confirmatory impact analyses. Both analyses found little to no effect on high school graduation, college enrollment, passing three or more credit-bearing courses, and college GPA.
- The regression analysis found a small increase (2 to 3 percentage points) in the rates of both enrolling in and passing nondevelopmental courses in both math and English. The effects were larger (3.5 to 6.0 percentage points) for the subsample of students who seamlessly enrolled in college.
- Using a multinomial course-level outcome, we examined how enrollment in lower- and upper-level developmental education was impacted by the FCCRI. This sample was limited to students who seamlessly enrolled in college, an outcome that the FCCRI did not have an impact on. We found that the treatment group was less likely to enroll in developmental education courses, especially in math, and the impact varied across baseline achievement.
  - Students at the upper end of the targeted range saw little to no impact on course-level enrollment in both math and English. Many of these students were college-ready by the time they enrolled in college regardless of whether they received any intervention.
  - The largest overall positive effect was among mid-performing targeted students in math, where enrollment in non-developmental education

courses increased by 10.7 percentage points. Most students moving to a higher course level were successful in completing the course.

Cost-benefit analysis

- The annual cost of the FCCRI (not including startup costs) was approximately $57 per targeted student in 2014, for a total of. $7.3 million in 2014 dollars based on a sub-sample of schools that participated in qualitative data collection.
- Of these costs, 63 percent were at the school level, 33 percent at the district level, and 3 percent at the state level. By ingredient, 68 percent of costs were on personnel, 30 percent on materials and equipment, and 1 percent on facilities.
- There was substantial variation in both costs and (gross) benefits by school and district.
    - Of nine schools at which interviews were conducted, the lowest costs (combined across all levels and ingredients) were $23 per student and the highest were $312 per student.
- Seven of eight districts had positive gross benefits from the FCCRI at one or more cutoffs (the eighth met sample size requirements at only one cutoff). Net benefits of the FCCRI were generally negative.
    - Average costs per student were greater than benefits for every district cutoff in six districts at the upper FCAT and PERT cutoffs in reading. Net benefits were negative for almost all districts at the upper FCAT cutoff in math.
    - Net benefits were positive for four out of eight districts at the PERT cutoff in math; however, the positive net benefits for these districts were smaller in magnitude than the negative net benefits at the other four districts.
    - If we assume that our cost analysis is representative of schools statewide, the FCCRI would need to move 7.5 percent of course-takers in either subject from passing an upper DE course to passing a for-credit course to be cost-effective.

Conclusions

The findings may have implications for researchers and policymakers considering similar college readiness programs in terms of lessons learned about the initiative's design and implementation. These findings may also help to explain why the FCCRI had few discernable effects on student outcomes.

Lessons learned about the design of the FCCRI:

- The FCCRI did not seem to adequately consider students' motivation to attend college. The initiative was focused on helping students become college-ready regardless of whether they intended to enroll in college. No additional advising or support was provided to non-college-bound students to improve their career options or to encourage them to consider postsecondary programs.
- The FCCRI likely targeted students from too wide range of achievement levels. Some higher-performing students took CRS courses at the expense of more rigorous grade 12 courses, while lower-performing students may have needed more than a single course to become college-ready. A single course cannot meet the needs of every student from such a wide range of achievement levels.
- The FCCRI was largely an unfunded mandate, resulting in varying reallocations of existing resources among schools and districts. This disparity led to uneven implementation across districts and schools, some of which may not have had the resources to serve all targeted students adequately.
- While FLDOE provided funding for college readiness testing and it approved CRS courses, this may not have been enough to successful implement the initiative. Teachers and administrators lacked guidance beyond broad standards on what CRS courses should over or how they would do so. This confusion likely contributed to lower and less uniform instructional quality that lacked alignment with the original intent of the initiative.
- It may have been more effective to begin participation in the FCCRI in earlier grade levels. The IES Practice Guide on pathways to college suggests offering college readiness assessments throughout high school and recommends that students begin preparing for college-level work by ninth grade (Tierney, Bailey, Constantine, Finkelstein, & Hurd; 2009).

Lessons learned about the implementation of the FCCRI:

- Schools and districts often did not follow state requirements for participation in the FCCRI, which may have been exacerbated by the lack of enforcement or sanctions for non-compliance. Under the voluntary FCCRI, most schools did not offer college readiness testing, CRS courses, or both. Although school-level compliance increased under the mandatory FCCRI, many targeted students did not participate.
- More communication was needed about how to implement the FCCRI, as teachers were often unaware of the content of the PERT or of how CRS courses were supposed to prepare students for college-level work. While FLDOE recommended that high schools collaborate with local colleges, we found very

little evidence of collaboration around CRS courses or college readiness more broadly. The few educators who did participate in cross-sector collaboration indicated that they found it very helpful, and those who did not participate expressed a desire for more opportunities to do so.

- Though CRS courses shared the same names and course numbers statewide, they were implemented differently across and sometimes even within schools. This made it difficult to evaluate the program as a whole. CRS courses will have little effect if they do not contrast with existing courses (e.g. some English CRS courses covered the same material in the same way as their schools' regular English 4 courses).

Contents

Executive Summary ..... **v**

Data and analysis v
Key findings vii
Conclusions x

Chapter 1: Introduction and Study Overview **1**

Chapter 2: Description of the Florida College and Career Readiness Initiative (FCCRI) **5**

Policy context 5
Theory of action 18

Chapter 3: Implementation of the FCCRI **23**

School, district, and state roles in implementing the FCCRI 24
Participation in college readiness testing 24
CRS course participation 32
Description of treatment and counterfactual conditions 36

Chapter 4: Confirmatory Impact Analysis of the FCCRI Using Regression Discontinuity **45**

Econometric methodology 45
Data and sample 51
Regression discontinuity validity requirements 53
Results from regression discontinuity analysis 56

Chapter 5: Exploratory Impact Analysis of the FCCRI Using a Before-After Regression Analysis **67**

Methodology 68
Data and sample 71
Sample analysis 71
Estimation Results 72
Limitations 78

Chapter 6: District Cost-Benefit Analysis **80**

Cost-benefit methods and data 81

Cost-benefit findings 86
Limitations 96
Implications 96

Chapter 8: Conclusions 98

Implementation analysis 98
Confirmatory impact analysis of the FCCRI 98
Exploratory impact analysis of the FCCRI 100
Cost-benefit analysis 101
Policy implications 102

Appendix A: Technical Appendix for the Confirmatory Impact Analysis 105

Descriptive statistics of student characteristics 105
Outcomes variables for confirmatory and exploratory impact analyses 110
McCrary density test results examining the integrity of the running variable ... 111
Attrition at the RD cutoffs 114
Baseline equivalence 115
Computing bandwidths through cross-validation 133

Appendix B: Technical Appendix for the Exploratory Impact Analysis 138

Appendix C: Technical Appendix for the Cost Analysis 140

References 143

List of Figures

Figure 1. Assessments used to target students for participation in the FCCRI .. 5
Figure 2. Timeline summarizing activities affecting the FCCRI 16
Figure 3. Logic model for the theory of action of the FCCRI 19
Figure 4. Distribution of FCAT levels for cohort M1 (targeted levels shaded in dark blue) 25
Figure 5. Characteristics of FCAT targeted students compared with those of students below target and above target 26
Figure 6. Percentage of schools offering college readiness testing, by cohort .. 26
Figure 7. Student-level participation rates in college readiness testing among FCAT targeted students 28
Figure 8. Participation rates in college readiness testing for targeted students in math and reading for cohort V3 and cohort M1, by FCAT level (targeted levels shaded in dark blue) 29
Figure 9. Tree diagram illustrating student participation and performance on the PERT: grade 11 students in 2011/12 31
Figure 10. Percentage of schools offering at least one CRS course, by cohort 33
Figure 11. Student-level participation rates in CRS courses for targeted students who scored below college ready on the CPT or PERT, by cohort 34
Figure 12. Student-level CRS course participation rates for targeted students, by placement level on the grade 11 PERT (developmental education courses shaded in dark blue) 35
Figure 13. Percentage of targeted students enrolling in each English course level from cohort V2 to cohort M2, by FCAT level 41
Figure 14. Percentage of targeted students enrolling in each math course level from cohort V2 to cohort M2, by FCAT level 42
Figure 15. Percentage of targeted students enrolling in each math course type from cohort V2 to cohort M2, by FCAT level. 43
Figure 16. Compliance with FCCRI assignment 49
Figure 17. Sequence of college-level courses in math and English at all Florida colleges. 50
Figure 18. Enrollment and passing by assessment score, cohort M1 60
Figure 19. Enrollment and passing by assessment score, cohort M2 62
Figure 20. Lowest level of math and English course taken in college for fall-starters in cohort M1, by FCAT level 67
Figure 21. Diagram of comparison and treatment groups for the regression analysis 69
Figure 22. Math course-level enrollment estimation results, by treatment status and FCAT level 76

Figure 23. English course-level enrollment estimation results, by treatment status and FCAT level ... 77
Figure 24. Pass rates for the first course taken in math and English, by treatment status and FCAT level ... 78

List of Tables

Table 1. Description of postsecondary course levels for developmental and nondevelopmental courses 11
Table 2. Timeline for cohort V1 through cohort M2 17
Table 3. Characteristics of schools that first adopted testing under the voluntary FCCRI versus those of schools that first adopted under the mandatory FCCRI 27
Table 4. Performance on PERT retakes for students who scored below college-ready on PERT in grade 11 in 2011/12 32
Table 5. Characteristics of schools that first adopted CRS courses under the voluntary FCCRI versus those of schools that first adopted under the mandatory FCCRI 33
Table 6. Assignment to treatment by assessment, cohort M1 46
Table 7. Assignment to treatment by grade 10 assessment, cohort M2 47
Table 8. Number of students testing multiple times by cohort, assessment, and subject 52
Table 9. Regression discontinuity results for the outcomes of high school graduation, seamless college enrollment, and nondevelopmental enrollment and passing, cohort M1 57
Table 10. Regression discontinuity results for the outcomes of high school graduation, seamless college enrollment, and nondevelopmental enrollment and passing, cohort M1 61
Table 11. Year 2 outcomes for cohort M1 63
Table 12. Year 3 outcomes for cohort M1 64
Table 13. Year 2 outcomes for cohort M2 65
Table 14. Control variables included in regression analysis models 70
Table 15. Sample formation, by targeted subject 71
Table 16. Rates of treatment, by subject and treatment status 72
Table 17. Before-after regression estimated marginal effect of treatment for the seamless college enrollee subsample 73
Table 18. Before-after regression estimated marginal effect of treatment for the seamless college enrollee subsample 74
Table 19. Cost-benefit student characteristics by district 81
Table 20. Characteristics of districts included in the cost study 84
Table 21. Remaining costs before completion of for-credit course. 88
Table 22. Cost savings per student at RD cutoffs 88
Table 23. Per-student annual cost of the FCCRI, by cost ingredient 89
Table 24. Replacement cost per student (including school, district, and state level) and total cost across FCCRI sites 91
Table 25. Net FCCRI benefit per student 93

Glossary

ADP	American Diploma Project
AICE	Advanced International Certificate of Education
AP	Advanced Placement
CBCSE	Center for Benefit-Cost Studies of Education
CCD	Common Core of Data
CPT	College Entry-Level Placement Test
CRS	college readiness and success
EDW	Education Data Warehouse
ELL	English language learner
EOC	end-of-course assessment
FCAT	Florida Comprehensive Assessment Test
FCCRI	Florida College and Career Readiness Initiative
FLDOE	Florida Department of Education
FRPL	free and-reduced-price lunch
GPA	grade point average
IB	International Baccalaureate
PERT	Postsecondary Education Readiness Test
PL	performance level (on the FCAT)
PRC	postsecondary readiness competencies
RD	regression discontinuity
SY	school year
WWC	What Works Clearinghouse

Chapter 1: Introduction and Study Overview

This study evaluates the implementation, impacts, and costs of the Florida College and Career Readiness Initiative (FCCRI). This statewide initiative consisted of testing grade 11 students statewide to determine their college readiness and requiring students who did not test college-ready to take math and/or English college readiness and success (CRS) courses in grade 12. The FCCRI addressed two interrelated nationwide concerns: how to make high school students aware of the standards to succeed in for-credit college courses and how to help students meet college entrance standards without setting high school graduation standards so high that many will not be able to graduate from high school (Betts & Grogger, 2003; Jacob, 2001; Lillard & DeCicca, 2001).

Prior to the adoption of the FCCRI, state policymakers were seeking a way to improve college access, readiness, and completion. In 2008, 54 percent of recent Florida high school graduates who enrolled in community colleges required remediation (Florida Department of Education, 2013). The chances of obtaining two-year degrees or completing courses of value in the workplace are much lower for students who need remediation than for students who do not (for example, Furchtgott-Roth, Jacobson, & Mokher, 2009; Martorell & McFarlin, 2011). Providing remediation is also very resource-intensive—State Impact Florida (2012) estimates that $168 million was spent annually on developmental education in Florida.

One of the reasons postsecondary remediation rates are so high is that many students who receive high school diplomas do not have the requisite skills to complete college-level work (for example, Achieve and The Education Trust, 2008; Boser & Burd, 2009; Strong American Schools, 2008). Further, these students often do not recognize that they lack the preparation necessary to complete (introductory) for-credit college courses and that they will be required to enroll in developmental education in college. This gap in understanding is particularly apparent for first-generation college-going students (Kirst, 2005; Kirst & Bracco, 2004). The theory of action for the FCCRI is that providing college readiness testing and CRS courses in high school may raise students' awareness of their academic deficiencies, motivate them to further develop college-level skills, and improve their ability to gain those skills in their senior year—a period when many students could take courses to adequately prepare for college but fail to do so (Kirst, 2001; National Commission on the High School Senior Year, 2001; Peterson, 2003; Rosenbaum, 2001).

This evaluation of the FCCRI aims to provide rigorous evidence about a potentially highly cost-effective means to address the national problems of students' underpreparation for college and lack of awareness of academic deficiencies. In the absence of the FCCRI, students may be more likely to take a grade 12 math or English course that is not as well aligned with the content of first year college courses. In English many students would otherwise be taking standard English IV, while in math many students would otherwise be taking courses such as Pre-Calculus, Financial Applications, Trigonometry, or Math Analysis.

Several additional features make an evaluation of the FCCRI important:

- The initiative uses a single college readiness test throughout a large and diverse state. This feature, which could be replicated elsewhere, establishes a common standard that all high schools and colleges can work toward meeting and fosters information sharing about what works.
- Rigorously evaluating and collecting information about the FCCRI's strengths and weaknesses from the point of view of high schools and districts can provide valuable guidance to over two dozen states that have implemented similar programs (Barnett, Fay, Trimble, & Pheatt, 2013).
- Florida routinely collects all the basic high school, college, and workforce data needed for the impact evaluation and related work. This study illustrates the advantages of conducting a comprehensive evaluation using information from a single, well-articulated K–20 and workforce data warehouse.

Overview of the FCCRI

The FCCRI began in 2008/09 with the passage of Senate Bill 1908, which required all public high schools to offer a common college readiness assessment to mid-performing grade 11 students with Florida Comprehensive Achievement Test (FCAT) achievement levels of 3 (out of 5) in reading and 3 or 4 (out of 5) in math. The law also required high schools to offer postsecondary preparatory instruction to grade 12 students who scored below the state-established college readiness benchmarks. However, student participation in both the college readiness testing and the postsecondary preparatory instruction was voluntary, so we refer to this program as the voluntary FCCRI. There were three cohorts of students under the voluntary FCCRI program, herein referred to as V1, V2, and V3. High school students scoring at or above the college readiness benchmarks were exempt from additional college testing and from developmental education courses if they entered a Florida college within two years of taking the test.

In 2011/12, House Bill 1255 made participation in the college readiness assessment mandatory for grade 11 students with FCAT achievement levels of 2 and 3 in reading and 2, 3, or 4 in math. Starting in 2012/13, CRS course participation also became mandatory for students who scored below college-ready on the assessment. We refer to this program as the mandatory FCCRI. There were two cohorts of students required to participate in both testing and CRS courses under the mandatory FCCRI, herein referred to as M1 and M2.

High schools, state colleges, and the Florida Department of Education (FLDOE) have all played a role in implementing the FCCRI. High schools were responsible for identifying students eligible for the college readiness testing, reviewing test results with the students, counseling students about CRS and general senior year course selection to improve college readiness, and providing CRS courses designed to build the skills needed to become college-ready. FLDOE was responsible for overseeing the FCCRI, organizing the testing, and reviewing postsecondary preparatory courses for approval. Through 2010/11, only state colleges were allowed to administer the college readiness assessment, and the colleges had to form partnerships with local high schools to test students. Starting in 2011/12, the state transitioned to a new postsecondary readiness assessment that could be administered directly by high schools.

Evaluation goals

The **first major goal** of the project was to evaluate the FCCRI's impact on students' short- and long-term educational outcomes. Short-term outcomes include high school graduation, college enrollment, enrollment and passing rates for a college-level course in the same subject area in which the student was targeted for FCCRI participation, completion of three or more for-credit college courses in any subject, and cumulative grade point average (GPA) by the end of the first year of college enrollment. Long-term outcomes are college persistence, enrollment and pass rates in non-developmental courses in subsequent years, transfer from two-year to four-year institutions, and completion of a credential (certificate or degree).

Two different nonexperimental approaches were used to examine the FCCRI's impact on student outcomes. First, a regression discontinuity design was used to examine the impact for the subset of students who were just above or just below the cutoff value for assignment to the FCCRI. The results from these analyses meet the What Works Clearinghouse (WWC) regression discontinuity design evidence standards with reservations (U.S. Department of Education, 2016). Second, a school-level fixed effects analysis was used to evaluate the FCCRI's impact for targeted students who were enrolled in schools prior to compliance with the FCCRI to students enrolled in the same schools after compliance. This analysis meets WWC (2017) group design evidence standards with reservations, and allowed us to examine the FCCRI's overall impact for students from a wider range of academic performance levels.

The **second major goal** of the project was to examine the FCCRI's implementation and collect feedback from state-, district-, and school-level educators. There were three primary purposes for collecting this information. First, the data would contextualize the results of the impact estimates. Educational interventions often experience lower program implementation levels than intended, and information about impediments can provide insight into factors associated with variation in program effectiveness. Second, the information would provide FLDOE with schools' perceptions of how well the initiative was meeting objectives and the major impediments to its successful implementation. The state does not have any mechanisms in place to receive systematic feedback on the program from a representative group of stakeholders. Instead, officials tend to hear about a range of complaints from a small number of high schools and districts. Our research seeks to fill this gap, leaving it to FLDOE to determine how to use the resulting information to improve program design or implementation. Third, the information collected on program implementation may provide useful insights to policymakers outside of Florida who are interested in improving or adopting similar programs.

The **third major goal** was to provide a cost analysis of the FCCRI. Student-level records were used to examine FCCRI costs associated with activities such as administering the PERT to targeted grade 11 students. We also collected administrative and qualitative data on the costs of personnel, facilities, and materials/equipment used to implement the FCCRI at the state, district, and school levels. This information on costs was combined with results from the regression discontinuity analysis on the impact of the FCCRI on student outcomes to conduct a cost-benefit analysis.

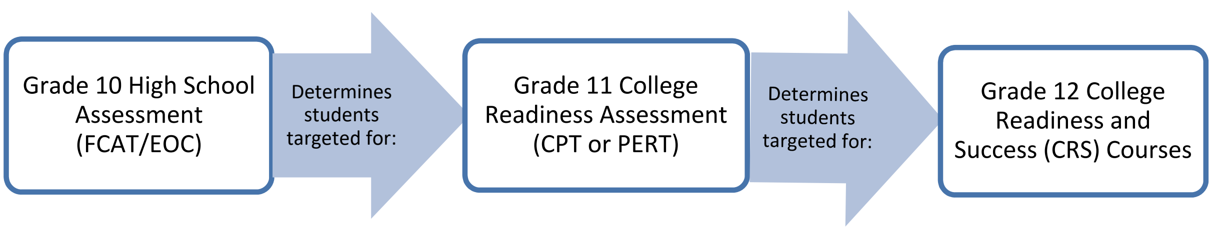
Structure of this report

This report is organized into seven chapters. Chapter 2 describes the policy context for the FCCRI and the underlying theory of action. Chapter 3 describes the FCCRI's implementation and changes that occurred between the voluntary and mandatory periods. Chapter 4 provides the results of the confirmatory impact analysis of outcomes for students near the cutoff values for assignment to the FCCRI using regression discontinuity analysis. Chapter 5 provides the results of the exploratory impact analysis of CRS course outcomes for students from a broader range of academic performance levels using a before-after regression analysis. Chapter 6 provides the results of a district cost-benefit analysis of the FCCRI. Chapter 7 summarizes conclusions from each evaluation component and discusses the policy implications.

Chapter 2: Description of the Florida College and Career Readiness Initiative (FCCRI)

The purpose of the FCCRI is "to improve the college and career readiness rates of high school students and reduce the percent needing postsecondary remediation in reading, writing, and mathematics after graduation" (Florida Department of Education, 2008, p. 1). This chapter begins by describing the state's policy context and the legislation that initiated and modified the FCCRI. Next, we examine each of the intervention's components: the grade 10 high school assessment, which identifies students to be targeted for the grade 11 college readiness assessment; the grade 11 college readiness assessment, which identifies students to be targeted for postsecondary preparatory instruction in grade 12; and the five grade 12 college readiness and success (CRS) courses offered in math, reading, or writing (Figure 1). We conclude with a logic model that illustrates the theory of action behind the FCCRI.

Figure 1. Assessments used to target students for participation in the FCCRI



Policy context

The Florida Department of Education (FLDOE) uses data to identify educational priorities and monitor student progress. The state's K–20 education data warehouse, one of the nation's most comprehensive state longitudinal data systems, can be used to track students from the K–12 education system into postsecondary education and the state's labor market. One of the impetuses of the FCCRI was a series of analytic reports from the mid-2000s indicating that many students were leaving high school unprepared for postsecondary education and that students who entered developmental education in college were unlikely to complete postsecondary credentials (Burdman, 2011). For example, a report from FLDOE (Florida Department of Education, 2005) found that 69 percent of students with a level 3 on the FCAT math (which indicates that the student is performing at grade level) and 25 percent of students with a level 3 on the FCAT reading (also at grade level) tested below the college-ready cut score on the placement test in the corresponding subject upon entrance at a Florida community college. The number of students scoring below the college readiness benchmarks caused particular concern because research had shown that only 52 percent of students who need remediation complete the required developmental education sequence (Florida Office of Program Policy Analysis and Government Accountability, 2007), a rate similar to those found in national studies (Bailey, Jeong, & Cho, 2010).

The state's unified education system facilitated the exploration of potential policy levers to improve high school graduates' college readiness. In Florida, the Division of Florida Colleges and the Division of Public Schools are housed together at the state's Department of Education and both report directly to the Commissioner of Education. This facilitates collaboration between the state's secondary and postsecondary sectors (Burdman, 2011). The state also has many standardized policies across all the postsecondary institutions within the Division of Florida Colleges, such as a common course numbering system and a statewide transfer agreement between two-year and four-year colleges. Unlike many states, Florida has the advantage that all public colleges have been required since 1992 to use a common placement test to determine students' readiness to enter a degree program (Bilsky, 2011). They are also required to use the same set of cut scores on the assessment to determine postsecondary course placement. This policy environment of collaboration and unification set the stage for the adoption of the FCCRI.

Senate Bill 1908: The voluntary FCCRI

Senate Bill 1908 was introduced in Florida's 2008 legislative session. The policy was modeled after smaller-scale collaborations between individual school districts and local community colleges to assess and improve high school students' college readiness (Burdman, 2011). The legislation required all high schools to evaluate students' college readiness before grade 12 in English, reading, and math. The program was targeted toward students in the midrange of performance on the FCAT who intended to enroll in a postsecondary degree program. If students met the minimum test scores as set by the State Board of Education to demonstrate college readiness, they were exempt from developmental education courses if they enrolled in a Florida college within two years of the test. The legislation also required high schools to advise students about their level of preparation and to provide access to college preparatory instruction before high school graduation to any students with scores below the college readiness benchmarks. Courses were to be provided "to the maximum extent practicable," either on the high school campus or online through Florida Virtual Schools. The legislation, signed into law by the governor on June 30, 2008, modified Florida Statute 1008.30, which regulates common college placement testing.

Senate Bill 1908 specified that the State Board of Education would develop or purchase the test used to assess high school students' college readiness. High schools would be required to use this test or an alternative approved by the Board. In 2008 the General Appropriations Act allocated $700,000 for community colleges to administer the college readiness assessment to high school students (Florida Department of Education, 2008). However, no funding was provided to develop courses for postsecondary preparatory instruction or to offer professional development for teachers of these courses (Burdman, 2011).

Student participation in the college readiness assessment and the postsecondary preparatory instruction was voluntary under Senate Bill 1908. The policy was targeted toward students with an intent to enroll in postsecondary education, as determined locally. This intent could be explicitly stated by students (for example, by writing their goals in their educational plan) or implied by their request to take the test (Florida Department of Education, 2008). State policymakers hoped that the potential benefits to students would provide enough of an incentive for students to volunteer to participate. If students did not meet the college readiness benchmarks on the assessment, they could identify their need for additional preparation, adjust course-taking in their senior year, and save money by reducing their need to complete developmental education courses in college (where such courses cost tuition but give no college credit). Students who did meet the college readiness benchmarks on the assessment could begin taking dual enrollment courses for college credit during their senior year and would have a guarantee that they would be exempt from developmental education if they enrolled in a Florida college within two years of taking the test.

House Bill 1255: The mandatory FCCRI

Florida Statute 1008.30 ("Assessment and Accountability") was revised again during the 2011 legislative session with House Bill 1255. The legislation removed the text about targeting the FCCRI toward students with "an interest in postsecondary education." It required high schools to administer the college readiness assessment to all students within a specified range of FCAT scores before they entered grade 12. The legislation also required high schools to provide preparatory postsecondary instruction to all students not meeting college readiness benchmarks on the assessment. Targeted students would be required to complete (but not necessarily pass) a CRS course to meet new high school graduation requirements. The curriculum for the CRS courses would require approval by the State Board of Education and had to align with Florida's new Postsecondary Readiness Competencies (PRCs), which were developed jointly by the Division of Florida Colleges and the Division of Public Schools.

House Bill 1255 also transferred the responsibility of administering the college readiness assessment from the public colleges to the high schools. The state indicated that "the testing responsibilities have been shifted to high schools to better serve students and expedite testing" (Florida Department of Education, 2008, p. 7). There were also concerns that the colleges would not have the capacity to accommodate the large increase in the number of students that would be taking the assessment under the revised policy (Burdman, 2011). New funds were appropriated in 2011 for districts to administer the college placement test to high school students.

House Bill 1769: The return of the voluntary FCCRI

In February 2015, FLDOE released an "Assessment Investigation" report led by Commissioner Pam Stewart. This report collected data from all Florida districts on the standardized assessments administered in each grade level, the reasons students are required to take each test, and how the test results are used. The report concluded with recommendations to reduce statewide testing requirements "in order for Florida to move forward with fewer, better assessments." One of these recommendations included eliminating the PERT as a state mandate and making student participation voluntary again. The rationale provided for this decision was that the PERT was no longer necessary because the state's recent adoption of more rigorous graduation requirements would ensure that students would graduate from high school college- and career-ready.

New legislation signed under House Bill 7069 in April 2015 eliminated the requirements for common placement testing and postsecondary preparatory instruction. Participation in both components became voluntary at the student level (as with the voluntary FCCRI) and also at the school level beginning in 2015/16. These changes occurred after the cohorts for our evaluation had completed high school and do not affect the impact analyses.

FCAT testing in grade 10

Part of the state's accountability system since 1998, the FCAT was initially administered to students in grades 4, 5, and 8, but was later expanded to cover grades 3–11. The FCAT consists of criterion-referenced tests in math, reading, science, and writing and aligns with Florida's Sunshine State Standards Benchmarks. Students receive a scale score ranging from 100 to 500 in each subject. The scale score is also converted into a developmental scale score from 0 to 3,000, which is used to track students' growth over time. Students also receive an achievement level based on their scale scores. The achievement levels range from a low of 1 to a high of 5, with a score of 3 indicating that the student is performing at grade level.[1] The cut scores for each level are set by FLDOE each year and are subject to change.

In 2010/11, Florida began to transition from the FCAT to a new assessment system that includes the FCAT 2.0 and a series of end-of-course (EOC) assessments. The FCAT 2.0 aligns with the Next Generation Sunshine State Standards, which are more rigorous than the prior standards. The scoring for FCAT 2.0 consists of scale scores, which are also referred to as developmental scale scores. Students also still receive an achievement level, which ranges from a low of 1 to a high of 5. Under the new

1. http://www.fldoe.org/faq/default.asp?Dept=179&Cat=0

assessment system, reading is the only subject for FCAT 2.0 for grade 10 students. Florida has also developed a series of EOC assessments that align with the Next Generation Sunshine State Standards but are taken at the end of a specific high school course rather than at a particular grade level. An EOC assessment for Algebra I, first offered in 2010/11 was taken by all middle school and high school students enrolled in a course that met the high school graduation requirement for Algebra I. In 2011/12 the state set standards and achievement levels for the Algebra I EOC assessment, and the test replaced the FCAT math as the high school math test in the state accountability system.

Students must receive a passing score on the FCAT, FCAT 2.0, and/or EOC assessment (the specific tests depend on the year the student entered grade 9) to meet high school graduation requirements. Students who entered grade 9 between 2000/01 and 2009/10 must receive an FCAT math scale score of 300 and either an FCAT reading scale score of 300 or an FCAT 2.0 reading scale score of 241. Changes to high school graduation requirements were phased in over time. For students entering grade 9 in 2009/10, the requirements included an FCAT reading scale score of 327 (an increase from the previous requirement of 300) and an Algebra I EOC score of level 3 or higher. Students in later cohorts were subject to additional graduation requirements that included EOC assessments in other subjects.

All students who scored below a level 3 on FCAT reading or math were required to be provided with additional diagnostic assessments to determine the nature of the student's difficulty, the areas of academic need, and strategies for appropriate intervention and instruction, as described in the student's individualized progress monitoring plan. Students who received a level 1 or 2 on the FCAT math or Algebra I EOC assessment were required to take an intensive or remedial course the following year. Students who received a level 1 on the FCAT reading were required to take an intensive reading course the following year, and students who received a level 2 on the FCAT reading were required to take either an intensive reading course or another content area course that the school determined would meet the students' needs. These intensive and remedial courses were separate from the college preparatory instruction mandated under Florida Statute 1008.30.

Students were allowed to retake the FCAT up to four times in high school. Several alternatives could be used to fulfill the high school graduation requirement, which included earning a comparable score on the ACT or SAT, completing a certificate of completion at a community college, or completing a GED. In addition, students with disabilities could request a waiver from the FCAT and EOC graduation requirements.

Students at the midrange of performance on the FCAT were targeted to take the college readiness assessment in grade 11, but only in the subject area in which they scored below the FCAT threshold for college readiness testing. This means that the only students who took both subjects of the college readiness assessment were those who scored below the FCAT threshold for college readiness testing in both subjects. Under

Senate Bill 1908 (voluntary FCCRI), students were eligible for college readiness testing if they had “scores at Level 2 or Level 3 on the reading portion of the grade 10 FCAT or Level 2, Level 3, or Level 4 on the mathematics assessments” (Florida Statute 1008.30, 2008). However, subsequent guidance from FLDOE recommended that schools initially limit participation to students with an FCAT level of 3 (reading and math) or 4 (math) and then expand the program to students with FCAT level 2 scores if funding was available (Florida Department of Education, 2008). Students scoring below FCAT level 3 were expected to take intensive or remedial courses for their FCAT performance instead of CRS courses. Students in the upper range of FCAT scores (FCAT level 4 or higher in reading and FCAT level 5 in math) were not targeted to take the college readiness assessment in grade 11 because they had already demonstrated high performance and were expected to meet college readiness criteria. A report from FLDOE (Florida Department of Education, 2005) found that over 90 percent of students with an FCAT level 4 in reading or FCAT level 5 in math who attended a Florida state college passed the corresponding subject test on the College Entry-Level Placement Test.

Under House Bill 1255 (mandatory FCCRI), students were required to take the college readiness assessment if they had “scores at Level 2 or Level 3 on the reading portion of the grade 10 FCAT or Level 2, Level 3, or Level 4 on the mathematics assessments” (Florida Statute 1008.30, 2011). This was a change from prior years in that participation became mandatory for students in the specified FCAT ranges, and the target population was expanded to include lower-performing students with FCAT scores of level 2 (which was not contingent upon the availability of funding). After the passage of House Bill 1255, the state transitioned from the FCAT math assessment to the Algebra I EOC assessment. This created a challenge for identifying which students should participate in the college readiness assessment in math, particularly during the first year after the Algebra I EOC assessment was administered. Only students who enrolled in an Algebra I course in grade 10 would have an EOC score to determine whether they were required to take the college readiness assessment in grade 11. Students who took Algebra I in grade 9 or in middle school would have completed the course before the EOC assessment was developed, and students who did not complete Algebra I until after grade 10 would not have EOC scores until after the administration of the college readiness assessment. For this first cohort under the new assessment system, students who did not take Algebra I in grade 10 were not required to take the college readiness assessment in math according to state statute, although districts may have had separate local requirements. For example, some districts test all students in grade 11 unless they are enrolled in an advanced math course such as an Advanced Placement (AP) or International Baccalaureate (IB) course.[2]

2. Based on a conversation during a bimonthly conference call between FLDOE and district PERT coordinators on October 22, 2012.

College readiness assessment in grade 11

During the timeframe for our analyses, State Board of Education Rule 6A-10.0315 identified the college readiness assessment to be used by all public colleges statewide for course placement and establishes the minimum scores for college readiness in entry-level courses (table 1). If college students did not meet these minimum scores, they were required to enroll in developmental education courses prior to completing 12 college credits. High school students who met these state-defined cut scores on the assessment were considered college-ready (Florida Department of Education, 2010). If they enrolled in the Florida College System within two years of taking the test, they did not need to take any developmental education courses.

The primary college readiness assessment used by the state changed after the FCCRI's inception from the Florida College Entry-Level Placement Test (CPT) to the Postsecondary Education Readiness Test (PERT), each of which are described in more detail below. For both the CPT and the PERT, state funding was been available to test students only once in grade 11 (Burdman, 2011). However, State Board of Education Rule 6a allowed students to take the test up to two times during high school.

Table 1. Description of postsecondary course levels for developmental and nondevelopmental courses

	Course Level	Description
Developmental	Lower-level developmental	Courses are reserved for students receiving the lowest scores on the college placement exam, and students taking these courses do not earn credits from them.
	Upper-level developmental	Courses are for students who are below the college readiness cut score but are not in the lowest range of scores. Students earn no credits for these courses.
Nondevelopmental	Transitional (math only)	Courses are designed for students above the developmental level who are still underprepared for college algebra. Students earn elective credits for these courses, but the courses do not satisfy the degree requirement for math. There is no equivalent course in English.
	Degree credit	Course satisfies part of the degree requirement in math or English.

Florida College Entry-Level Placement Test (CPT)

Florida public colleges began using the Florida College Entry-Level Placement Test (CPT) in 1995. The CPT is a version of the College Board's Accuplacer assessment, which is widely used in other colleges across the country (Florida College System, 2012b). The CPT is a computer-adaptive test with sections in reading, writing, and math that are used to determine placement in college courses in the corresponding subject areas. The State Board of Education allows students to substitute scores from other approved assessments, and many students at four-year colleges use SAT or ACT scores instead of the CPT (Bilsky & Tappen, 2010). By July 1997, all colleges were also required to use the same minimum cut scores to determine college readiness, although colleges were permitted to set higher cut scores if they wanted.

When high schools were required to begin offering the CPT after Senate Bill 1908, they had to partner with local colleges to fulfill this responsibility. The state's contract with the College Board allowed only Florida colleges to administer the test. In the early years of the FCCRI, most of the tests were administered by college personnel at the high school campus in paper-and-pencil format because many high schools did not have the equipment or security required for the computer-based CPT (Burdman, 2011). In some districts, the students traveled to a college campus to take the computer-based test. Each district had to develop a plan with a college that included details about the test format (paper-and-pencil or computer-based), test location, test times, student transportation, staffing, and plans for reporting the test results to the high school (Florida Department of Education, 2008). Testing dates varied by district, but the state recommended that students take the CPT in the late fall or early spring of the junior year. This would allow them to continue to develop skills in their junior year before taking the test, while also providing enough time for students to modify their course registrations for the senior year.

Development of a new assessment for college placement

Florida started to develop a new college placement assessment based on recommendations from the Go Higher Florida! Task Force in 2008 (Bilsky, 2011). FLDOE brought together K–20 faculty from across the state to review new benchmarks that were being developed as part of the state's involvement in Achieve's American Diploma Project (ADP). They also identified college readiness competencies within the ADP benchmarks and identified gaps in K–12 preparation. This was used to develop new Postsecondary Readiness Competencies (PRCs), which were then aligned with the K–12 Sunshine State Standards. In 2009 the faculty met again to develop exemplar test items for each PRC. These items were sent to test developers, who were invited to send sample test questions for a new college readiness assessment. The faculty reviewed the proposals and selected the one that they felt most closely aligned with the PRCs. In January 2010, McCann Associates was awarded a contract to develop a new assessment.

In July 2010, Florida adopted the Common Core State Standards in English Language Arts and mathematics. The faculty met again to review these standards and revise the PRCs. The revised PRCs were used to develop the new test items for the PERT. A pilot test of the PERT was administered to over 10,000 students at Florida Colleges, and the results from the pilot were used to create the PERT item bank. In August 2010, faculty reviewed all items for quality and alignment with the PRCs. Since that time, K–20 faculty have continued to review all new test items on the PERT before they are used.

Florida colleges began transitioning to the PERT as the primary assessment for college placement in October 2010. All colleges were expected to use it as the primary placement test by fall 2011 (Burdman, 2011). Public high schools also made the transition to using the PERT as the college readiness assessment for grade 11 students during the 2011/12 school year.

Postsecondary Education Readiness Test (PERT)

The PERT comprises three subtests in reading, writing, and math, with 30 items on each section (Florida College System, 2012b). It is a computer-adaptive test that starts with a question at an average ability level and then adjusts the difficulty of subsequent questions based on the accuracy of the responses to other questions in the same competency. Students are not allowed to bring their own calculators, but an on-screen calculator is available for some of the questions on the math subtest.

During the timeframe for our analyses, the PERT was used to determine postsecondary course placement. Students could not “fail” the test, but if they did not meet the state-established cut score, they had to complete developmental coursework before they could enroll in college-level courses. The scale scores ranged from 50 to 150 in each subject area, and the college-ready cut scores were 113 in math, 104 in reading, and 99 in writing.[3] The initial cut scores were developed by trying to match the distribution of scores on the CPT with the distribution of scores on the pilot of the PERT. PERT scores were available to students and district administrators as soon as the computer-adaptive test was completed.

Each district decided when it would offer the PERT to eligible students, but enough time had to be allowed for students to be placed into the appropriate courses in grade 12 (Florida College System, 2012b). Schools were required to provide Internet-connected computers for students to complete the computer-adaptive test, a secure and monitored location for students to take the test, and authorized test administrators to monitor the test-taking. The PERT could be administered during regular school hours or at another time. The test was untimed, but the average completion time was 37 minutes in math, 33 minutes for writing, and 1 hour and 3 minutes for reading on the computer-based tests (Florida Department of Education, 2012). The length of the test created a challenge during the initial phases of implementation because the test took much longer than anticipated. The PERT took

3. According to Florida State Board of Education Rule 6a, equivalent college readiness scores for the CPT, SAT-I, and ACT can be substituted for the PERT. If students take more than one of these tests, the highest test score from each test will be accepted.

nearly twice as long to administer as the CPT, which could be completed in 90 minutes (Burdman, 2011).

Although almost all high schools administered the PERT under the FCCRI, districts could request permission from FLDOE to use their grant money to purchase an alternative college readiness assessment. In 2011/12 three districts used alternate assessments. The Florida School for the Deaf and Blind used the ACT because there were no accommodations for the PERT in the first year of testing. Two additional districts, Brevard Public Schools and Putman County School District, also opted to use the ACT instead of the PERT to assess college readiness (personal communication with Sandy Dilger, FLDOE).

CRS courses in grade 12

The purpose of providing CRS courses in grade 12 was to "prepare students for entry-level college credit courses as well as gainful employment" (Stewart, 2011, p.1). Some high schools began offering these courses in 2008/09, but other schools needed additional time to prepare. All high schools were expected to offer courses by 2009/10. Students who scored below the college-ready cut scores on the postsecondary readiness assessment were advised to enroll in these courses in 2008/09 to 2011/12, but their participation was voluntary. Beginning in 2012/13, all students not meeting the college-ready cut scores were required to enroll in a CRS course in the corresponding subject area. If students scored above college-ready on the postsecondary readiness assessment, they might still have been advised to take the CRS courses, or they might have been advised to take an even more challenging dual enrollment course to further improve their readiness for college-level work. Course-taking was thus likely to be modified somewhat in the senior year regardless of student performance on the college readiness assessment.

In 2008, the state approved standardized course codes for two postsecondary readiness courses in math: Math for College Success (state course code 1200410) and Math for College Readiness (state course code 1200700). Math for College Success was a half-credit course designed for students at the lowest levels of preparation. The Math for College Readiness Course was a full-credit course designed for students scoring at or slightly above the college-ready cut scores. It aligned with a college elective credit (Intermediate Algebra) and was not categorized as developmental education. The state provided a list of standards to be covered in each course but did not provide any curriculum. High schools were advised to use curriculum from their local community colleges (Florida College System, 2012b).

In 2009, the state approved standardized course codes for a writing course (Writing for College Success—state course code 1009370) and two reading courses (Reading for College Success—state course code 1008350, and English 4: Florida College Prep—state course code 1001405) (Bilsky & Tappen, 2009). The two "Success" courses both counted as a half credit and were designed for students at the lowest levels of preparation. The English 4 course was a full-credit course designed for students scoring at or slightly above the college-ready cut scores in reading or writing. As with the math CRS courses, the state provided standards to be covered in each course but did not provide any curriculum.

The standards for all courses were revised in 2011 to align with the state's Postsecondary Readiness Competencies and Common Core State Standards, and schools were expected to offer revised courses that cover these standards no later than 2012/13 (Stewart, 2011). Schools were required to offer at least one course in each subject, and the courses were supposed to be taught by a teacher certified in the course subject area. All five of the courses counted toward elective credits at the high school and fulfilled the requirement for postsecondary preparatory instruction for students scoring below college-ready on the CPT or PERT. Math for College Readiness and English 4 also counted as a full credit toward the subject area high school graduation requirement (Florida College System, 2012b). Since these courses counted toward the high school graduation requirement, all students were eligible to enroll, regardless of whether they were targeted for participation under the FCCRI.

Students in voluntary cohort 1 through mandatory cohort 1 of the FCCRI who completed CRS courses in high school were required to retake the CPT or PERT and meet the college-ready cut scores to enroll in credit-bearing college courses (Florida College System, 2012b). This testing could occur any time prior to college matriculation; some high schools offered students a retest of the CPT or PERT at the end of the CRS courses, but they were not required to do so.

However, a 2013 postsecondary policy change affecting college placement went into effect during the first year of college enrollment for the second mandatory cohort of the FCCRI. Beginning in 2014/15, with enactment of Senate Bill 1720, colleges are no longer allowed to require recent high school graduates to take placement exams for college entry or enroll in noncredit developmental education courses, even if their high school PERT scores are below college-ready. (The legislation allows colleges to require nontraditional students to take placement exams and enroll in developmental education.) Colleges are also required to implement alternative developmental education strategies, as well as plans for comprehensive advising programs to inform students about these voluntary options. These changes were driven by a legislative response to perceptions about a lack of student success in conventional developmental programs at state colleges.

Summary of changes

Figure 2 provides a timeline that summarizes activities affecting the FCCRI. Table 2 provides a timeline for each year from 2008/09 to 2012/13 for voluntary cohort 1 (V1) through mandatory cohort 2 (M2), the cohorts included in our impact evaluation. This shows when various changes to the FCCRI went into effect for students in the analytic sample.

Figure 2. Timeline summarizing activities affecting the FCCRI

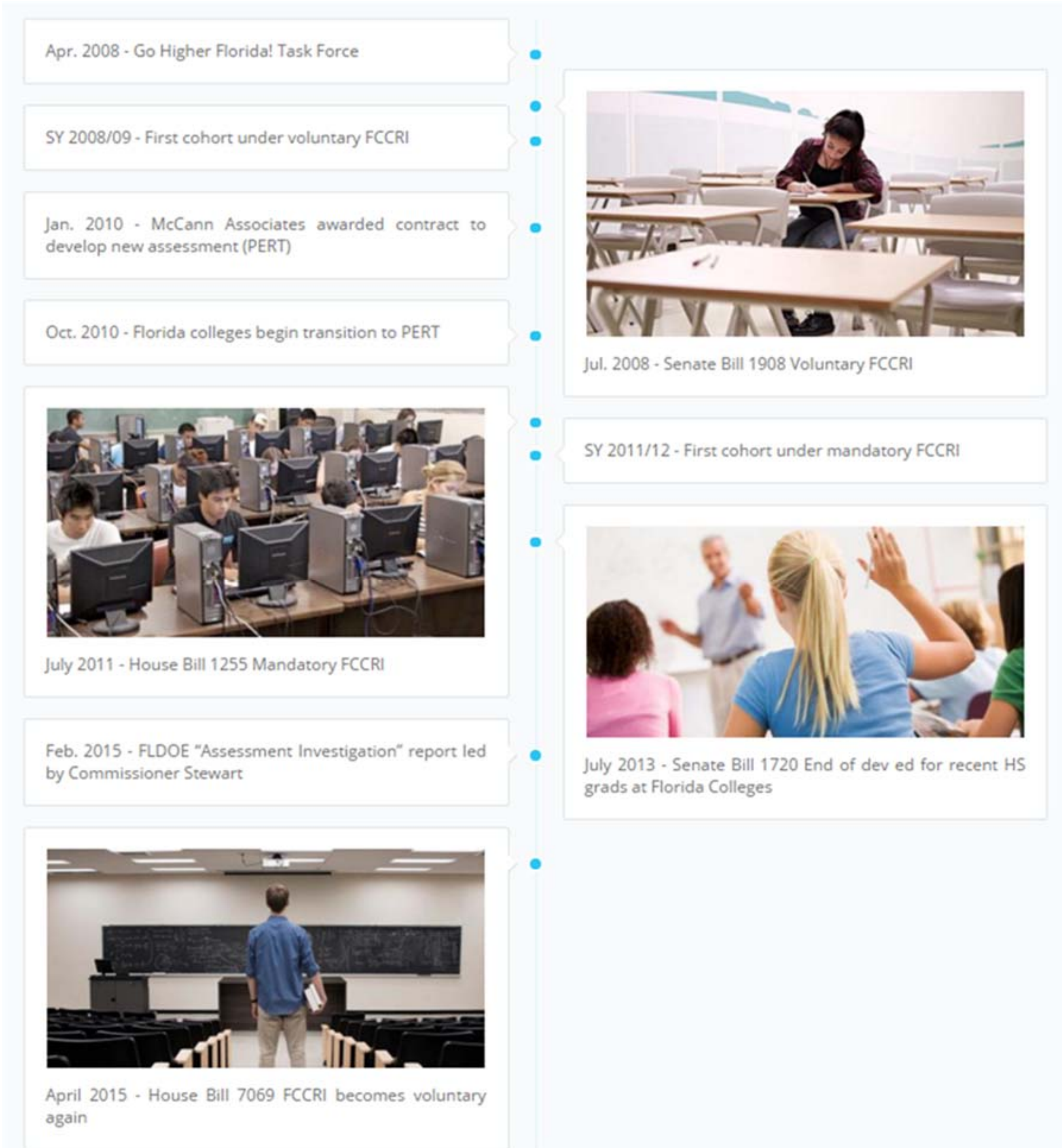


Table 2. Timeline for cohort V1 through cohort M2

	Voluntary Cohort1 (V1)	Voluntary Cohort 2 (V2)	Voluntary Cohort 3 (V3)	Mandatory Cohort 1 (M1)	Mandatory Cohort 2 (M2)
2007/08	Grade 10: FCAT Math and Reading				
2008/09	Grade 11: CPT (Voluntary)	Grade 10: FCAT Math and Reading			
2009/10	Grade 12: CRS courses (Voluntary)	Grade 11: CPT (Voluntary)	Grade 10: FCAT Math and Reading		
2010/11	College: Year 1	Grade 12: CRS courses (Voluntary)	Grade 11: CPT (Voluntary)	Grade 10: FCAT Math and Reading	
2011/12	College: Year 2	College: Year 1	Grade 12: CRS courses (Voluntary)	Grade 11: PERT (Mandatory)	Grade 10: Algebra I EOC, FCAT Reading
2012/13	College: Year 3	College: Year 2	College: Year 1	Grade 12: CRS courses (Mandatory)	Grade 11: PERT (Mandatory)
2013/14	College: Year 4	College: Year 3	College: Year 2	College: Year 1	Grade 12: CRS courses (Mandatory)
2014/15	College: Year 5	College: Year 4	College: Year 3	College: Year 2	College: Year 1 (developmental education became optional)

Theory of action

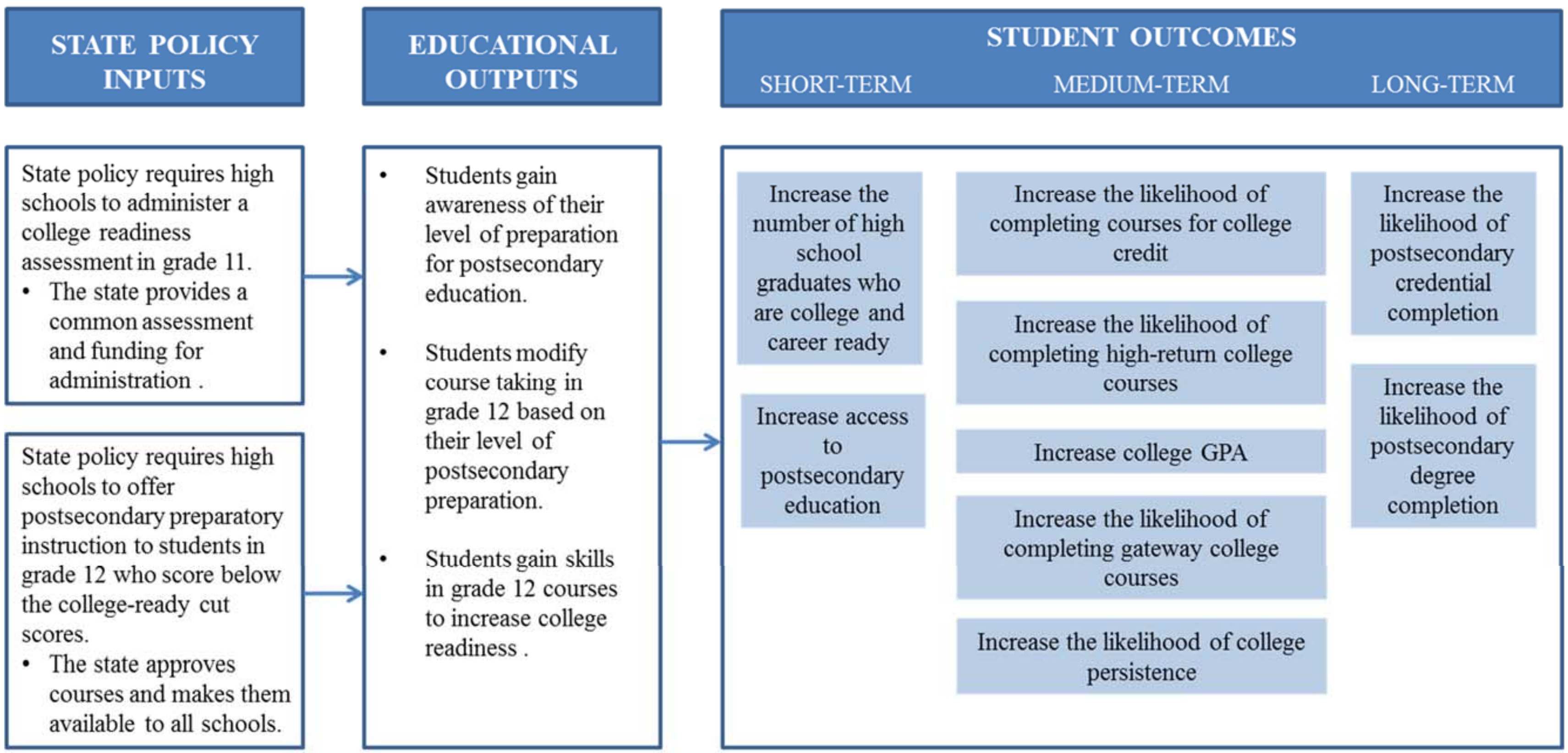
The theory of action for the FCCRI is that providing college readiness testing and postsecondary preparatory instruction in high school may raise students' awareness of their academic deficiencies and motivate them to acquire college-level skills in their senior year. The logic model shown in Figure 3 illustrates the FCCRI's key components as it is intended to be implemented, the mediators (or educational outputs) through which the intervention is designed to work, and the short-, medium-, and long-term outcomes the program is expected to affect.

College readiness assessment in high school

The first issue that the FCCRI seeks to address is students' lack of awareness of the skills needed to succeed in college. Approximately 90 percent of students in grade 9 expect to complete some form of education after high school (Kirst, 2005; National Commission on the High School Senior Year, 2001); however, for some students there is a disconnect between wanting to go to college and being prepared for college. Most students believe that earning a high school diploma demonstrates that they are ready for college. A study from Strong American Schools (2008) found that of the students required to take developmental education classes in college, 80 percent believed that they were college-ready and were angry, surprised, and embarrassed that they were unable to enroll in credit-bearing courses. Ninety-five percent of these students reported that they did all or most of the work required of them in high school. Students in low-performing schools (Boser & Burd, 2009) and first-generation college attendees (National Commission on the High School Senior Year, 2001) are particularly unaware of the gap between the knowledge and skills required to graduate from high school and the knowledge and skills required for college.

This disconnect between wanting to attend college and being college-ready is furthered by high school assessments. As a result of the No Child Left Behind Act, all states have assessments (usually in grades 9 or 10 at the high school level) to test student achievement in English, math, and science. Most students believe that passing these tests indicates that they are ready for college (Conley, 2007). Yet high school curriculum, standards, and assessments are largely unaligned with college entrance and placement exams (Kirst, 2005; National Commission on the High School Senior Year, 2001). Many high schools place such an emphasis on students passing the state test that high school teachers align their curriculum with the test content rather than postsecondary academic standards (Conley, 2007). High school teachers often do not realize that their content is unaligned. For example, a study by McCormick and Lucas (2011) found that the majority of high school math teachers believe that they are meeting the appropriate state standards to prepare students for college, whereas the majority of college professors believe that students come to college unprepared to take college-level math courses.

Figure 3. Logic model for the theory of action of the FCCRI



There are problems both with the lack of information that students receive about the necessity of testing college-ready and with when the tests are required to be taken. High schools and colleges often fail to communicate to students that they must first take a placement test to determine if they are eligible to enroll in credit-bearing courses once they enroll in college (Boser & Burd, 2009; Kirst, 2001; Rosenbaum, 2001). Further, waiting to assess students' college readiness until after they leave high school sends a signal to students that achievement in high school is not important for attending college (Rosenbaum, 2001). One way to help students understand their lack of college preparedness is through greater use of college readiness assessments in high school. The Institute of Education Sciences (IES) practice guide *Helping Students Navigate the Path to College: What High Schools Can Do* (Tierney, Bailey, Constantine, Finkelstein, & Hurd, 2009) recommends using assessments to inform students about their level of college preparedness and to help them overcome their academic deficiencies.

The logic model shows that the first main policy input for the FCCRI is a state policy that requires high schools to administer a college readiness assessment in grade 11. To implement this policy, the state provides a common assessment and funding for the administration of the assessment. This component of the intervention is designed to increase students' awareness of their level of postsecondary preparation and to allow students to modify course-taking in grade 12 based on this level. This increase in awareness of college readiness is hypothesized to improve student outcomes in postsecondary education. In addition, the early assessment could motivate more students to go to college by providing a signal of college-readiness. This could occur if students did not initially intend to go to college because they thought they were unprepared, but then received a college-ready score in grade 11.There is some evidence that district- or statewide use of assessments that predict college readiness (such as the ACT) in high school is associated with improved college outcomes, such as higher college enrollment rates (Hurwitz, Smith, Niu, & Howell, 2015; Hyman, 2017) and a reduction in the need for developmental education courses (Kurlaender, Jackson, & Howell, 2016).

Postsecondary preparatory instruction

The second issue that the FCCRI seeks to address is the rigor of course-taking in the senior year. The majority of students attending postsecondary education are enrolled in open-access institutions that allow any students with a high school diploma to enroll, regardless of high school performance. Yet it is generally accepted that many students who receive high school diplomas do not have the requisite skills to successfully complete college-level work (for example, Achieve and The Education Trust, 2008; Boser & Burd, 2009; Strong American Schools, 2008). Students do not understand that completing high school requirements and being admitted to a postsecondary institution do not indicate college readiness. As a result, about 70 percent of students go on to a postsecondary institution, but the percentage completing a bachelor's degree has not risen much above what it was in 1950 (Kirst, 2001).

Although many high schools offer opportunities for students to go beyond the high school graduation requirements by taking advanced courses, students often fail to take advantage of these opportunities. Kirst (2001) argues that students do not understand the influence that advanced high school courses can have on postsecondary degree attainment. Because the college admissions process begins in the first semester of the senior year, most colleges do not take into account grades earned during students' last year in high school. By the second semester of their senior year, students know where they will attend college and the remainder of the school year is perceived as a time to relax and have fun. Kirst (2001) refers to this as the “senior slump” and argues that there is a disconnect between K–12 and postsecondary education that disincentivizes students to work hard during their senior year to academically prepare for college. Rosenbaum (2001) also finds that students who plan to attend colleges with open admissions policies put less effort into their high school work because they do not understand the relevance of their effort to their plans. In addition to increasing motivation, the CRS courses may also help to improve the alignment between the content of grade 12 courses and first year college courses.

Although there is limited research on the effect of providing postsecondary preparatory instruction in high school, several studies have examined more broadly the influence of high school courses on postsecondary outcomes and found that high school curriculum influences college persistence and postsecondary degree attainment. Students who participate in rigorous high school classes and attend college are more likely than their counterparts to persist at a postsecondary institution, to remain enrolled at their initial institution, and to pursue a bachelor's degree track if they transfer institutions (Attewell & Domina, 2008; Horn & Kojaku, 2001; Jacobson & Mokher, 2009). There is also a strong correlation between high school curriculum and bachelor’s degree completion; students who take more rigorous high school courses are more likely to complete a bachelor's degree. Furthermore, the correlation of high school curriculum with postsecondary degree attainment is higher than that of high school test scores or class rank among college enrollees (Adelman, 1999, 2006). These findings suggest that policy initiatives such as the FCCRI, which are designed to improve the rigor of high school course-taking, may have a positive effect on students’ postsecondary outcomes. Courses that students would otherwise be taking may only fulfill the minimal requirements for high school graduation. Moreover, postsecondary preparatory instruction may be even more influential than other advanced high school courses, which may not cover the specific skills needed for college.

The second major component of the FCCRI seeks to ensure that all students have access to courses that will prepare them with the skills needed for college by requiring all high schools to offer postsecondary preparatory instruction in reading, writing, and math to grade 12 students who score below the college-ready cut scores on the college readiness assessment. The state approves high school courses that align with both the K–12 state standards and the PRCs from the Division of Florida Colleges. These courses are made available to all public high schools statewide.

In the absence of statewide policies to make courses available to all high schools, there may be disparities in access to and participation in rigorous courses for some student subgroups. Nationwide, Latino/a students and students from the lowest socioeconomic status are less likely to attend high schools that offer advanced math classes (Adelman, 2006). Black and Latino/a students who do have access to higher level classes are much less likely than White and Asian/Pacific Islander students to take rigorous high school courses (Horn & Kojaku, 2001).

Further, students in rural schools often do not have the same level of access to rigorous courses as do students in urban or suburban schools (Anderson & Chang, 2011; Levin, 2007). Rural schools tend to have smaller enrollments (Jimerson, 2006), which makes it difficult to provide courses for interested students. Rural schools also typically have limited financial resources to allocate to rigorous courses given the lower property tax base and associated lower per-pupil expenditure (Johnson & Strange, 2007; Picciano & Seaman, 2009). This means that some students may lack the opportunity to further improve their preparation for college while they are still enrolled in high school.

The intent of providing postsecondary preparatory instruction in all high schools is to allow students to modify course-taking in grade 12 based on their level of postsecondary preparation and to help students gain skills in their grade 12 courses to increase college readiness. This is hypothesized to influence short-term outcomes, including the number of high school graduates who are college- and career-ready; medium-term outcomes, such as the likelihood of completing courses for college credit; and long-term outcomes, such as the likelihood of postsecondary credential or degree completion.

Chapter 3: Implementation of the FCCRI

This section defines the roles of schools, districts, and the state in implementing the FCCRI; summarizes student- and school-level participation in the college readiness testing and CRS course-taking components of the FCCRI; and describes differences in the treatment and counterfactual conditions.

The qualitative data on implementation comes from an extensive feedback analysis conducted by the research team during the first two years of the evaluation, which is described in greater detail in our previous reports (Mokher & Jacobson, 2014; Mokher, Jacobson, Rosenbaum, & Lalonde, 2013). In the first year, we administered a statewide survey of CRS teachers using a stratified sample of schools (N=225 respondents). We also conducted small-group discussions with CRS teachers in eight cities to collect more detailed feedback (N=60 participants). In the second year of data collection, we resurveyed CRS teachers who participated in the year 1 survey to see whether their perceptions changed as the initiative matured. We also wanted to get feedback from a broader range of stakeholders beyond CRS teachers, so we conducted site visits at high schools, district offices, and state colleges in six counties, which included interviews about the FCCRI's implementation and perceptions of its effectiveness (N=80 interviews).

The quantitative data on implementation comes from student-level records from the Florida K–20 Education Data Warehouse (EDW), which follow all Florida public school students from grade 10 through postsecondary education as long as they remain in Florida and attend a public college or university. The sample includes students enrolled in 474 Florida public high schools. These high schools serve students pursuing standard diplomas, are run by one of Florida's 67 districts or chartered by these districts, and are covered by Florida Department of Education's standardized rating and data reporting system. Schools that are included in the FLDOE rating system serve at least 30 students who are enrolled for the entire school year and have valid FCAT scores in reading and math in the current and previous year. Since the FCAT was used to identify which students were targeted to take the college placement test in grade 11, we included in our sample high schools that received a rating and excluded schools with an insufficient number of FCAT scores. We compared changes over time in participation in the college readiness test and CRS course components of the FCCRI as the initiative shifted from voluntary to mandatory participation for targeted students. We also examined differences in grade 12 course-taking by college-ready status before and after the mandatory FCCRI.

School, district, and state roles in implementing the FCCRI

The FCCRI was a complex initiative involving a range of educational stakeholders at the state and local levels. Each organization played a specific role in FCCRI implementation, sharing information and interacting, when appropriate, with organizations at other levels.

FLDOE provided common high school assessments through the FCAT and EOC tests. It also granted funding to administer the college readiness test to students with targeted scores on these high school assessments. Additionally, FLDOE set standards for CRS courses and communicated these requirements to district-level administrators, who in turn were responsible for communicating FCCRI-related information from FLDOE to high schools.

Districts had autonomy in making curricular decisions, providing textbooks and other instructional materials to use in CRS courses, and organizing professional development for teachers. They also were responsible for providing guidance and oversight of PERT administrations and for reporting test score results to the state.

High school administrators directly implemented PERT testing and CRS courses, the key components of the FCCRI. They also provided guidance to CRS teachers. CRS teachers were largely left on their own, however, to develop lesson plans and select which texts and materials to use in their classes.

State colleges administered the CPT assessment in grade 11 during the voluntary FCCRI, but the responsibility for college readiness testing shifted to high schools with the transition to the PERT assessment. Although state colleges had no direct role in implementing CRS courses, they had extensive experience in crafting developmental education programs with very similar goals as those of the CRS courses. In contrast to the high schools, state colleges had developmental education departments that provided detailed lesson plans for their instructors; a wide range of teaching materials; and extensive support services such as tutors, computer labs, and separate study skills courses. There was some informal collaboration around CRS courses between some high schools and state colleges, but it tended to be limited to a small number of participants and short in duration.

Participation in college readiness testing

Students scoring in the midrange of the FCAT were identified to take the college readiness test in grade 11. Each year from 2008/09 to 2011/12 (corresponding to V1 through M1), approximately 84 percent of students were targeted in math and 57 percent were targeted in reading; the majority of the student population were thus affected by this initiative. Our sample includes 83,603 targeted students in reading and 125,054 targeted students in math in M1 (Figure 4); sample sizes are similar for the other cohorts, with one exception. Due to the change to the Algebra I EOC assessment for cohort M2, only 15,322 students were tested and 10,574 (69 percent) received a score in the targeted range.

Figure 4. Distribution of FCAT levels for cohort M1 (targeted levels shaded in dark blue)

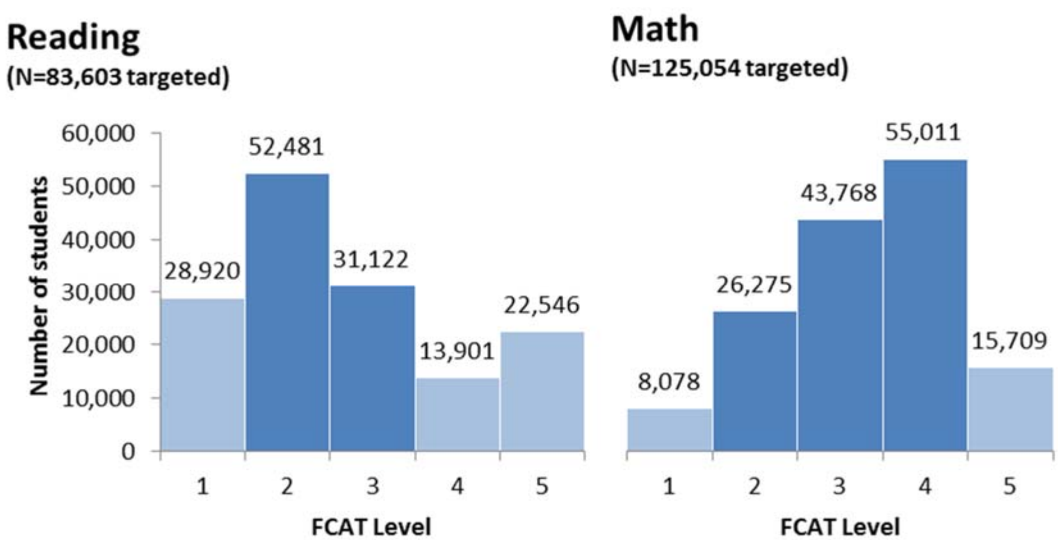
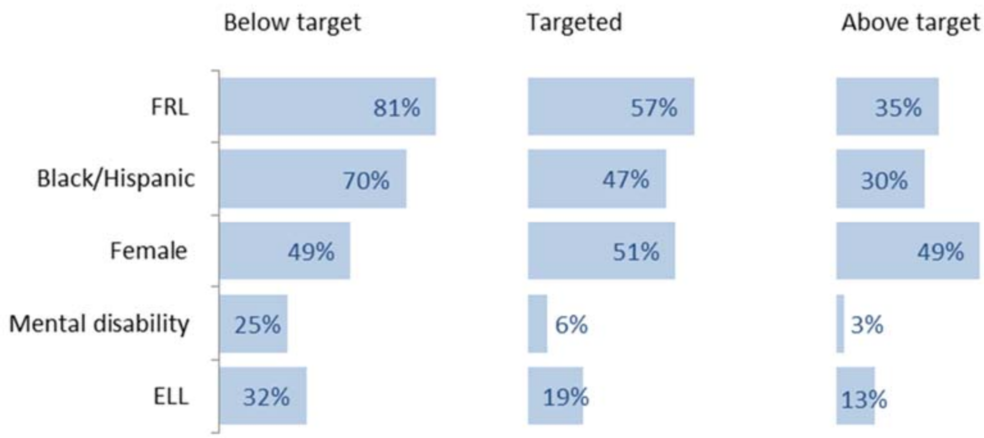


Figure 5 compares the characteristics of targeted students with those of students below and above target. The percentage of students classified as free and reduced-price lunch (FRPL) status, Black/Hispanic, mental disability, and ELL were highest in the below target group and lowest in the above target group, with targeted students in the middle of the distribution. All three groups were similarly distributed by gender, with about half of all students classified as female. We also examined student characteristics by subject area, and found that all characteristics for targeted students were similar (within 2 percentage points) in math and reading.

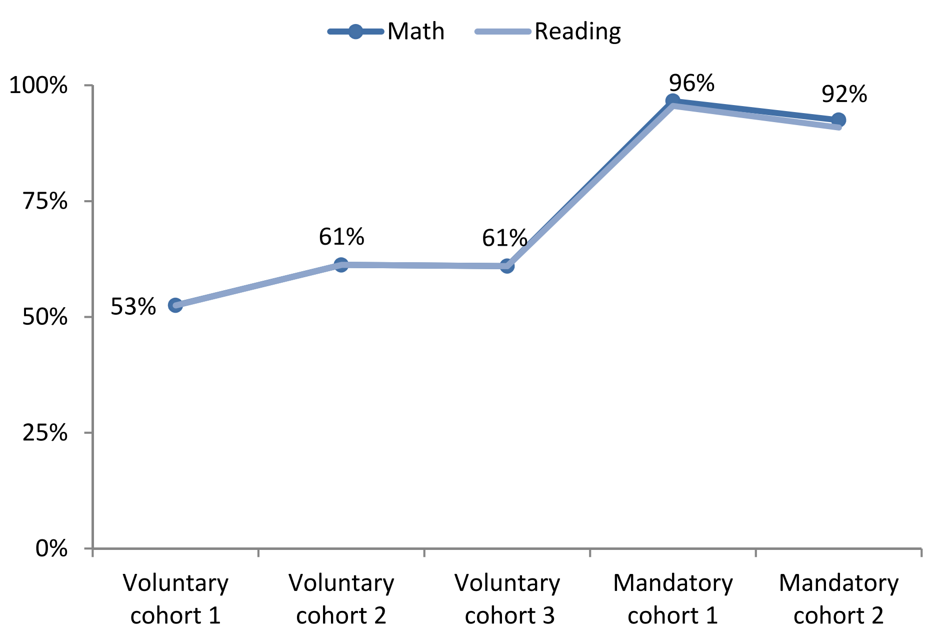
In the early years of the initiative, many high schools did not offer college readiness testing even though they were supposed to. In the first voluntary cohort, only about half of schools (53 percent) offered college readiness testing. This rose to 61 percent in the next two years of the voluntary program, and then increased sharply to over 90 percent once the FCCRI became mandatory (Figure 6). However, participation was still less than 100 percent because even in the mandatory period there was no enforcement or sanctions for schools that did not comply with the requirements of the FCCRI.

Figure 5. Characteristics of FCAT targeted students compared with those of students below target and above target



NOTE: Data include students from cohort V3 and cohort M1 (N=672,316). The targeted group includes students who were targeted in math and/or reading based on their grade 10 FCAT scores.

Figure 6. Percentage of schools offering college readiness testing, by cohort



There are few differences between schools that first participated in college readiness testing under the voluntary FCCRI and schools that first participated under the mandatory FCCRI (Table 3). Voluntary adopters had a lower percentage of Black and Hispanic students than mandatory adopters (43.5 percent versus 56.6 percent). Yet other student demographic characteristics, the percentage of students targeted under the FCCRI, and school locale (urban, suburban, rural) were all within 3 percentage points for schools in the two groups. Conversations with FLDOE staff revealed that schools with existing relationships with local colleges were most likely to participate in college readiness testing under the voluntary FCCRI because high schools needed a college to partner with to administer the CPT.

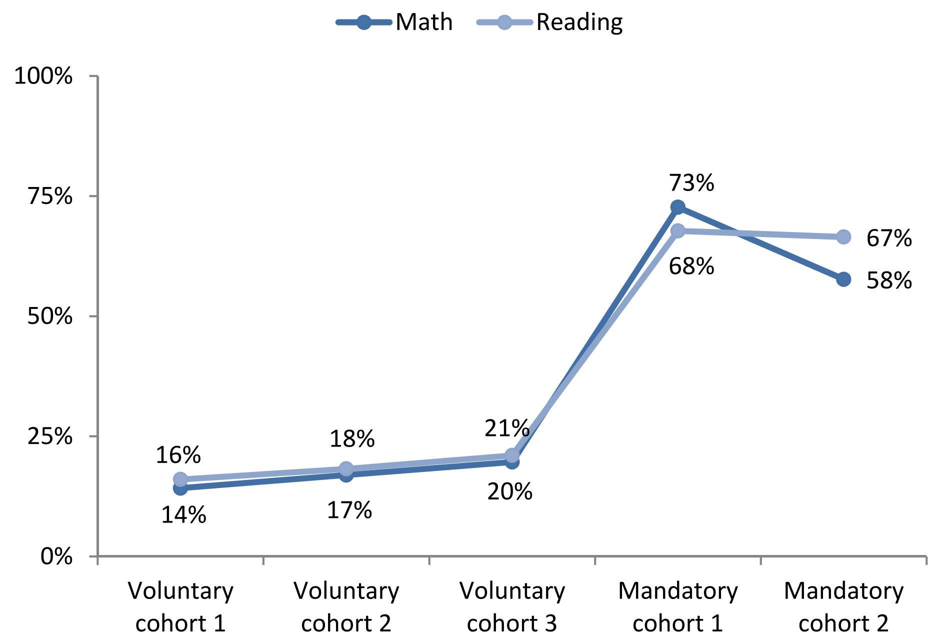
Table 3. Characteristics of schools that first adopted testing under the voluntary FCCRI versus those of schools that first adopted under the mandatory FCCRI

	Voluntary adopters	Mandatory adopters	Difference
Percent of students targeted...			
Math	83.9%	83.7%	0.2%
Reading	55.0%	54.3%	0.7%
Student demographics			
Percent FRPL	49.7%	51.8%	-2.1%
Percent Black/Hispanic	43.5%	56.6%	-13.0%
Percent female	49.8%	51.8%	-2.0%
Percent mental disability	11.6%	11.0%	0.6%
Percent ELL	5.5%	8.1%	-2.7%
School locale			
Urban	23.7%	26.4%	-2.6%
Suburban	48.8%	48.4%	0.4%
Rural	27.5%	25.3%	2.2%

NOTE: There were 374 schools that ever participated in testing under the voluntary FCCRI and 91 schools that first participated in testing under the mandatory FCCRI. Eight schools did not participate in testing under the voluntary or mandatory FCCRI (not shown here).

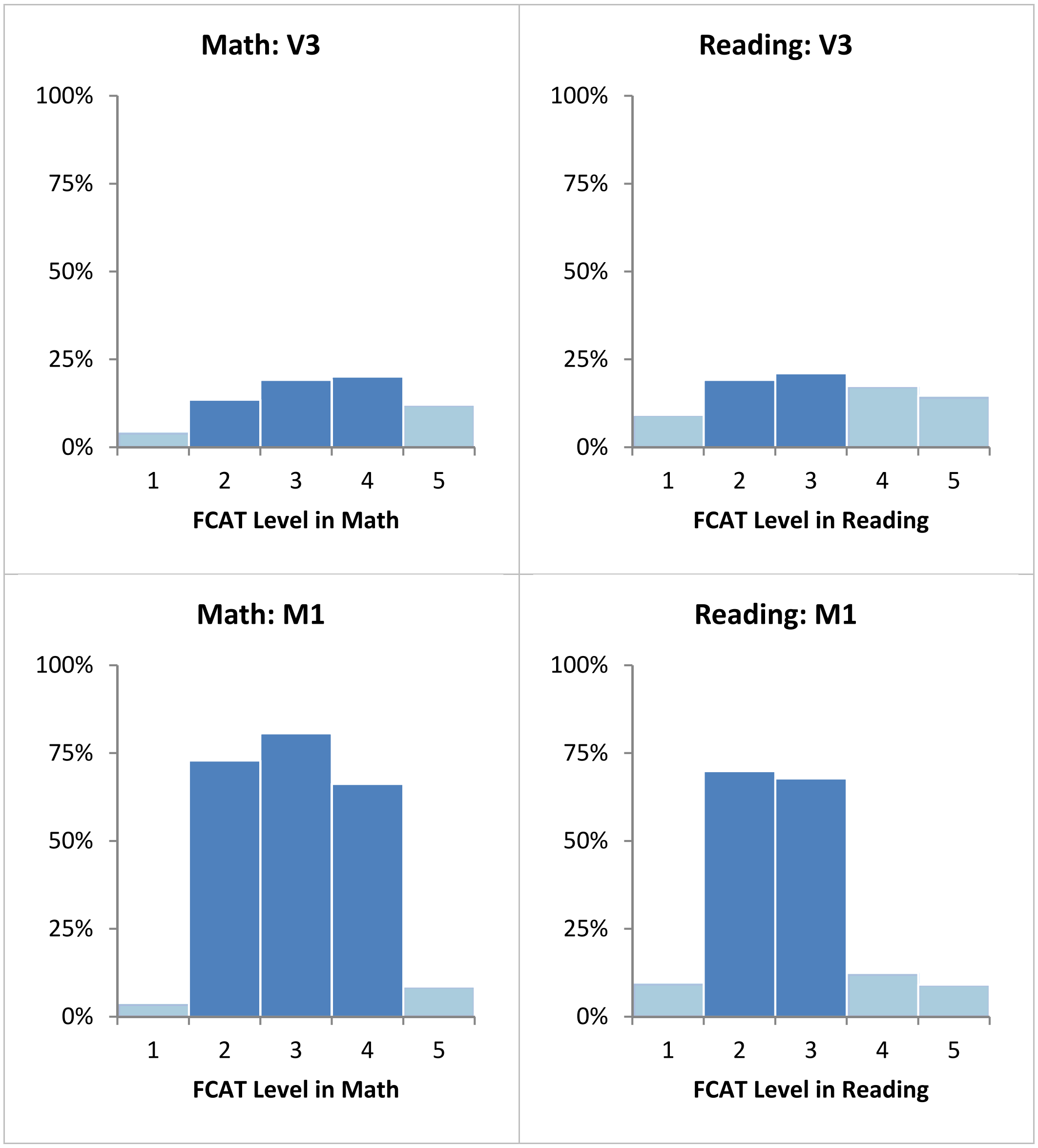
Only about 15 percent of targeted students participated in the college readiness testing during voluntary cohort 1 (Figure 7). This participation rate increased slightly during each year of the voluntary program, reaching about 20 percent by voluntary cohort 3. When the program became mandatory, student-level participation rates more than tripled to 68 percent in reading and 73 percent in math. However, this also means that about one-quarter to one-third of students who should have been tested under the mandatory FCCRI did not participate. While reading participation rates remained similar from M1 to M2, there was a decline in math participation rates which may have been due to confusion about which students were required to be tested after the switch from the FCAT math to the Algebra I EOC.

Figure 7. Student-level participation rates in college readiness testing among FCAT targeted students



Although FCCRI policy specified that only students who achieved mid-level FCAT scores in reading and math were required to take a college readiness assessment, some students across all five FCAT levels were tested during the voluntary and mandatory FCCRI (Figure 8). When the program was voluntary, between 14 and 20 percent of students targeted to take the math test based on their FCAT scores (levels 2, 3, and 4) and about 20 percent of students targeted to take the reading test (levels 2 and 3) did. In the first mandatory cohort, participation rates in the college readiness testing for targeted students ranged from 66 percent (FCAT 4) to 81 percent (FCAT 3) in math, and 68 percent (FCAT 3) to 70 percent (FCAT 2) in reading. In M2 after the change in the grade 10 math assessment, the majority of targeted students still took the college readiness test, but participation rates dropped slightly with a range of 57 percent (FCAT 4) to 63 percent (FCAT 2). In addition, during both the voluntary and mandatory periods about 5 to 10 percent of students whose FCAT scores didn't require it were tested anyway, even though they were not targeted for participation under FCCRI policy. The reason for this is likely a combination of some students choosing to take the test voluntarily, and other students being required to take the test because school staff were uncertain about which students needed to participate.

Figure 8. Participation rates in college readiness testing for targeted students in math and reading for cohort V3 and cohort M1, by FCAT level (targeted levels shaded in dark blue)



In our small-group discussions, teachers explained why their high schools deviated from the testing criteria during the mandatory program, which may also explain variation during the voluntary program. In four districts, some schools administered the PERT to all grade 11 students, regardless of their FCAT achievement. Other high schools in three districts exempted the highest and/or lowest FCAT performers from taking the PERT, although the FCAT cutoff could differ from that defined in Florida

Statute 1008.30. For example, at one high school, all students who did not achieve an FCAT level 5 in reading took the PERT. In contrast, at another high school, the PERT was administered "to all juniors who intend to go into college prep, dual enrollment, or AP (Advanced Placement)," so many of its highest-performing students were tested. Another high school offered the PERT to students who had not passed the FCAT so they could try to obtain a concordance score to fulfill the high school graduation requirement. Additionally, respondents from several districts specified that receiving a concordance score on the ACT/SAT exempted students from taking the PERT.

Retesting students in CRS courses

The FCCRI legislation did not require students who scored below college-ready on the PERT in grade 11 to retake the PERT in grade 12, but many students did so. Policies at the classroom, school, or district level affected whether students retook the PERT.

At the district level, Flagler, Gulf, and Hamilton Counties reported using the PERT to count toward a certain percentage of final course grades in CRS courses (Florida Department of Education, 2015). At the school level, we found that some schools administered a retest of the PERT to seniors enrolled in CRS courses just prior to the end of the school year. Certain schools required seniors to retake the PERT multiple times until they passed it (Mokher, Jacobson, Rosenbaum, & LaLonde, 2013).

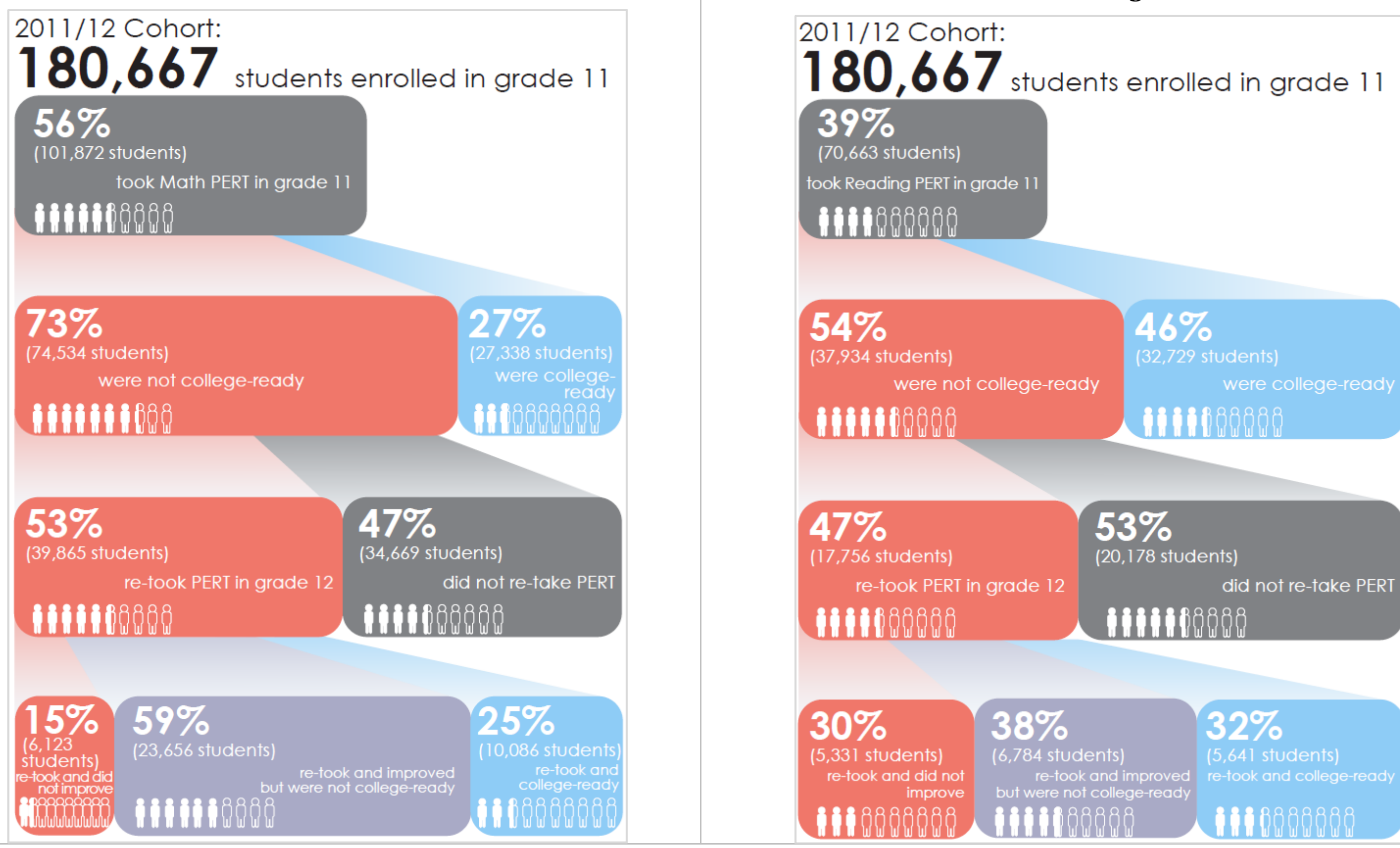
Similarly, site visits and interviews we conducted found instances of teachers using the PERT as part of final course grades in their CRS courses, but choosing to do so on their own, rather than as part of a school or district policy (Mokher & Jacobson, 2014; Mokher, Jacobson, Rosenbaum, & LaLonde, 2013).

Only about half of students who did not score college-ready in grade 11 retook the PERT in grade 12 in the first mandatory cohort, so results on PERT retake performance should be interpreted with caution as not representing all grade 12 students. Students who retook the PERT tended to have lower academic performance, so it may have been more challenging for them to score college-ready than for nonretakers (Lee, 2016).

In grade 11, some 27 percent of students who took the PERT scored college-ready in math (Figure 9). About half of the students who scored below college-ready in grade 11 retook the PERT in grade 12. Despite taking the math exam again, only a quarter of those students scored college-ready, although an additional 59 percent of them did improve their score.

Figure 9. Tree diagram illustrating student participation and performance on the PERT: grade 11 students in 2011/12

Math | **Reading**



Source: Author's calculations based on student-level data provided by the Florida Education Data Warehouse.

Some 46 percent of students who took the PERT in grade 11 scored college-ready in reading. Of the students who scored below college-ready, 47 percent took the reading PERT again in grade 12, and 32 percent scored college-ready. Another 38 percent of retakers improved their score, but not enough to meet college readiness criteria in reading. Although some of these gains in PERT scores may be attributed to students' improvement in college-level skills, there may also be other causes, such as regression to the mean if students scored unusually low on their first attempt and closer to their true ability on the retest.

Although the majority of students did not retest college-ready, many still made large gains. Among students who improved but still did not score college-ready, about 45 percent improved their scores by 10 points or more (Table 4).

Table 4. Performance on PERT retakes for students who scored below college-ready on PERT in grade 11 in 2011/12

	Math (N=39,865)		Reading (N=17,756)	
Performance in grade 12	Number	Percent	Number	Percent
Scored college-ready	10,086	25.3%	5,641	31.8%
Retook and improved but still not college-ready:	23,656	59.3%	6,784	38.2%
Improved 1 to 3 points	3,960	16.7%	1,363	20.1%
Improved 4 to 6 points	4,739	20.0%	1,213	17.9%
Improved 7 to 9 points	4,587	19.4%	1,051	15.5%
Improved 10 or more points	10,370	43.8%	3,157	46.5%
Retook and scored the same/worse	6,123	15.4%	5,331	30.0%

Source: Author's calculations based on student-level data provided by the Florida Education Data Warehouse.

CRS course participation

There were initial delays in CRS course implementation, with only 8 percent of schools offering CRS courses in reading and 27 percent of schools offering CRS courses in math during the first year of the voluntary program (Figure 10). The percentage of schools offering these courses increased each year. Once CRS courses became mandatory for targeted students, over 90 percent of schools participated by offering at least one CRS course in each subject area.

Schools that first adopted CRS courses during the voluntary FCCRI were less likely to be in rural locales than schools that adopted during the mandatory FCCRI (24.1 percent versus 29.1 percent, Table 5). Voluntary adopters also had a higher percentage of Black and Hispanic students than mandatory adopters (48.9 percent versus 44.8 percent). Other student demographic characteristics and the percentage of students targeted under the FCCRI were similar among schools that adopted during the voluntary and mandatory FCCRI.

Figure 10. Percentage of schools offering at least one CRS course, by cohort

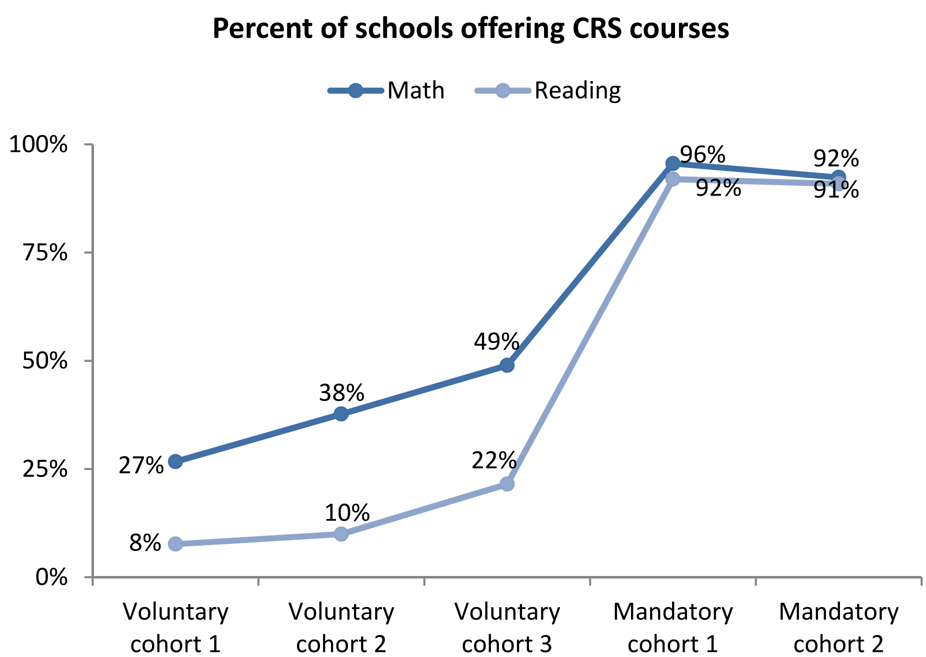


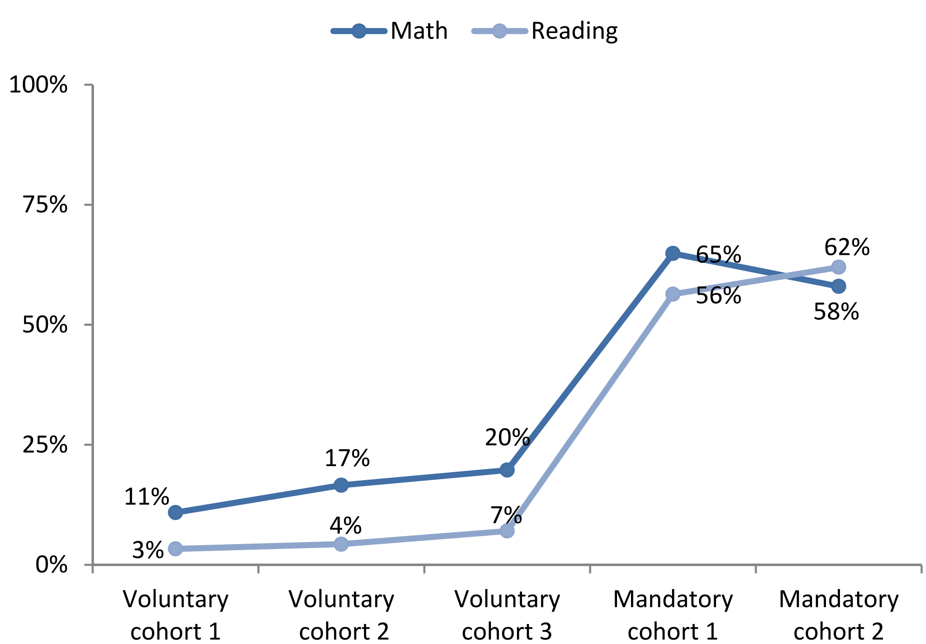
Table 5. Characteristics of schools that first adopted CRS courses under the voluntary FCCRI versus those of schools that first adopted under the mandatory FCCRI

	Voluntary adopters	Mandatory adopters	Difference
Percentage of students targeted...			
Math	84.0%	84.5%	-0.6%
Reading	54.8%	55.5%	-0.7%
Student demographics			
Percent FRPL	49.8%	51.4%	-1.5%
Percent Black/Hispanic	48.9%	44.8%	4.1%
Percent female	49.7%	51.4%	-1.7%
Percent mental disability	11.2%	12.0%	-0.8%
Percent ELL	6.5%	5.7%	0.8%
School locale			
Urban	25.3%	22.9%	2.4%
Suburban	50.6%	47.9%	2.7%
Rural	24.1%	29.1%	-5.0%

NOTE: There were 162 schools that ever participated in testing under the voluntary FCCRI and 292 schools that first participated in testing under the mandatory FCCRI. Twenty schools did not participate in testing under the voluntary or mandatory FCCRI (not shown here).

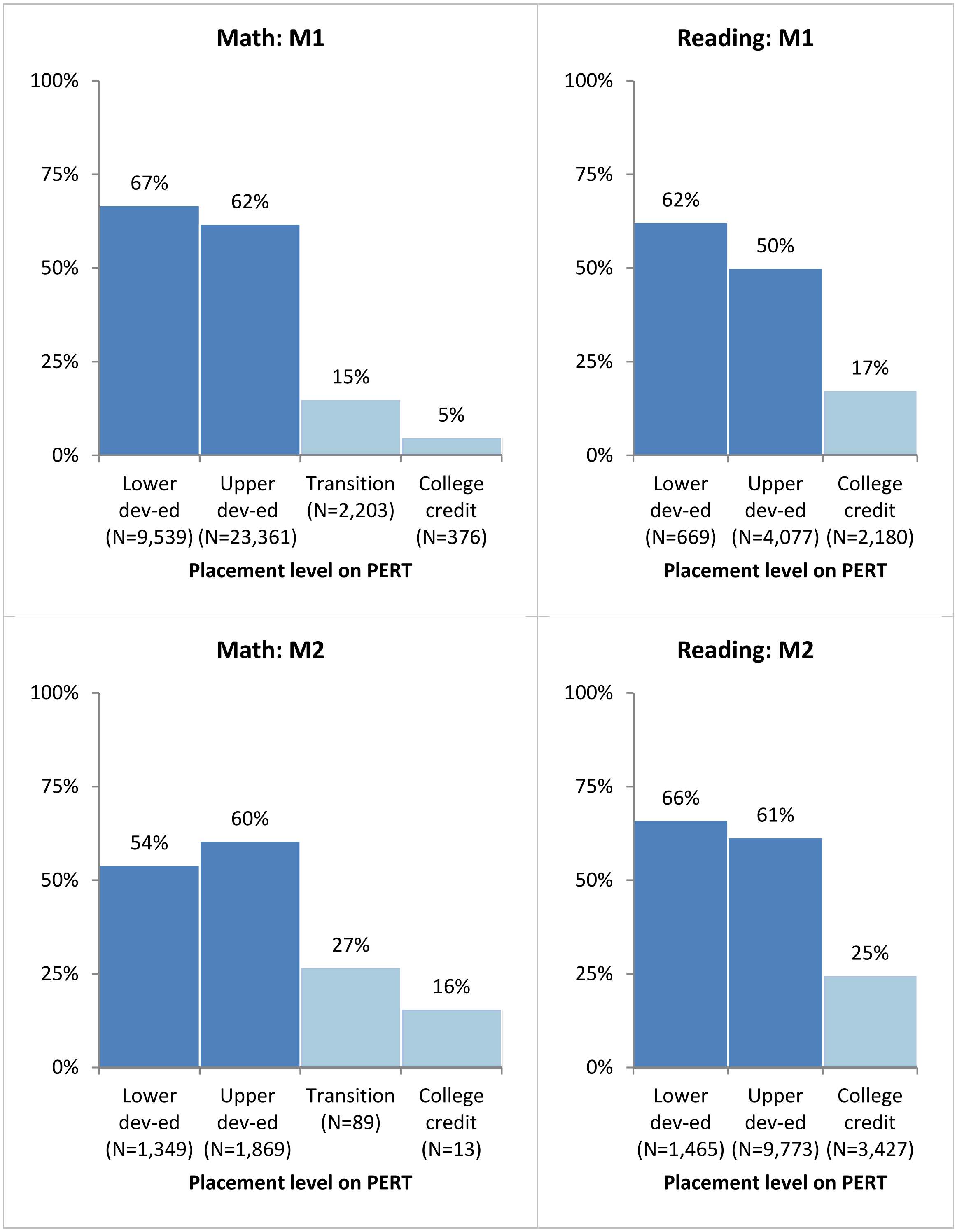
Student participation in CRS courses was also low in the FCCRI's early years. In voluntary cohort 1, participation rates were 3 percent in reading and 11 percent in math for targeted students who scored below college-ready on the CPT (Figure 11). CRS course enrollments increased during each year of the voluntary FCCRI, but still fewer than one-quarter of targeted students participated by the end of the voluntary period. When CRS courses became mandatory, participation rates rose to about one half to two-thirds of targeted students who scored below college-ready on the PERT. Even though this was a large increase from the voluntary FCCRI, nearly one-third to one-half of students who should have enrolled in CRS courses did not. It is not clear how much of this non-compliance is due to schools not attempting to implement the requirement versus schools that implemented the requirement but allowing many exemptions.

Figure 11. Student-level participation rates in CRS courses for targeted students who scored below college ready on the CPT or PERT, by cohort



CRS course participation rates tended to decline as student performance on the college readiness test improved (Figure 12). Participation rates were highest for targeted students who received a score on the grade 11 PERT corresponding with placement in upper level or lower-level developmental education courses at state colleges Additionally, between 5 to 27 percent of targeted students chose to enroll in CRS courses even though they received a college-ready score on the PERT corresponding with placement in nondevelopmental courses.

Figure 12. Student-level CRS course participation rates for targeted students, by placement level on the grade 11 PERT (developmental education courses shaded in dark blue)



The qualitative data revealed several reasons students may have enrolled in CRS courses even if they were not targeted for participation. During the small-group discussions, teachers from schools in four districts explained that all students who were absent or truant and therefore missed taking the PERT in grade 11 were placed in CRS courses in grade 12. Furthermore, some high-performing students on the FCAT or PERT chose to enroll in CRS courses even though they were not required to do so. Eliminating standard-level grade 12 English courses in some districts limited students to either taking a challenging honors or AP course or taking the less challenging CRS course. However, this trend appears to be limited to a small number of school, as the overall number of courses offered in grade 12 by subject area remained relatively similar over time. In four of the eight small-group discussions, teachers described that students who performed well on the FCAT but did not want to do the extra work of a higher-level English class chose to enroll in the CRS course.

In some cases, administrators placed students in CRS courses. Teachers from two schools in the same district described that students who were struggling in more advanced classes were placed into CRS courses. One teacher stated that the administration placed students in her math CRS course because they needed an additional math credit to achieve the four credits required to graduate. Another teacher noted that administrators also occasionally moved students out of the CRS courses if they achieved a concordance score on the ACT or SAT after the course began.

Description of treatment and counterfactual conditions

The conditions for the college readiness testing component for the FCCRI were straightforward. Students in the targeted range of achievement levels in grade 10 FCAT were assigned to participate in the college readiness test in grade 11, while students above or below this targeted range were not assigned to participate in the testing.

Yet when we examine the CRS component of the FCCRI, there are important distinctions to make between the treatment and counterfactual conditions. Although targeted students who scored below college-ready on the college readiness assessment were assigned to participate in CRS courses in grade 12, there was no single course assignment for students who scored above college-ready. In this section we examine the content and goals of the CRS courses, as well as startup challenges in the development of CRS courses, to provide a better understanding of the treatment conditions. We also examine differences in the types of courses taken in grade 12 by students who scored just above and just below college-ready.

CRS course content

FLDOE approved five CRS courses. Three were one-semester courses: Mathematics for College Success, Reading for College Success, and Writing for College Success. Two were two-semester courses: Mathematics for College Readiness and English 4: College Prep. Schools were required only to offer one course in math and one course in English Language Arts. Most of the high schools we visited offered only the two-semester courses that count toward graduation requirements in their respective subject areas.

Each CRS course had standards set by FLDOE that defined the topics to be covered in the course. Districts, high schools, and teachers had considerable discretion in how they implemented these courses, however. This means that even though all CRS courses share the same label, their content may have differed substantially across classrooms.

Overall, the difficulty level of the courses varied widely across schools. A counselor in one medium-size district high school said that the English CRS course was "a step above regular English IV." Yet teachers in other districts noted that they spent a lot of time reviewing basic skills because many students in their CRS courses had not yet passed the FCAT.

Guidance to schools from FLDOE (Florida Department of Education, 2008) about the CRS courses said that, "The postsecondary readiness curriculum offered by the community college will be used to develop the high school course. The development of the high school curriculum should mirror the college course and will require secondary-postsecondary collaboration where subject area faculty, both high school and college, make joint decisions regarding the curriculum and student outcomes" (p. 4). In practice, we found very little evidence of secondary-postsecondary collaboration around the CRS courses. There were also mixed views about how similar the CRS courses were to developmental education courses at state colleges. One teacher in a large district who is also an adjunct at a state college said, "It's the same content but in a different order." Another instructor at the same state college disagreed and said, "[The high school] courses are supposed to mimic ours, but they don't. We're not sure they're using the state competencies."

Math CRS courses

Many math CRS courses took the form of review of basic skills of Algebra I and II. By the time they took the math CRS course, most students had taken Algebra I and II and Geometry—three of the four math courses required for high school graduation. One teacher commented that the students "really, really get it"—because this was the second time they were receiving instruction in the content of the course. For some teachers, the emphasis of these courses was on refining students' Algebra II skills to prepare them for college-level Algebra. Some of the teachers noted that students in the math CRS courses would otherwise have been taking courses such as Pre-Calculus, Financial Applications, Trigonometry, or Math Analysis. One interviewed teacher said CRS courses were "absolutely" a better fit for some students. Another teacher said CRS courses were "less scary" for them than Pre-Calculus and students were "happier" there.

One of the schools we visited used a different approach. The school offered a one-semester Math for College Success course for students who did not pass the PERT, focusing almost entirely on ACT test prep. The school also offered a two-semester Math for College Readiness course that was "much more rigorous" and went beyond Algebra II content. This approach was not observed in any of the other schools that participated in the site visits, however.

English Language Arts CRS courses

English Language Arts CRS courses tended to be similar to a traditional English IV course. Many of the teachers interviewed approached the CRS courses in the same way they would other English Language

Arts courses but placed more emphasis on integrating ACT or PERT preparation and study skills into the CRS course. Other teachers followed the same curriculum used for their honors English IV course but implemented it at a slower pace.

The level of rigor also varied by school. In one rural district, students in the English Language Arts CRS course did not read a single novel during the entire school year because the teacher struggled to get students to read even a few pages.

Given the similarities between English Language Arts CRS courses and the English IV courses, some high schools dropped the traditional English IV course in favor of the CRS course for seniors. Honors English or AP English still were offered in these schools, but any student not taking these courses enrolled in the English/language arts CRS course to satisfy the requirement to complete the four English Language Arts courses needed for graduation. If we look at changes in specific course offerings from 2010/11 to 2012/13 among the 474 schools in our administrative data from EDH, the most common courses eliminated were English 4 (77 fewer schools), English 4 Honors (41 fewer schools) and Reading: 9-12 (29 fewer schools).

Goals of CRS courses

We found that CRS courses had several goals associated with preparing students for life after high school, including helping students to test college-ready, develop the academic skills needed in college, develop soft skills needed in college, develop career-related skills, and decide what they want to do after high school. Districts and teachers placed different levels of emphasis on each of these goals as they implemented the CRS courses, however.

Developing academic skills

The foundation of many CRS courses was to provide students with the academic skills and knowledge needed to perform at the college level. Surveyed teachers put a strong emphasis on developing the academic skills of students, with 86 percent saying it was a “very important” or “important” focus.

Many of the interviewed teachers placed a high emphasis on this goal because of the importance of students’ having a strong base of content knowledge when entering college. Many teachers believed developing students’ academic skills was the most important goal in preparing students for life after graduation. Teachers commented that fundamental reading and writing skills prepared students for a variety of college and career opportunities over their lifetimes regardless of what a student might do after high school.

Testing college-ready

Teachers also highly emphasized ensuring that students test college-ready. Of teachers surveyed, 80 percent reported that test preparation was an “important” or “very important” focus in the course. Although many teachers used the PERT to gauge students’ college readiness, students also could test college-ready using the ACT, SAT, or Accuplacer. This high level of emphasis on testing college-ready was consistent across the large, medium-size, and small districts. Site visits and interviews also confirmed this emphasis on test preparation.

Teachers commented that they highly emphasized students' testing college-ready for two reasons. First, it gave students options after high school, even for students not planning to attend college. Second, their district placed significant importance on getting students to test college-ready, sometimes at the expense of other goals of the CRS course.

Not all teachers highly emphasized this goal, however. Some of the reasons teachers gave during site visits were that they did not have much guidance on the PERT; that they had to focus on basic deficiencies rather than on college readiness because students were too far behind academically; and that they believed students were overtested or got test preparation in other settings.

Other goals: Soft skills, postsecondary plans, and career-related skills

Whereas CRS teachers highly emphasized academic content and students testing college-ready, other goals received lower priority, such as developing the soft skills needed in college, helping students decide what to do after high school, and developing career-related skills. Fewer than 50 percent of teachers surveyed focused on those goals. The main reason they gave was that there was no specific mention of these goals in the state standards for the CRS courses, and many times there was no expectation that teachers should cover this kind of content. Thus teachers did not need to include these skills in their courses in order to comply with implementation as intended by state.

Many of the teachers who were integrating these goals were doing so on their own initiative because they believed it was important for student development in a course about college readiness. Teachers in our interviews who highly emphasized developing soft skills stressed that students would not be able to succeed in college without good study habits, time management skills, and knowing how to use technology in academic settings, especially to complete academic tasks such as writing a research paper. For the career and college choice goals, teachers believed it was important to expose students to the various opportunities available to them after high school so that they could make informed decisions about what profession to pursue.

Although many interviewed teachers wished they could dedicate more time to these goals, there simply was not enough time available because of the strong emphasis on academic content and testing. Some of the teachers noted that they could spend only a few days in an entire semester on some of these goals. The rest of the time was spent on academics and test preparation.

In addition, the soft skill and college and career planning goals tended to overlap with other parts of the high school experience, including students' interactions with counselors. As a result, some teachers believed that less time needed to be spent on those goals in the CRS courses.

Startup challenges in developing CRS courses

As with any large statewide initiative, implementation is a complex endeavor and it takes time to work out the most effective mechanisms for translating ideas into actions. In the feedback we collected in the first year of the mandatory FCCRI program, teachers noted several implementation issues: lack of preparation time for new CRS courses, lack of adequate course materials and the inability of school and district leaders to provide funds for adequate materials or textbooks, variations or lack of clarity in criteria used to place students in CRS courses, and ambiguities in course content and goals. Many districts could not afford to purchase new materials and generally had to wait until funds became available on regular cycles to replace out-of-date texts and other materials with content specifically geared to the college readiness courses. CRS teachers noted that they were left largely on their own in developing lesson plans and materials. In most high schools, because no person was responsible for developing curricula, pacing guides, and materials, the CRS courses were not taught in a uniform way. These findings suggest that the challenges with initial program implementation could lead to negative impacts on student outcomes in the first year.

Also during the first year of the mandatory FCCRI, CRS teachers had expressed concern that they lacked information about what topics were covered on the PERT, to what depth each topic should be covered in preparing students to take the test, and how well the PERT aligned with the knowledge and skills students need to succeed in college. The PERT was a new assessment and information was limited in comparison with the resources available on the PERT's predecessor, the CPT. The CPT, which was developed by the College Board, closely resembled the Board's Accuplacer used nationwide, which had a broad array of study guides and practice tests.

In the second year of the mandatory FCCRI, we surveyed CRS teachers about the extent to which they believed these impediments had changed from the prior year. The largest improvement was in the receipt of information about the PERT, with 37 percent of teachers reporting that the impediment had been reduced. Several factors contributed to this large positive difference. In the intervening year, the company that created the PERT released more information about it; the state more widely disseminated this information, especially to districts; and districts had considerably more time to provide this information to teachers. The research team also worked with state colleges and FLDOE to develop a social media site on Edmodo.com that disseminates information about the PERT, as well as practice tests and review materials. A possible additional factor, however, is that teachers could have asked their 2012/13 students to describe the test and the corresponding areas that were well covered, and poorly covered, by their CRS courses.

About 20 percent of teachers reported an improvement in the lack of teaching materials between the first and second year of the mandatory FCCRI, which could potentially contribute to differences in the impact of the FCCRI in M1 and M2. However, most teachers still reported that they were largely on their own to develop materials for their CRS courses. There were also large differences by subject area, with 55 percent of English teachers reporting that lack of teaching materials was an "important" or "very important" impediment, compared with only 30 percent of math teachers.

Differences in grade 12 course-taking by college-ready status

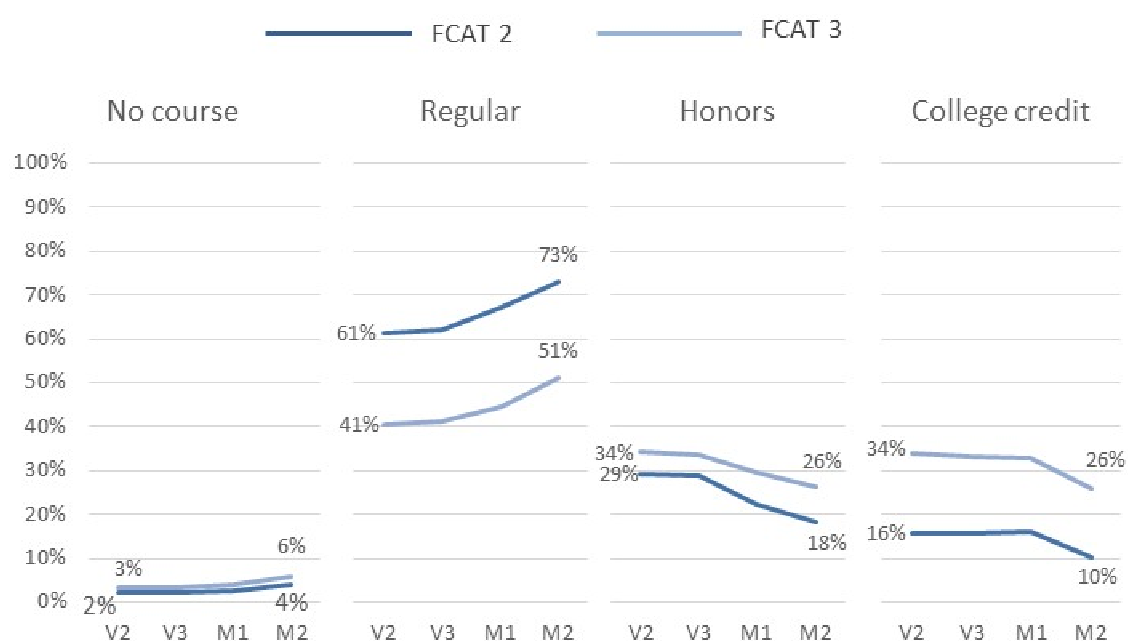
Next, we looked at changes over time in the types of courses taken in grade 12 for targeted students at each FCAT level. This indicates the types of courses that students substituted for CRS courses as the FCCRI expanded. We exclude the V1 cohort from this analysis because these students had different high school graduation requirements, which likely affected course-taking independent of the FCCRI. V1 students were required to complete only three credits in math, whereas students in V2 to M2 were required to complete four credits. Students in all cohorts were required to complete four credits in English.

FLDOE classifies courses using three level designations: (1) basic, (2) regular, and (3) advanced. In this analysis, honors-level courses are distinguished from college credit–bearing courses in FLDOE's "advanced" category. College credit–bearing courses include Advanced Placement (AP), International Baccalaureate (IB), Advanced International Certificate of Education (AICE), and dual enrollment courses. Within the FLDOE classification, the CRS courses are categorized as "regular" courses. If students enrolled in more than one course in the same subject area during grade 12, we used the highest-level course.

We found that from V2 to M2, enrollment in regular English courses for all targeted students rose by approximately 10 percentage points (Figure 13).[4] Most of the increase in regular course enrollment was due to declines in honors course enrollment, although there were also declines in enrollment among college credit courses. Only a small percentage of targeted students in all cohorts and FCAT levels (3 percent or less) did not take an English course in grade 12.

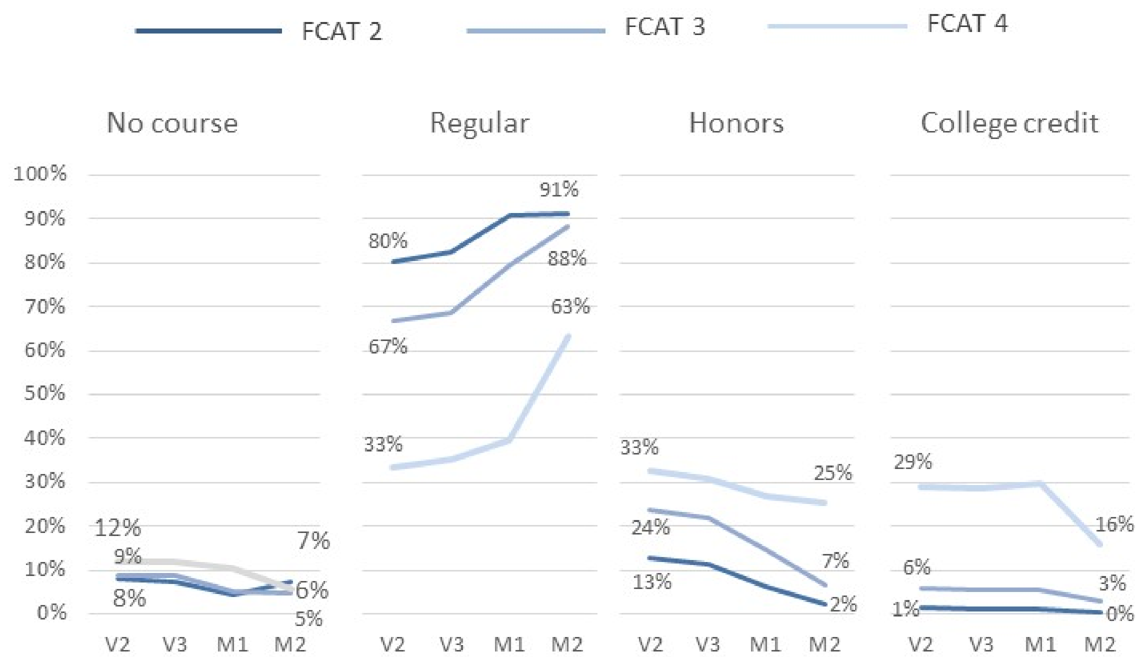
We observed a similar trend in math courses—regular course enrollments increased over time while college credit and honors-level course enrollments declined (Figure 14). The greatest change in math course levels occurred from cohort M1 to M2. This is because only students who took Algebra I in grade 10 were required to take the state assessment in cohort M2, so more advanced students who took Algebra I in earlier grade levels are not included in these figures like they were for the other cohorts.

Figure 13. Percentage of targeted students enrolling in each English course level from cohort V2 to cohort M2, by FCAT level



[4] In both math and English, 1 percent or less of targeted students enrolled in a grade 12 course below the regular level (not shown in the figures).

Figure 14. Percentage of targeted students enrolling in each math course level from cohort V2 to cohort M2, by FCAT level

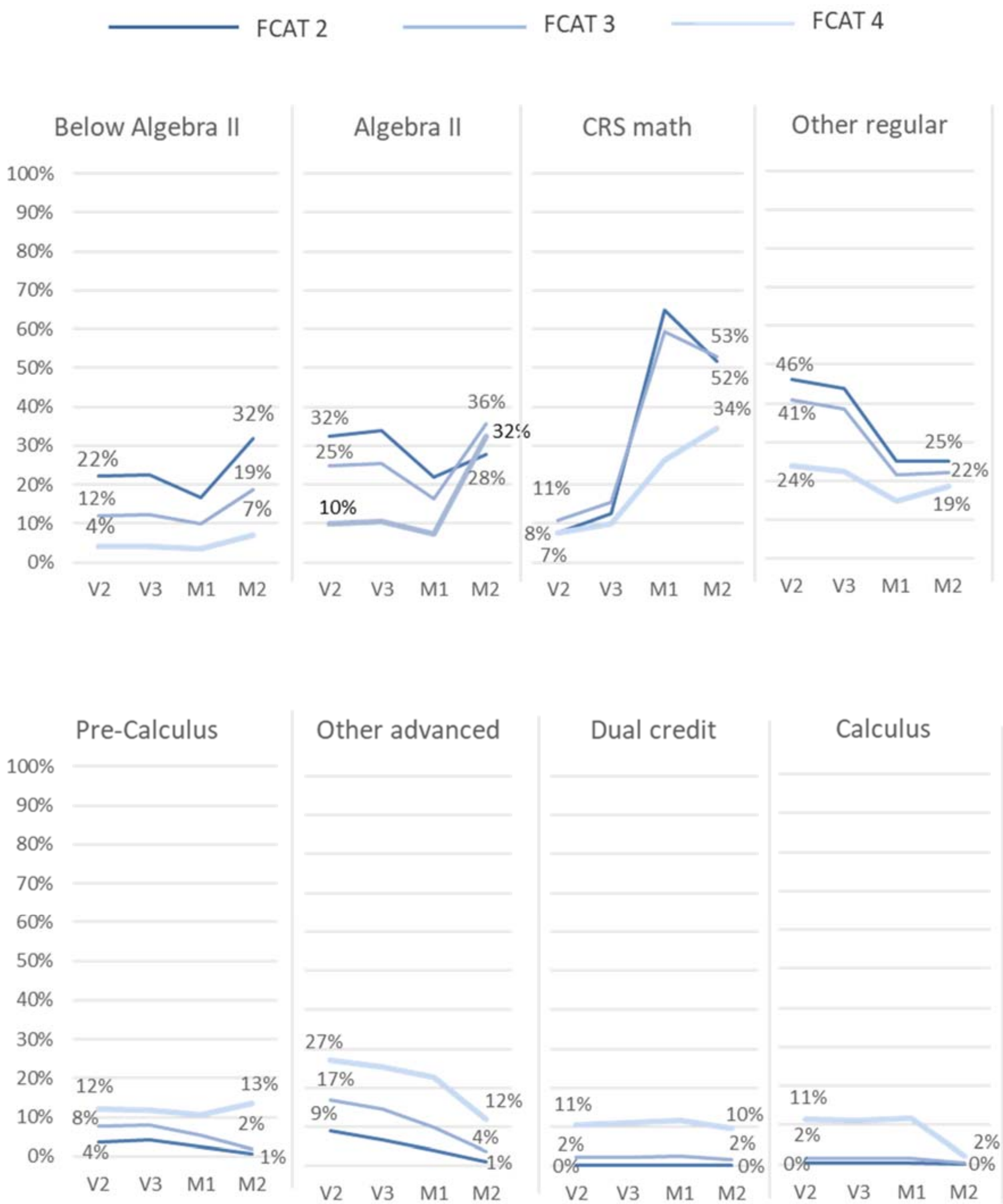


In math, students also had the option to choose among different course types such as Geometry and Calculus. We also examined the change in the percentage of targeted students enrolling in each type of math course from V2 to M2 (Figure 15).[5] As expected, the greatest change is in the category of CRS courses. For FCAT 2 and FCAT 3 students, CRS course enrollment increased more than 40 percentage points, from around 10 percent in V2 to more than 50 percent in M2. FCAT 4 students had a smaller increase of 27 percentage points (from 7 percent in V2 to 34 percent in M1). The greatest decreases in math course enrollment for FCAT 2 and 3 students were in the categories of "other regular math" (which includes courses such as Financial Algebra, Analytic Geometry, and Liberal Arts Math) and Algebra II. For FCAT 4 students, there were declines in enrollment in "other regular math", "other advanced math" (which includes courses such as AP Statistics and Honors Probability), and calculus.

The changes over time in grade 12 course-taking may have important implications that differ by FCAT level. Higher-performing targeted students were more likely to be taking CRS courses at the expense of honors-level and other advanced courses. This suggests that the FCCRI could have a negative impact on postsecondary outcomes since students required to take CRS courses would likely experience a decline in the rigor of course-taking in grade 12 relative to their peers who were not required to take these courses. On the other hand, lower-performing targeted students may have taken CRS courses in lieu of other regular courses such as liberal arts math. These other courses may not have prepared students as well for college-level work, so the CRS courses may have been a better option for these students.

[5] English courses are not represented due to the much higher number of courses offered and the lack of a clear course hierarchy.

Figure 15. Percentage of targeted students enrolling in each math course type from cohort V2 to cohort M2, by FCAT level.



The changes in grade 12 course-taking patterns may also influence student outcomes because of peer effects (e.g. Sacerdote, 2011). Under the mandatory FCCRI, fewer targeted students enrolled in honors courses. This means there would have been more above-target students (FCAT4s and 5s in reading, and FCAT 5s in math) remaining in the honors courses, so these courses may have become more rigorous. Within the CRS courses, the quality of peer effects may have decreased for higher-performing targeted students and increased for lower-performing targeted students, relative to the courses they would have otherwise taken.

Lastly, we checked whether there were any declines over time (from V1 to M1) in the percentage of students who ever took Algebra II in high school. Algebra II is often considered a critical course for college success, and we wanted to check whether there were any unintended consequences of the FCCRI if students in MI were less likely to take Algebra II because they were required to enroll in CRS courses. This does not appear to be the case, as participation rates in Algebra II increased slightly during each year of the FCCRI. However, this increase is unlikely to be due to the FCCRI, and may be attributed to other changes in high school graduation requirements over time. For V1 through M1, the highest math course required was Algebra 1. M2 was the first cohort required to take both Algebra I and Geometry (or higher), and students graduating in spring 2017 were required to enroll in math courses up to Algebra II. There may have been a gradual increase over time in Algebra II enrollment because schools were gearing up for the increased graduation requirements, so they may have been more apt to advise all students to take Algebra II.

Chapter 4: Confirmatory Impact Analysis of the FCCRI Using Regression Discontinuity

This chapter describes the confirmatory impact analyses, which used a regression discontinuity design to assess the FCCRI's effect on student outcomes. We begin by describing the econometric methodology and the data used in the analyses. Then we test the validity requirements for regression discontinuity and present the results.

We restricted our primary analysis to students who were first-time grade 11 students in 2011/12. This was the first mandatory cohort (M1) for which targeted students were required to take a college readiness assessment in grade 11 and CRS coursework in grade 12 if they scored below college-ready. Although participation in these activities rose over the prior three voluntary cohorts, it remained too low for meaningful analysis. Short-term outcomes were examined one year after high school graduation, and longer-term outcomes were examined two and three years after high school graduation. The longer-term analyses both let us observe some of the same outcomes over time and let us observe additional sets of outcomes, such as persistence and transfer, that could not be observed over the course of a single year.

Our secondary analysis uses first-time grade 11 students in 2012/13 (cohort M2), though different high school assessments and postsecondary remediation policies mean that results are not directly comparable across cohorts. This was the last cohort where targeted students were required to participate in both the college readiness assessment and the CRS courses. Short-term outcomes were examined one year after high school graduation and long-term outcomes were examined two years after high school graduation.

Econometric methodology

We measured the FCCRI's impact using regression discontinuity (RD) methodology. RD analysis is typically used when assignment to a policy treatment is determined by whether a continuously valued variable (the "running variable") has crossed a predetermined cutoff. RD gained popularity as a method used for studying the impact of election outcomes (DiNardo & Lee, 2004; McCrary, 2008). Within education contexts it has been used when determining eligibility for financial aid (Leeds & DesJardins, 2015; Van der Klaauw, 2002) or class rank–based programs (Daugherty, Martorell, & McFarlin, 2014). This analysis separately estimated for 1) the impact of assignment to participate in the PERT in grade 11, 2) the impact of assignment to a CRS course in grade 12, and 3) the impact of placement in developmental coursework prior to college enrollment (since students scoring college-ready on the PERT in grade 12 were exempt from enrolling in developmental education coursework) among students at each margin.

RD analysis isolates the impact of the policy being analyzed without capturing outside or irrelevant factors. Since true experiments (where students are randomly assigned to treatment and control groups) are difficult to implement in policy settings, RD is one way to compare students who were "as good as" randomly assigned to these groups. Often, researchers worry that treatment and control groups are not randomly determined—that the treatment group is systematically different in ways that are observed (such as by race, gender, or socioeconomic status) or typically not observed (such as by motivation or work ethic)—and that findings reflect systematic features of the treatment group rather than the policy being evaluated. As part of RD analysis, researchers verify that observed characteristics do not change discontinuously at the cutoff for treatment, often by looking in a very narrow range around the cutoff. Although unobserved characteristics by definition cannot be checked, showing that individuals are not bunching on either side of the cutoff for treatment (for example, that particularly motivated or high-ability students are not manipulating their FCAT scores to avoid having to participate in the FCCRI) can minimize these concerns. Ultimately, if assignment to the treatment group is the only thing that changes noticeably at the cutoff for treatment, any difference in student outcomes should be attributable to that treatment.[6] This approximates an experiment, because students very near the cutoff are assigned to treatment or control based on a method that is "as good as random" from a statistical point of view.

The main drawback of RD is that its results apply only near the cutoff for treatment and are not generalizable to all students. For instance, the impact of taking a CRS course should be quite different at different points in the FCAT distribution—students with very low FCAT scores might need additional remedial coursework before CRS courses would be useful and those with very high FCAT scores would have likely already mastered the material, but those with midrange FCAT scores might both understand the material and benefit from it. The strengths of RD—focusing on a specific area where student characteristics are fairly stable—prevent the estimates it produces from being generated elsewhere. As a result, RD results should be discussed only within the context of the cutoff for treatment and should not be presented as an overall impact.

Table 6 shows how FCAT performance levels (PLs) are used to assign students to college readiness testing and how college readiness tests are used in grade 11 to assign students to college readiness coursework and in grade 12 (and onward) to assign students to developmental education coursework. Although the assessments group students into broad categories, students also receive scaled scores—between 100 and 500 on the FCAT and between 50 and 150 on the PERT—that function as nearly continuous measures of student achievement. Using these scaled scores, it is possible to compare students on either side of a proficiency level cutoff who have extremely similar profiles and differ primarily in their assignment to treatment.

Table 6. Assignment to treatment by assessment, cohort M1

FCAT		
	Math	**Reading**
PL1	Scores 100–286 At risk for failing to graduate Not assigned to take PERT	Scores 100–286 At risk for failing to graduate Not assigned to take PERT
PL2	Scores 287–314 Eligible for graduation at 300 Assigned to take PERT	Scores 287–326 Eligible for graduation at 300 Assigned to take PERT

[6] Multiple treatments determined at the same cutoff must be evaluated together or not at all, unless there is a clear method of separating their impacts.

FCAT		
PL3	Scores 315–339 Assigned to take PERT	Scores 327–354 Assigned to take PERT
PL4	Scores 340–374 Assigned to take PERT	Scores 355–371 Not assigned to take PERT
PL5	Scores 375–500 Not assigned to take PERT	Scores 372–500 Not assigned to take PERT
grade 11 PERT		
	Math	**Reading**
Not College-Ready	Scores 50–112 Assigned to CRS course	Scores 50–103 Assigned to CRS course
College- Ready	Scores 113–150 Not assigned to CRS course	Scores 104–150 Not assigned to CRS course
grade 12 PERT		
	Math	**Reading**
Not College-Ready	Scores 50–112 Assigned to developmental coursework	Scores 50–103 Assigned to developmental coursework
College-Ready	Scores 113–150 Not assigned to developmental coursework	Scores 104–150 Not assigned to developmental coursework

Table 7 shows the grade 10 cutoff values for cohort M2. Outcomes shown for the Algebra 1 EOC assessment apply only to students taking it at the end of grade 10; policy regarding students who take it at the end of grade 9 is unclear, and the data do not show discontinuities in PERT-taking based at any grade 9 scores. In reading, data for cohort M2 reflect an updated scoring scale for the FCAT (though its content remained unchanged). PERT assessments used the same score cutoffs for both cohorts.

Table 7. Assignment to treatment by grade 10 assessment, cohort M2

Grade 10 assessments for cohort M2		
	Algebra 1 EOC	FCAT Reading
PL1	Scores 325-374 At risk for failing to graduate Not assigned to take PERT	Scores 188-227 At risk for failing to graduate Not assigned to take PERT
PL2	Scores 375-398 Assigned to take PERT	Scores 228-244 Assigned to take PERT
PL3	Scores 399-424 Eligible for graduation at 399 Assigned to take PERT	Scores 245-255 Eligible for graduation at 245 Assigned to take PERT
PL4	Scores 425-436 Assigned to take PERT	Scores 256-270 Not assigned to take PERT
PL5	Scores 437-472 Not assigned to take PERT	Scores 271-302 Not assigned to take PERT

We used a sharp RD design with a local linear framework, formally written as:

$$Y_i = \alpha_1 \widetilde{R_i} + \gamma_i * \mathbf{1}\{\widetilde{R_i} \geq 0\} + \alpha_2 * \mathbf{1}\{\widetilde{R_i} \geq 0\} * \widetilde{R_i} + X_i\beta + \varepsilon_i$$

where Y_i is an outcome of interest for individual i, $\widetilde{R_i} = R_i - c$ is the running variable R_i recentered around cutoff c, X_i is a vector of individual characteristics, and ε_i is a mean-zero error term. $\mathbf{1}\{\widetilde{R_i} \geq 0\}$ is an indicator variable equal to 1 if the running variable is at or above the cutoff and zero if it is below. The parameter of interest in this equation is γ_i, the impact on individual i of being in the treatment group; $E[\hat{\gamma}]$ is the average treatment effect at the cutoff for treatment when compliance is perfect and the intent-to-treat effect at the cutoff otherwise.[7] In our context, the running variable, R_i, is a student's scale score on either the FCAT or the PERT; depending on the regression being run, c may be the cutoff between FCAT proficiency levels or the PERT score required for college readiness in a particular subject area.[8]

Sharp RD analysis at the FCAT cutoffs can measure the impact of telling students at those cutoffs to take the PERT (the intent to treat), but does not provide the impact of actually taking the test. If people on both sides of the cutoff complied perfectly with their treatment status—if all those below the cutoff did not take the PERT and those above it did—then sharp RD estimates would measure the impact of the treatment itself; however, compliance with assignment to PERT testing or CRS course-taking was far from perfect.[9] While a fuzzy RD framework would better capture impacts on students who comply with treatment, we use sharp RD to provide a clearer picture of how the FCCRI functioned from a policymaker's perspective.

Figure 16 shows how each of the two cohorts complies with each stage of the FCCRI. Depending on the graph, each data point shows either the percent of students at a given FCAT or EOC score who take the PERT or the percent of students at each PERT score who take the corresponding CRS course. Across both subjects and both cohorts, there is a discontinuous jump in PERT-taking at the low FCAT margin, a discontinuous drop in PERT-taking at the high FCAT margin, and a discontinuous drop in CRS-taking at the PERT's college readiness cutoff. However, many students are "never-takers," who do not take the PERT or CRS courses even when assigned to do so, or are "always-takers," who do the opposite.[10]

In practical terms, the FCAT's "mandatory" testing cutoffs are decent (but far from perfect) predictors of whether students take the PERT, and the PERT's college readiness cutoffs are decent (but far from perfect) predictors of whether students take CRS courses. The FCCRI did not narrowly target students on the margin of college attendance, as many students targeted for PERT testing did not meet FCAT

---

[7] "Compliance" refers to whether students assigned to treatment receive it (and vice versa); this study and others use the term without positive or negative connotation.

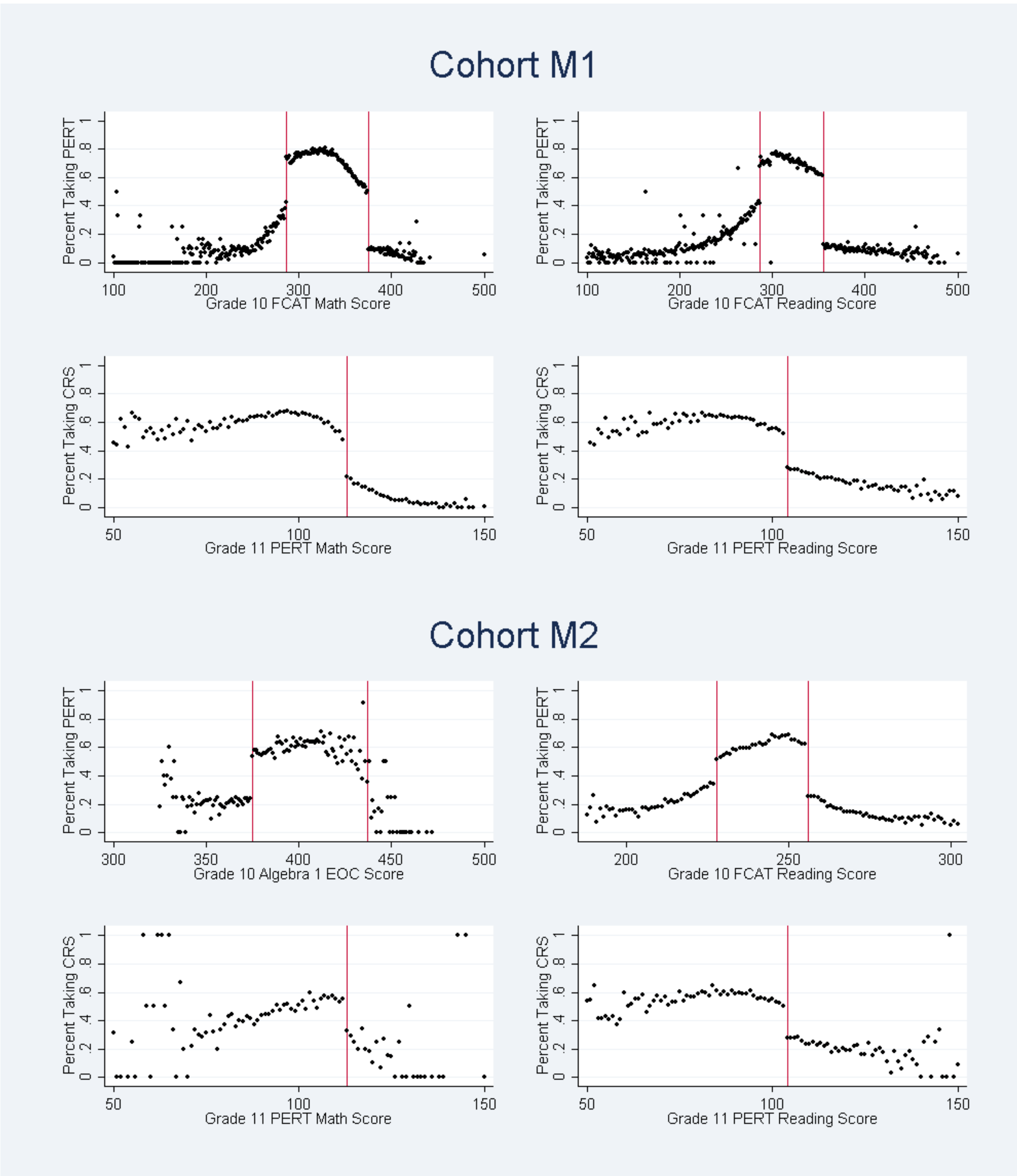
[8] For the upper FCAT margin and the grade 11 PERT, where the policy treatment goes to students scoring below a particular cutoff, we modify the RD equation above to contain $\mathbf{1}\{\widetilde{R_i} \leq 0\}$.

[9] Individuals who receive or reject treatment no matter their assignment may each have systematically different observed and unobserved characteristics from those who comply with their treatment status. The impact of forcing them to (not) receive treatment may therefore be systematically different from that for compliers.

[10] There is only a small discontinuity in CRS-taking at low FCAT margins and none at high FCAT margins. The former may be due to opportunities for noncompliance or insufficient course offerings, while the latter is because students with high FCAT scores are very likely to score college ready on the PERT. This cautions against expecting discontinuities at FCAT cutoffs—students barely in the treatment group are being told that they could become college-ready and are taking a college readiness assessment (both of which might affect their motivation) but are not substantially more likely than students barely outside the treatment group to receive assistance in becoming college-ready.

requirements for high school graduation.[11] RD analyses at FCAT thresholds may capture motivation or information effects from taking the PERT, while those at PERT thresholds are more likely to capture the result of preparation from enrolling in the CRS courses.

Figure 16. Compliance with FCCRI assignment



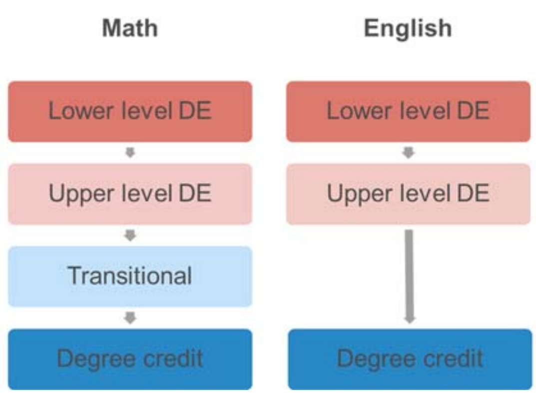
[11] Very few of these students passed the PERT or seamlessly enrolled in college.

Outcome variables

The short-term outcomes we study occurred after grade 12. These include whether a student graduated high school with any type of diploma, enrolled in a postsecondary institution in the fall semester after on-time graduation (seamless fall enrollment), enrolled in a nondevelopmental course in the relevant subject, and passed a nondevelopmental course in the relevant subject. All outcomes and their definitions are presented in Appendix A (Table 9A).

Figure 17 summarizes the sequence of math and English courses for the college coursetaking outcomes. Math is unique in offering a transition course (Intermediate Algebra) that counts for elective credit but not for degree or major requirements. We study nondevelopmental rather than degree credit coursework because the cut score for placement into Intermediate Algebra is the same score used to assign students to CRS math in grade 12. There is no transition course in English; students at the college readiness cutoff in reading are recommended for degree credit coursework.

Figure 17. Sequence of college-level courses in math and English at all Florida colleges.



Note: DE refers to developmental education.

Data for cohort M1 only allows us to examine nondevelopmental enrollment and pass rates only through a second year of postsecondary enrollment. This accounts for students who failed in their initial attempt(s) at nondevelopmental courses, those who were unable to take one (e.g. if they required two semesters of remediation), and those who chose to defer taking one. Though enrollment and pass rates at any score can only rise over time, discontinuities could grow or shrink depending on whether these rates rise faster on either side of a cutoff.

We also examine the impact of the FCCRI on several other longer-term outcomes for both cohorts. First, we measure persistence by whether M1 students enrolled in a postsecondary institution during 2013/14 were still enrolled during 2014/15 or 2015/16 and by whether M2 students enrolled during 2014/15 were still enrolled during 2015/16.[12] Second, we say that students transferred if they first

[12] Since our data do not cover out-of-state students, those who transfer out-of-state are not counted as persisting.

enrolled exclusively at a two-year school but later enrolled at a four-year school.[13] Third, we observe whether an individual received any sort of credential. In many cases, these will be certificates and two-year degrees; several higher-level degrees are also listed, but these are likely limited to a small group of students who completed a substantial number of college credits while still in high school. As a result, we do not separate degrees by type.

Running variables

Interpretation of treatment will vary depending on the assessment (FCAT/EOC or PERT), though the FCAT/EOC treatments have different effects at each cutoff. At FCAT/EOC cutoffs in both subjects, treatment consists of being targeted for PERT testing. The treatment group is the one being targeted by an FCCRI provision, meaning that the treatment group in either subject is above the low FCAT cutoff but below the high FCAT cutoff. When grade 11 PERT scores are the running variable, treatment is below the cutoff and consists of assignment to CRS courses. When grade 12 PERT scores are the running variable, treatment is applied above the college readiness cutoff and consists of placement into for-credit college courses. Because treatments at the upper FCAT margin and for the grade 11 PERT apply below cutoffs, graphs at these margins should be read from right to left.

The four assessments—the FCAT, EOC, grade 11 PERT, and grade 12 PERT—cover the interval between initial targeting and college course placement. Although the grade 12 PERT is not part of the FCCRI, including it has two key advantages. First, it allows us to contrast the effects of the FCCRI against those of a known policy with predictable effects on student outcomes. Second, it allows us to see how many students whom CRS courses prepare for college (in the loose sense of scoring "college-ready" after taking these courses) are actually capable of passing college-level coursework.

Data and sample

We used student-level data from FLDOE's K–20 Education Data Warehouse. Our initial sample for M1 consisted of all students first enrolled in grade 11 in 2011/12 (N=180,662). We omitted students who transferred to an out-of-state, private, or home school; withdrew from school for medical reasons; or who did not have a school enrollment record (N=11,202). Students must have had valid assessment scores on the FCAT in grade 10 (N=158,330 in math, N=158,658 in reading). We merged in data on race/ethnicity, sex, free or reduced-price lunch (FRPL) status, English language learner (ELL) status, cumulative high school GPA through grade 10, and college outcomes (casewise deletion of students with missing background data gives N=145,580 in math and N=145,754 in reading).

Our initial sample for cohort M2 consists of all students first enrolled in grade 11 in 2012/13 (N=178,140). We follow the same steps as for cohort M1, dropping those who exit our sample (N=11,133) and keeping those with valid FCAT/EOC scores (N=15,093 in math, N=168,842 in reading).

[13] We exclude students who enrolled in both levels during the first year, as this could reflect co-enrollment rather than transfer. We also ignore transfers among two-year colleges or among four-year colleges. Our measure is therefore a lower bound for the number of actual transfers.

Casewise deletion of students with missing background data gives N=11,336 in math and N=141,538 in reading.

As students may retake any of these assessments, we used initial scores on the PERT (grouped by grade to differentiate scores affecting CRS enrollment from those affecting developmental college coursework). Since our data did not include FCAT testing dates, we used students' lowest FCAT scores as a proxy. While students' highest scores are more likely to be the ones that determine whether they are assigned to take the PERT, initial scores cannot be manipulated through selective retesting.

To address whether retesting is an issue in practice, Table 8 shows retesting patterns by assessment and cohort. Most analyses using grade 11 PERT scores use only students targeted for PERT testing, and most analyses using grade 12 PERT scores use targeted students who scored below college-ready on the grade 11 PERT and then complied with assignment to a CRS course. This applies to the grade 11 PERT for cohort M1 (N=89,145 in math, N=51,268 in reading), the grade 12 PERT for cohort M2 (N=29,246 in math, N=10,043 in reading), and the PERT reading section for cohort M2 (N=52,625 in grade 11, N=12,500 in grade 12). However, because so few students from cohort M2 took the Algebra 1 EOC assessment at the end of grade 10, PERT analyses for cohort M2 do not condition on prior assessment scores or on CRS-taking (N=57,671 in grade 11, N=53,094 in grade 12).

Table 8. Number of students testing multiple times by cohort, assessment, and subject

Number of tests, M1	FCAT Math	Reading	PERT – grade 11 Math	Reading	PERT – grade 12 Math	Reading
1	128,911	101,369	84,922	49,435	23,891	8,503
2	16,648	44,328	3,626	1,640	4,544	1,318
3	20	45	473	164	736	307
4	1	12	94	23	66	13
5	-	-	26	4	9	2
6	-	-	4	2	-	-
Total	145,580	145,754	89,145	51,268	29,246	10,043

Number of tests, M2	Algebra 1 EOC	FCAT Reading	PERT – grade 11 Math[a]	Reading	PERT – grade 12 Math[a]	Reading
1	10,741	83,389	55,561	51,120	41,099	9,873
2	572	58,075	1,953	1,384	8,779	1,975
3	23	60	150	110	2,173	478
4	-	14	6	11	739	131
5	-	-	1	-	222	27
6+	-	-	-	-	82[b]	16[c]
Total	11,336	141,538	57,671	52,625	53,094	12,500

[a] As very few students took the Algebra 1 EOC in grade 10, these counts do not condition on prior assessment performance or CRS enrollment.

[b] Of these 82 students, 62 tested six times, 11 tested seven times, five tested eight times, two tested nine times, one student tested 10 times, and one student tested 11 times.

[c] Of these 16 students, seven tested six times, eight tested seven times, and one tested eight times.

Summary statistics for cohorts M1 and M2 are contained in Appendix A (Tables 1A through 8A). Although results from these tables do not reflect program impacts, demographics follow broad patterns—students from underserved populations (minority students, FRPL students, ELL students, and those with disabilities) are disproportionately unlikely to appear in the highest proficiency levels, whereas Asian students and those ever participating in gifted/talented programs are disproportionately likely to do so. Female students are underrepresented among the highest proficiency levels in math, but are equally represented across the full distribution of reading scores.[14] The underserved populations discussed above are also underrepresented above the passing score on both PERT subjects and in both grades. PERT-taking increases substantially in both subjects at the low FCAT margins. Drops of similar magnitude occur at the high FCAT margins. No other characteristics vary to this extent at any cutoff. CRS enrollment is also broadly consistent with the design of the FCCRI: the probability of enrolling in CRS courses increases sharply, plateaus, and then declines as FCAT scores increase.

Regression discontinuity validity requirements

The What Works Clearinghouse has three criteria for a study to qualify as RD; sharp RD studies must meet four further sets of criteria to meet evidence standards (What Works Clearinghouse, 2015). To qualify as an RD study, policy treatment must be based on a running variable; this is true of the FCCRI. The running variable must be ordinal, with a sufficient number of unique values; the FCAT/EOC and PERT are ordinally scored, with many possible scores above and below each cutoff. Finally, no other policies may be implemented at the same cutoff value.[15] This is certainly true for the FCAT on a statewide level, as there are no statewide policies uniformly affecting students at either of its cutoffs. It may not be true at individual districts or schools, as some might institute interventions for students in PL1, who are at risk for dropping out of high school; however, there is no clear indication of this in our estimates.[16] The high school graduation cutoff is located sufficiently far from proficiency level cutoffs that bandwidth selection should be able to avoid any confounding effects. The grade 11 PERT is more challenging, as it both assigns students to CRS courses and determines college readiness. We cannot separately identify positive effects from well-run CRS courses and discouragement effects from being told that one is not college ready, and can therefore estimate only a net effect. We present an overview of RD validity requirements here; a fuller treatment is available in the interim report (Mokher, Leeds, Harris, & Geraghty, 2016). Supporting tables and figures are presented in Appendix A, and additional details are available upon request.

[14] Female students are equally represented on either side of the math PERT cutoff in both grades.

[15] This is waived if the combined impact of all policy interventions is estimated; however, this often either defeats the purpose of the proposed study or necessitates an impractical number of additional considerations.

[16] Teachers or guidance counselors might use FCAT or PERT scores in making course recommendations, but RD analyses of senior year course-taking patterns do not show discontinuities unrelated to the FCCRI.

Integrity of the running variable

The first validity requirement for RD is that the running variable must be immune to manipulation; for instance, if motivated students systematically avoided CRS courses in favor of more rigorous ones, CRS courses might seem less effective. Manipulation requires that students know the cutoffs for treatment, have incentives to change their running variable values, and have the means to do so.

Although cutoff scores for each assessment are public information, students may not have the motive or means to manipulate their scores. Some students may be indifferent to their course placement, while those who do care may find ways to avoid complying. Some students are not eligible to retest—students on the FCAT can retest only if they score below the high school graduation requirement—and retesting may be too costly or time-consuming for many students who are eligible. Even then, perfect manipulation of test scores is challenging at best. Teachers and administrators might have motivation to manipulate scores, but as they neither take nor grade assessments, doing so would require blatant and unjustifiable falsification of results. The test is administered and scored by independent contractors, who have no incentive to modify scores or treatment statuses. However, to minimize the impact of retesting, we use student's first scores on the PERT in our analysis; since our data do not contain FCAT or EOC testing dates, we use students' lowest scores to proxy for their first scores.

Further evidence against manipulation is presented in density tests in Appendix A, (tables 10A and 11A, and figures 1A and 2A) which do not show students clustered at most FCCRI cutoffs. McCrary density test indicate a discontinuity in the reading section of the grade 11 PERT, but the corresponding graph does not show a clear discontinuity. Nevertheless, reviewers may wish to take estimates using the grade 11 reading PERT for cohort M1 with caution. There is also a small discontinuity at the upper FCAT cutoff in reading for cohort M1, but it is significant only at the 10 percent level, rather than the 5 percent level specified in WWC guidelines. For cohort M2, there is possible evidence of a discontinuity at the upper EOC cutoff; however, graphical analyses show that this may be due to overfitting in our tests, and we do not include this cutoff in our estimates anyway due to low sample size, making the issue moot.

Attrition

For RD estimates to be valid, there cannot be excessive overall attrition or differential attrition by treatment status. Attrition is defined in this study as transferring to an out-of-state, private, or home school; withdrawing for medical reasons; or lacking an enrollment record. Within a narrow range of all cutoffs, both types of attrition are low enough to avoid substantial bias. Overall attrition in both subjects and for both cohorts is less than 10 percent at all cutoffs and the difference between treatment and control groups is 2.5 percentage points or less at all cutoffs (Appendix A, table A12). With overall attrition at 10 percent, the WWC standards' strictest criteria permit up to 6.3 percentage points of differential attrition. As a result, our data are well within acceptable boundaries.

Continuity of outcome-running variable relationship

An effect caused by a policy must be entirely attributable to that policy; absent the policy, the outcome would trend smoothly through the cutoff for treatment. This is impossible to directly verify for any policy but implications of this condition can be tested by showing that baseline covariates do not vary discontinuously at any cutoff and that outcome variables do not have unexplainable discontinuities.

Baseline covariates must be stable at each cutoff to establish the balance of treatment and control groups. Discontinuities will both confound our estimates themselves and make it more plausible that unobserved characteristics are also doing so. The WWC specifies that if differences in baseline characteristics at the cutoff are greater than 0.25 standard deviations, the samples are too dissimilar to conduct an analysis (WWC, 2015). Baseline equivalence must be established for pretreatment achievement and socioeconomic status (WWC, 2015). We use FCAT scores (where applicable), cumulative high school GPA through grade 10, and FRPL status to account for these characteristics.

For cohort M1, we find effect sizes ranging from 0.0000 to 0.1274 standard deviations, shown in Appendix A (Tables 13A through 16A). Since no effect sizes were greater than 0.25 standard deviations, these tests do not invalidate our estimates. However, because 18 out of 184 estimates were both statistically significant and had effect sizes greater than 0.05, our estimates control for a full set of covariates covering race/ethnicity, gender, disability, gifted status, ELL status, native English-speaking status, and poverty.[17] Including relevant variables may decrease standard errors as well.

For cohort M2, 11 estimates had effect sizes greater than 0.25 standard deviations, violating RD criteria (Appendix A, Tables 17A through 20A). However, these estimates were either at the upper Algebra 1 EOC cutoff (where there were too few students for meaningful analysis) or were for seamless college enrollees at the lower Algebra 1 EOC cutoff (where there were too few students to return a coefficient in our impact analyses). As a result, the estimates that violate RD criteria are not presented here. As before, because some of the remaining estimates had effect sizes greater than 0.05 standard deviations, we include a full set of covariates in all of our impact analyses.

A second type of suggestive evidence of the continuity of the outcome-running variable relationship is to show that outcome variables are continuous or explainably discontinuous (i.e. due to known policy interventions using the same running variable) away from any cutoffs. Appendix A (Figures 3A to 12A) provides a graphical analysis displaying the relationship between each of the outcomes used in the analysis and the forcing variables. Unexplainable discontinuities in outcome variables would make it harder to argue that the FCCRI caused discontinuities at program cutoffs. Discontinuities do arise at FCAT cutoffs for high school graduation, but because they represent a major hurdle to college success, discontinuities there are well explained. They are sufficiently far from the low FCAT cutoffs used in this analysis that bandwidth selectors should avoid them if needed.

[17] Although native English-speaker status is closely related to ELL status, they are not perfectly correlated—some students are listed as both ELL and native English speakers, and some as neither. The latter case might represent students whose first languages were not English but who were able to integrate themselves into English-language schooling without requiring additional supports (for example, if they immigrated prior to kindergarten); it is unclear what circumstances would lead native English-speaking students to require additional language support.

Functional form and bandwidth

RD estimation must be conducted using an appropriate functional form and bandwidth. Our analyses use local linear estimation—a fixed functional form with variable bandwidths—allowing the slope of the estimates to vary on either side of the cutoff.

Optimal bandwidth choice reflects mean-variance tradeoffs. Using a narrow bandwidth can lead to large standard errors and imprecise estimates. Using a bandwidth allows individuals far from the cutoff to affect estimates at the cutoff—it will generally return more precise estimates, but these may be biased by nonlinearities in the outcome variable.

There is no single bandwidth selection method—several "optimal" bandwidth selectors exist in the economics literature. We use a cross-validation method initially presented by Imbens and Lemieux (2008) and further explicated by Lee and Lemieux (2010). The selection procedure and its implementation are detailed in Appendix A, under the subheading "computing bandwidths through cross-validation". We estimated effects separately for each subject, assessment, and cutoff.

Summary

All of our estimates at least partially satisfy the standards for integrity of the forcing variable, sample attrition, continuity, and bandwidth/functional form. This means that the study is eligible to meet WWC RDD standards with reservations.

Results from regression discontinuity analysis

We present estimates covering the following sets of effects:

- The intention to treat with PERT testing, based on FCAT scores.
- The intention to treat with CRS coursework, based on grade 11 PERT scores, conditional on being targeted based on FCAT scores.
- The intention to treat with placement into developmental college coursework, based on grade 12 PERT scores, conditional on being targeted based on FCAT scores, scoring below college-ready on the grade 11 PERT, and enrolling in a CRS course.[18]

We begin with year one results for short-term outcomes, first for cohort M1 and then for cohort M2. We then present longer-term outcomes with year two and three results for cohort M1, and year two

[18] For simplicity, the rest of the section refers to PERT estimates without mentioning conditional statements; however, these conditional statements do apply to all results presented below. The sole exception is that PERT results for cohort M2 in math do not condition on prior assessment scores or CRS course enrollment.

results for M2. As mentioned earlier, the results from cohorts M1 and M2 are not directly comparable due to differences over time in high school assessments and postsecondary remediation policies.[19]

Results for short-term outcomes in year one

The impact of assignment to each stage of the FCCRI for cohort M1 is presented in Table 9, broken into panels for math and English. Rows within each panel represent RD cutoffs. Each column represents an outcome—high school graduation (completion of any type of diploma in the cohort's graduation year), seamless postsecondary enrollment (in a public Florida postsecondary institution in the fall semester after graduation), enrollment in a nondevelopmental course in the given subject, and passing a nondevelopmental course in the given subject. For the last two outcomes, we restrict our sample to seamless enrollees in Florida public postsecondary institutions, as these outcomes are most meaningful for this group of students. When estimating the impact on passing a nondevelopmental course, we do not restrict our sample to students who enrolled in these courses; if both enrollment and passing estimates use the same sample, we can directly compare their coefficients. So, for example, if the coefficient on enrollment were 0.10 and the coefficient on passing were 0.05, this would mean that there were half as many new passes as new enrollees.[20] All estimates include baseline covariates; estimates without covariates are available upon request.

Table 9. Regression discontinuity results for the outcomes of high school graduation, seamless college enrollment, and nondevelopmental enrollment and passing, cohort M1

Model Version	Outcome: High School Diploma or Equivalent	Outcome: Seamless College Enrollment	Outcome: Nondevelopmental Enrollment	Outcome: Nondevelopmental Pass
Math				
FCAT, Low Margin	-0.0008 (0.0097) [bw = 20]	0.0076 (0.0126) [bw = 20]	-0.0290 (0.0194) [bw = 19]	-0.0212 (0.0155) [bw = 16]
FCAT, High Margin	-0.0011 (0.0035) [bw = 10]	0.0053 (0.0122) [bw = 10]	-0.0252 (0.0112) [bw = 13]	-0.0239 (0.0164) [bw = 14]
Grade 11 PERT	-0.0024 (0.0033) [bw = 8]	-0.0010 (0.0059) [bw = 19]	0.0100 (0.0098) [bw = 12]	-0.0042 (0.0140) [bw = 13]
Grade 12 PERT	0.0001 (0.0024) [bw = 18]	0.0011 (0.0114) [bw = 19]	0.3057 *** (0.0233) [bw = 14]	0.1812 *** (0.0202) [bw = 16]

[19] Readers should be aware that with 178 regression results shown in the remainder of this chapter, we would expect approximately 18 to be statistically significant at the 10 percent level or higher by random chance alone. A Bonferroni correction for all 178 regressions would require individual results to be significant at a level of p=0.00057 or less; the only results that survive this correction are those for the grade 12 PERT. While a Bonferroni correction is a very conservative approach to solving multiple comparisons issues, all estimates eliminated by a Bonferroni correction of 178 would also be eliminated by a Bonferroni correction of 4.

[20] We cannot say to what extent changes in passing are attributable to students whom the FCCRI induced to enroll in nondevelopmental courses versus those who would have enrolled absent the FCCRI.

Model Version	Outcome: High School Diploma or Equivalent	Outcome: Seamless College Enrollment	Outcome: Nondevelopmental Enrollment	Outcome: Nondevelopmental Pass
English				
FCAT, Low Margin	-0.0027 (0.0052) [bw = 15]	0.0008 (0.093) [bw = 15]	0.0425 ** (0.0191) [bw = 20]	0.0218 (0.0169) [bw = 20]
FCAT, High Margin	-0.0004 (0.0038) [bw = 19]	0.0077 (0.0102) [bw = 19]	0.0048 (0.0081) [bw = 18]	-0.0036 (0.0096) [bw = 19]
Grade 11 PERT	-0.0011 (0.0025) [bw = 20]	-0.0029 (0.0087) [bw = 18]	-0.0225 * (0.0125) [bw = 18]	-0.0025 (0.0098) [bw = 20]
Grade 12 PERT	-0.0014 (0.0042) [bw = 18]	0.0088 (0.0218) [bw = 18]	0.1676 *** (0.0418) [bw = 7]	0.1352 *** (0.0268) [bw = 7]
Sample	Students within bw	Students within bw	Seamless college enrollees within bw	Seamless college enrollees within bw

Note. Results are for cohort M1. All estimates reflect the impact of assignment to treatment—this means the impact of being just above the lower FCAT cutoff, just below the upper FCAT cutoff, just below the grade 11 PERT cutoff, and just above the grade 12 PERT cutoff. Point estimates are listed in percentage points, rather than percentages or elasticities. Standard errors (clustered by district) are in parentheses. Bandwidths (in FCAT or PERT points) are in brackets. N varies by specification – for FCAT, N ranges from 4,607 to 32,810; for grade 11 PERT, N ranges from 22,835 to 67,630; for grade 12 PERT, N ranges from 2,246 to 25,552. * = significant at the 10% level, ** = significant at the 5% level, *** = significant at the 1% level.

There is no statistically significant effect from any stage of the FCCRI on high school graduation, though this may be because students at upper FCAT and PERT cutoffs are highly likely to graduate from high school anyway. Since FCAT scores near the low cutoffs do not satisfy graduation requirements, information about college readiness may be less relevant to students at this cutoff than high school–level remediation. Targeted students at FCCRI cutoffs were equally likely to seamlessly enroll in college as similar students who were not targeted. At the low FCAT cutoff in math, only about a quarter of students seamlessly enrolled in a two-year or four-year college; in reading, the corresponding number is close to 40 percent. At the upper FCAT cutoff in math, close to three-quarters of students enrolled seamlessly; only about 60 percent did so at the upper cutoff in reading. In both subjects, approximately 60 percent of students at the college readiness cutoffs on the grade 11 PERT enrolled seamlessly; on the grade 12 PERT, this number is closer to 50 percent.

As the FCCRI had no statistically significant impact on high school graduation or seamless college enrollment, we restrict our nondevelopmental enrollment and passing estimates to seamless postsecondary enrollees. Selection on outcomes is generally discouraged because it can create differences between the treatment and comparison groups—if the FCCRI had caused many motivated but academically marginal students to enroll in college, college students targeted by the FCCRI might be very different from those who were not. However, since the FCCRI's impact on high school graduation and seamless college enrollment are both statistically insignificant and small in magnitude, balance between the two groups should be preserved. We therefore focus on seamless enrollees for our remaining results.

The outcome variable for course enrollment is equal to one if the first postsecondary course a student took in the given subject was for college credit; it is equal to zero if the student did not take a course in the given subject during their first year or took a developmental course prior to (or concurrent with) a nondevelopmental course. The outcome variable for passing also assigns a value of zero to students who met the above requirements but did not pass. Since there are three reasons a student would have an outcome value of zero—not taking any courses in the indicated subject, taking a developmental course, or failing to pass a nondevelopmental course—not all zero values should be interpreted in the same way. However, defining variables this way allows for a common scale and set of students across outcome variables.

The math FCAT and grade 11 PERT had little effect among students near policy cutoffs on enrolling in or passing a nondevelopmental course. The only statistically significant effects were that seamless enrollees targeted for the FCCRI at the low FCAT reading margin were 4.25 percentage points more likely to enroll in a nondevelopmental English course, while students at the grade 11 PERT reading cutoff were 2.25 percentage points less likely to do so.[21] There was no statistically significant effect on nondevelopmental passing at FCAT or grade 11 PERT cutoffs in either subject. No statistically insignificant estimates were greater than three percentage points, so even if these effects are not true zeroes they are unlikely to be large. For comparison, students who scored college-ready on the grade 12 math PERT were over 30 percentage points more likely to enroll in a nondevelopmental course but only 18 percentage points more likely to pass. This means that for every 10 new nondevelopmental enrollments, there were only six new passes. In English, targeted students were 16.8 percentage points more likely to enroll in nondevelopmental English than their non-targeted peers, but only 13.5 percentage points more likely to pass—for every 10 new enrollments, there were only eight new passes.

Figure 18 shows patterns of enrollment and passing by FCAT and PERT scores for both subjects. In both subjects, enrollment and passing trend smoothly through the FCCRI cutoffs—while there are discontinuities at graduation cutoffs in reading, they are not attributable to the FCCRI and so we do not draw attention to them in the figures presented here.

Table 10 shows the impact of the FCCRI on cohort M2 within the first year of on-time high school graduation. It is organized similarly to Table 9, but with three major differences. First, we do not present results for the high margin of the grade 10 Algebra 1 EOC, as there were too few students near this cutoff for statistically meaningful results. Second, PERT results for math do not condition on prior assessment scores or course enrollment. Finally, since very few students near the low cutoff for the Algebra 1 EOC enrolled seamlessly in college, the regressions used for other estimates cannot identify the FCCRI's impact on non-developmental enrollment and passing.

Unlike for cohort M1, there is one statistically positive impact on high school graduation; however, it is at the grade 12 PERT cutoff, which does not actually reflect the impact of the FCCRI.[22] Like for cohort M1, there is no effect on seamless college enrollment.

[21] However, the former result may be caused by interference from the FCAT's graduation requirement.

[22] This could represent reverse causality if qualifying for for-credit courses motivated students to complete high school. However, this does not explain why the result appears only in English and for cohort M2. Furthermore, if it were due to reverse causality, we would also expect a discontinuity in seamless college enrollment.

Figure 18. Enrollment and passing by assessment score, cohort M1

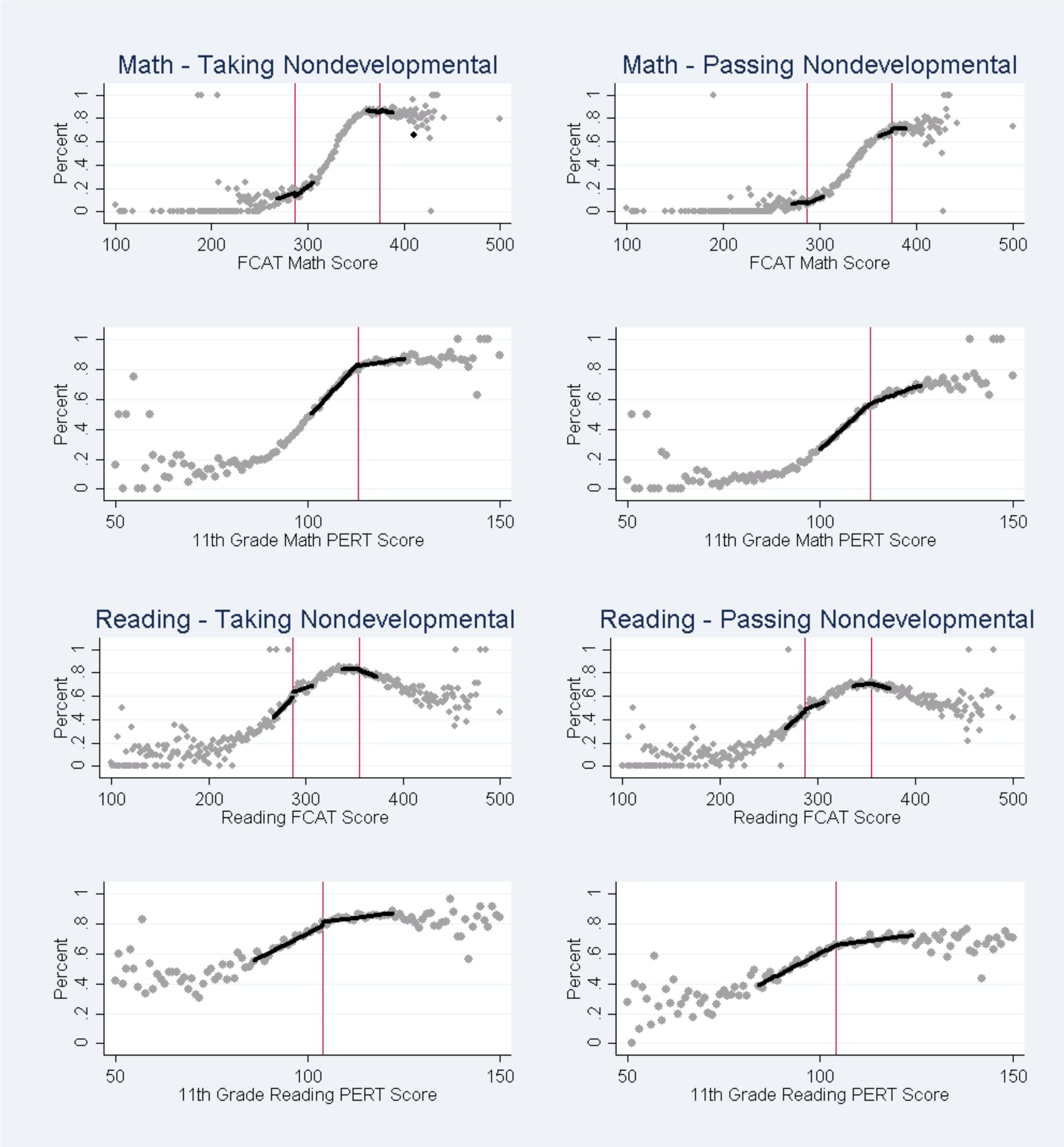


Table 10. Regression discontinuity results for the outcomes of high school graduation, seamless college enrollment, and nondevelopmental enrollment and passing, cohort M2

Model Version	High School Diploma or equivalent	Seamless College Enrollment	Nondevelopmental Enrollment	Nondevelopmental Pass
	Math			
Algebra 1 EOC, Low Margin	0.0123 (0.0263) [bw = 10]	0.0386 (0.0299) [bw = 10]	N/A	N/A
Grade 11 PERT (Unconditional)	0.0043 (0.0048) [bw = 8]	0.0162 (0.0107) [bw = 10]	-0.0228** (0.0108) [bw = 10]	-0.0148 (0.0142) [bw = 9]
Grade 12 PERT (Unconditional)	-0.0001 (0.0036) [bw = 9]	0.0092 (0.0100) [bw = 10]	0.0686*** (0.0118) [bw = 10]	0.0435*** (0.0131) [bw = 10]
	English			
FCAT 2.0, Low Margin	-0.0002 (0.0060) [bw = 8]	0.0126 (0.0123) [bw = 8]	-0.0182 (0.0140) [bw = 8]	-0.0045 (0.0191) [bw = 8]
FCAT 2.0, High Margin	0.0007 (0.0040) [bw = 7]	-0.0081 (0.0097) [bw = 7]	0.0066 (0.0096) [bw = 8]	0.0280** (0.0116) [bw = 8]
Grade 11 PERT	0.0040 (0.0047) [bw = 10]	0.0004 (0.0083) [bw = 10]	-0.0337*** (0.0104) [bw = 10]	-0.0177 (0.0146) [bw = 10]
Grade 12 PERT	0.0137** (0.0058) [bw = 10]	0.0046 (0.0187) [bw = 10]	0.0669** (0.0315) [bw = 9]	0.0552** (0.0259) [bw = 9]
Sample	Students within bw	Students within bw	Seamless college enrollees within bw	Seamless college enrollees within bw

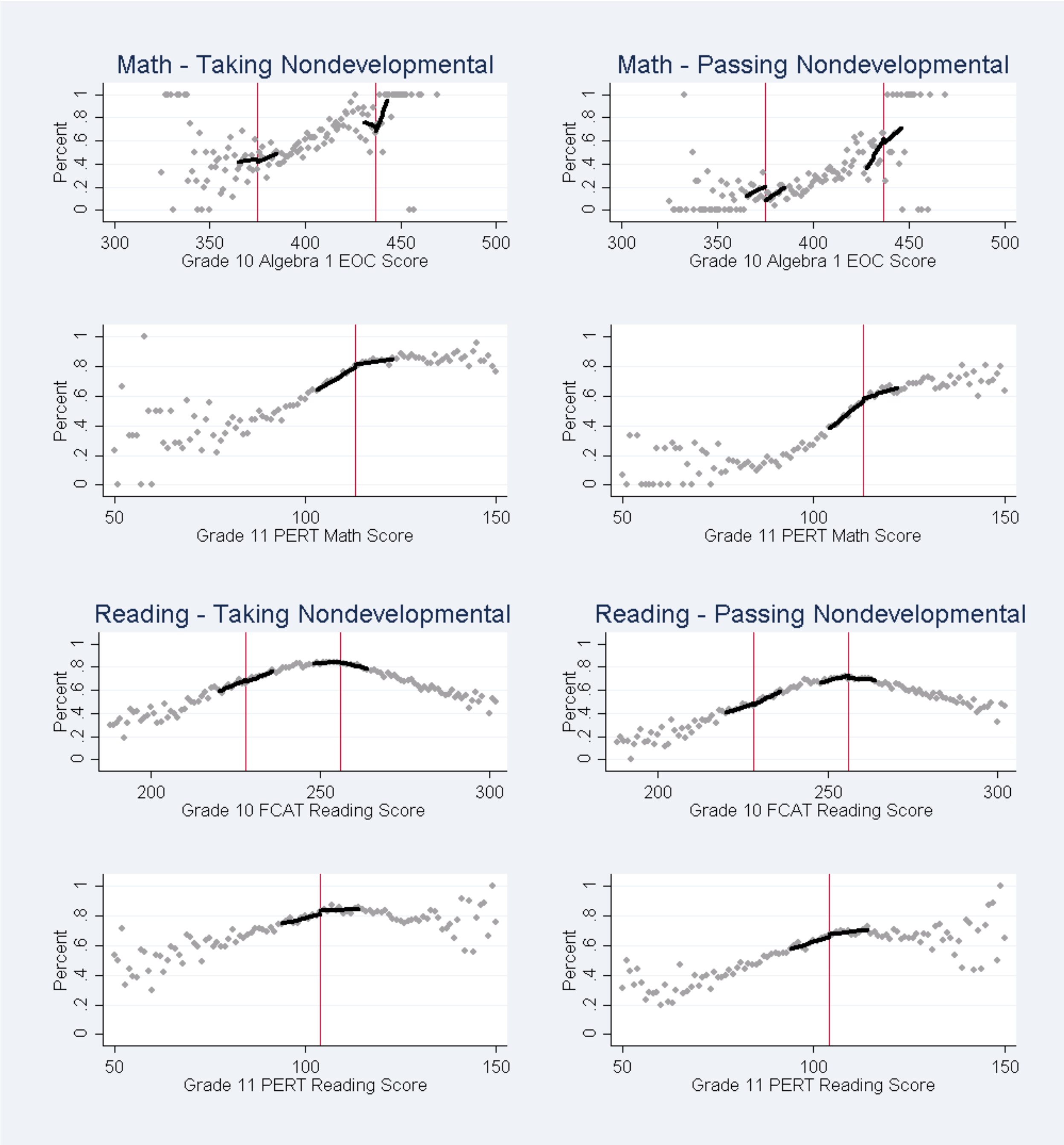
Note. Results are for cohort M2. All estimates reflect the impact of assignment to treatment—this means the impact of being just above the lower FCAT cutoff, just below the upper FCAT cutoff, just below the grade 11 PERT cutoff, and just above the grade 12 PERT cutoff. Point estimates are listed in percentage points, rather than percentages or elasticities. Standard errors (clustered by district) are in parentheses. Bandwidths (in FCAT or PERT points) are in brackets. N varies by specification – for FCAT, N ranges from 4,607 to 32,810; for grade 11 PERT, N ranges from 22,835 to 67,630; for grade 12 PERT, N ranges from 2,246 to 25,552. * = significant at the 10% level, ** = significant at the 5% level, *** = significant at the 1% level.

Targeted students at the PERT cutoffs in each subject were slightly less likely to enroll in nondevelopmental courses—2.3 percentage points less likely in math, and 3.4 percentage points less likely in reading. However, while corresponding point estimates for pass rates are negative (and over half the size of the point estimates for enrollment), they are not statistically significant. The reverse is true at the upper FCAT margin in reading—there is no statistically significant effect on nondevelopmental English enrollment, but there is a 2.8 percentage point increase in passing, meaning that while we do not observe any shifts in course enrollment, students targeted at this margin were more likely to pass courses in which they were already likely to enroll.

Discontinuities for cohort M2 at the grade 12 PERT cutoffs were less than a quarter as large as for cohort M1 in math and less than 40 percent as large in reading. This likely reflects the elimination of developmental education requirements. However, the impact on passing remains a stable fraction of the impact on enrollment—as with cohort M1, for every 10 new nondevelopmental enrollments, there were six new passes in math and eight new passes in English.

Figure 19 shows patterns of enrollment and passing for cohort M2 on the FCAT/EOC and the grade 11 PERT.

Figure 19. Enrollment and passing by assessment score, cohort M2



Results for longer-term outcomes in years two and three

Table 11 shows year two outcomes for M1. In math, no stage of the FCCRI had a detectable impact on persistence to a second year of college. Students narrowly assigned to CRS courses in math were slightly less likely to transfer from a two-year school to a four-year school (a difference of 0.62 percentage points). Neither assignment to take the PERT nor to take a CRS course had an effect on non-DE enrollment or passing or on the probability of receiving a degree after two years. Students who narrowly scored college-ready on the grade 12 PERT were only 5.1 percentage points more likely than those who did not to enroll in a non-DE math class and were not statistically significantly more likely to pass (versus 30.6 and 18.1 percentage points respectively after one year). In other words, students who were not college-ready by the end of grade 12 largely caught up to their college-ready peers in taking and passing for-credit math courses after two years.

In English, students were no more likely to persist to year two or transfer to a four-year institution. Students narrowly assigned to English CRS courses were no more likely to enroll in a non-DE English course within two years but were 1.6 percentage points more likely to pass. For comparison, students who narrowly scored college-ready on the grade 12 PERT were 4.8 percentage points more likely to enroll in a non-DE English course but were not statistically significantly likely to pass. Students at the low FCAT margin were 1.4 percentage points more likely to receive any form of degree; estimates at other cutoffs were both smaller in magnitude and statistically insignificant.

Table 11. Year 2 outcomes for cohort M1

Model Version	Outcome Persistence to year 2	Transfer two-to-four	Non-DE Enrollment	Non-DE Pass	Any degree
Math					
FCAT, Low Margin	-0.0169 (0.0227) [bw = 19]	0.0007 (0.0028) [bw = 20]	-0.0279 (0.0244) [bw = 19]	0.0014 (0.0209) [bw = 19]	0.0051 (0.0036) [bw = 16]
FCAT, High Margin	-0.0138 (0.0094) [bw = 14]	0.0060 (0.0050) [bw = 14]	-0.0058 (0.0100) [bw = 14]	-0.0111 (0.0157) [bw = 14]	0.0126 (0.0091) [bw = 14]
Grade 11 PERT	0.0021 (0.0054) [bw = 20]	-0.0062** (0.0024) [bw = 20]	0.0056 (0.0079) [bw = 15]	0.0035 (0.0084) [bw = 20]	-0.0029 (0.0035) [bw = 20]
Grade 12 PERT	0.0023 (0.0130) [bw = 17]	0.0013 (0.0031) [bw = 20]	0.0512*** (0.0116) [bw = 20]	0.0322 (0.0208) [bw = 15]	0.0102* (0.0053) [bw = 16]
English					
FCAT, Low Margin	0.0095 (0.0136) [bw = 19]	-0.0027 (0.0037) [bw = 19]	0.0121 (0.0147) [bw = 19]	0.0050 (0.0173) [bw = 20]	0.0140*** (0.0044) [bw = 20]
FCAT, High Margin	0.0011 (0.0116) [bw = 19]	-0.0087 (0.0054) [bw = 19]	-0.0020 (0.0084) [bw = 19]	-0.0088 (0.0089) [bw = 19]	-0.0044 (0.0070) [bw = 19]
Grade 11 PERT	-0.0010 (0.0098) [bw = 20]	-0.0023 (0.0047) [bw = 20]	0.0019 (0.0082) [bw = 20]	0.0163* (0.0097) [bw = 20]	-0.0002 (0.0041) [bw = 18]

Model Version	Outcome Persistence to year 2	Transfer two-to-four	Non-DE Enrollment	Non-DE Pass	Any degree
Grade 12 PERT	-0.0037 (0.0232) [bw = 9]	-0.0054 (0.0043) [bw = 20]	0.0475** (0.0182) [bw = 15]	0.0238 (0.0226) [bw = 20]	0.0072 (0.0058) [bw = 20]
Sample	Seamless enrollees	Seamless enrollees	Seamless enrollees	Seamless enrollees	Seamless enrollees

Note. Results are for cohort M1. All estimates reflect the impact of assignment to treatment—this means the impact of being just above the lower FCAT cutoff, just below the upper FCAT cutoff, just below the grade 11 PERT cutoff, and just above the grade 12 PERT cutoff. Point estimates are listed in percentage points, rather than percentages or elasticities. Standard errors (clustered by district) are in parentheses. Bandwidths (in FCAT or PERT points) are in brackets. N varies by specification. *p<.10. **p<.05. ***p<.01.

Table 12 shows year three outcomes for M1. In math, no outcomes were statistically significant at any cutoff. This is unsurprising, as few outcomes were significant after one or two years, and impacts on non-DE enrollment and passing mostly faded by the end of year two. In English, there were no effects at either FCAT cutoff. Students narrowly targeted for English CRS courses were 1.8 percentage points more likely to persist to a third year of college and 1.0 percentage points more likely to be enrolled in a four-year institution after two years. Students who narrowly scored college-ready in grade 12 had outcomes indistinguishable from those of students who narrowly did not.

Table 12. Year 3 outcomes for cohort M1

Model Version	Outcome Persist Anywhere	Enrolled Two-Year	Enrolled Four-Year	Transfer two-to-four	Any Award
Math					
FCAT, Low Margin	-0.0080 (0.0244) [bw = 19]	-0.0019 (0.0104) [bw = 20]	0.0028 (0.0031) [bw = 14]	0.0018 (0.0030) [bw = 19]	0.0110 (0.0152) [bw = 20]
FCAT, High Margin	-0.0092 (0.0136) [bw = 10]	-0.0196 (0.0128) [bw = 10]	0.0110 (0.0169) [bw = 10]	0.0065 (0.0077) [bw = 10]	-0.0145 (0.0159) [bw = 10]
Grade 11 PERT	-0.0014 (0.0083) [bw = 19]	0.0096 (0.0084) [bw = 13]	0.0030 (0.0074) [bw = 10]	-0.0034 (0.0043) [bw = 16]	0.0056 (0.0078) [bw = 19]
Grade 12 PERT	0.0013 (0.0157) [bw = 18]	-0.0080 (0.0103) [bw = 14]	0.0073 (0.0051) [bw = 16]	0.0071 (0.0058) [bw = 14]	0.0207 (0.0139) [bw = 18]
English					
FCAT, Low Margin	0.0072 (0.0175) [bw = 15]	0.0027 (0.0090) [bw = 15]	0.0043 (0.0076) [bw = 15]	0.0034 (0.0055) [bw = 15]	0.0064 (0.0134) [bw = 15]
FCAT, High Margin	0.0138 (0.0107) [bw = 19]	-0.0103 (0.0102) [bw = 19]	0.0150 (0.0095) [bw = 19]	-0.0073 (0.0056) [bw = 19]	0.0139 (0.0115) [bw = 19]
Grade 11 PERT	0.0179* (0.0100) [bw = 20]	0.0001 (0.0085) [bw = 19]	0.0102** (0.0050) [bw = 20]	0.0045 (0.0034) [bw = 20]	0.0022 (0.0120) [bw = 20]

Model Version	Outcome				
	Persist Anywhere	Enrolled Two-Year	Enrolled Four-Year	Transfer two-to-four	Any Award
Grade 12 PERT	-0.0113 (0.0265) [bw = 9]	0.0071 (0.0137) [bw = 17]	0.0030 (0.0080) [bw = 20]	0.0004 (0.0061) [bw = 19]	0.0007 (0.0287) [bw = 9]
Sample	Seamless enrollees	All students	All students	Seamless enrollees	Seamless enrollees

Note. Results are for cohort M1. All estimates reflect the impact of assignment to treatment—this means the impact of being just above the lower FCAT cutoff, just below the upper FCAT cutoff, just below the grade 11 PERT cutoff, and just above the grade 12 PERT cutoff. Point estimates are listed in percentage points, rather than percentages or elasticities. Standard errors (clustered by district) are in parentheses. Bandwidths (in FCAT or PERT points) are in brackets. N varies by specification. *p<.10. **p<.05. ***p<.01.

Table 13 shows year two outcomes for M2. There were no positive effects—the one statistically significant effect was that students narrowly assigned to CRS courses were 0.8 percentage points less likely to transfer from a two-year to a four-year institution.[23] In English, students narrowly assigned to take the reading PERT were 2.8 percentage points more likely to persist to a third year and 2.6 percentage points more likely to be enrolled in a two-year college. However, students narrowly assigned to take a CRS English course were 1.1 percentage points less likely to be enrolled in a four-year institution; the opposite effect holds for students who narrowly scored college ready on the grade 12 PERT.

Table 13. Year 2 outcomes for cohort M2

Model Version	Outcome				
	Persist Anywhere	Enrolled Two-Year	Enrolled Four-Year	Transfer two-to-four	Any Award
Math					
Algebra 1 EOC, Low Margin	0.0018 (0.1032) [bw = 10]	-0.0015 (0.0346) [bw = 10]	0.0029 (0.0030) [bw = 10]	-0.0001 (0.0016) [bw = 10]	0.0125 (0.0332) [bw = 8]
Grade 11 PERT (Unconditional)	-0.0112 (0.0114) [bw = 10]	0.0052 (0.0088) [bw = 10]	-0.0113 (0.0071) [bw = 10]	-0.0083** (0.0038) [bw = 10]	-0.0136 (0.0102) [bw = 10]
Grade 12 PERT (Unconditional)	-0.0043 (0.0146) [bw = 10]	0.0077 (0.0096) [bw = 9]	0.0022 (0.0039) [bw = 7]	0.0043 (0.0026) [bw = 6]	0.0050 (0.0059) [bw = 9]
English					
FCAT 2.0, Low Margin	0.0281* (0.0162) [bw = 8]	0.0257** (0.0123) [bw = 8]	0.0038 (0.0036) [bw = 6]	0.0008 (0.0016) [bw = 8]	-0.0040 (0.0051) [bw = 8]
FCAT 2.0, High Margin	0.0038 (0.0092) [bw = 7]	-0.0050 (0.0091) [bw = 7]	-0.0093 (0.0063) [bw = 7]	-0.0039 (0.0034) [bw = 7]	0.0065 (0.0057) [bw = 7]
Grade 11 PERT	-0.0139 (0.0084) [bw = 10]	-0.0025 (0.0104) [bw = 10]	-0.0107* (0.0059) [bw = 10]	-0.0008 (0.0018) [bw = 10]	-0.0020 (0.0063) [bw = 10]

[23] These results may also incorporate more PERT-takers exempt from CRS courses than in our M1 analyses.

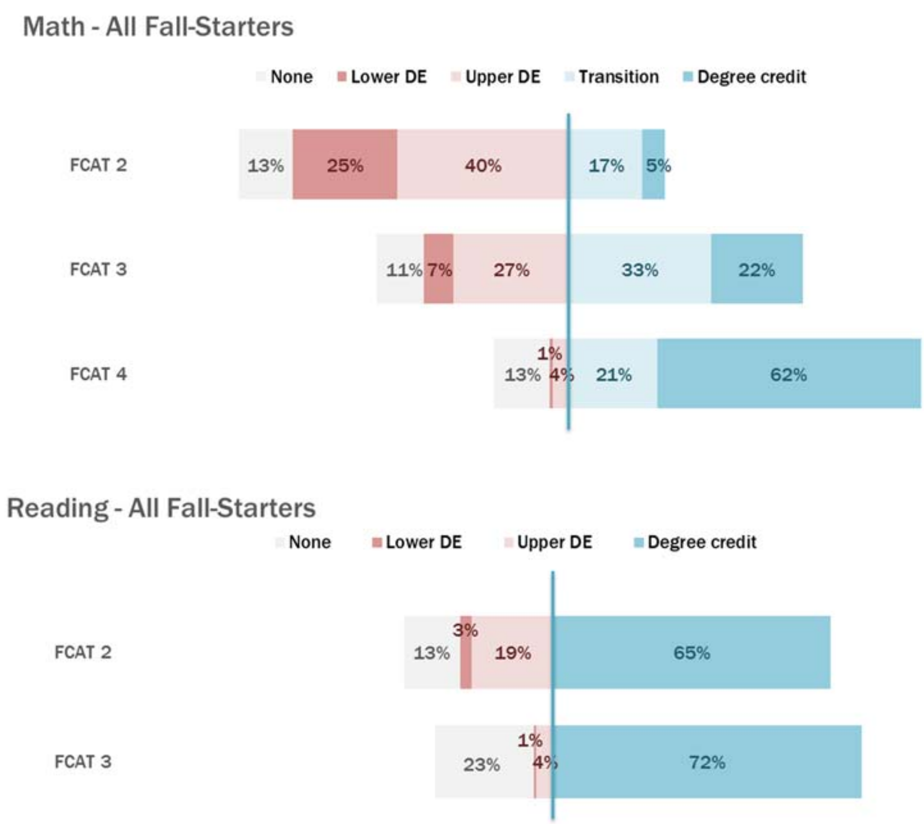
Model Version	Outcome: Persist Anywhere	Enrolled Two-Year	Enrolled Four-Year	Transfer two-to-four	Any Award
Grade 12 PERT	0.0334 (0.0221) [bw = 10]	0.0207 (0.0216) [bw = 10]	0.0110* (0.0061) [bw = 10]	-0.0001 (0.0030) [bw = 10]	-0.0062 (0.0151) [bw = 9]
Sample	Seamless enrollees	All students	All students	Seamless enrollees	Seamless enrollees

Note. Results are for cohort M2. All estimates reflect the impact of assignment to treatment—this means the impact of being just above the lower FCAT cutoff, just below the upper FCAT cutoff, just below the grade 11 PERT cutoff, and just above the grade 12 PERT cutoff. Point estimates are listed in percentage points, rather than percentages or elasticities. Standard errors (clustered by district) are in parentheses. Bandwidths (in FCAT or PERT points) are in brackets. N varies by specification. *p<.10. **p<.05. ***p<.01.

Chapter 5: Exploratory Impact Analysis of the FCCRI Using a Before-After Regression Analysis

This chapter describes the exploratory impact analyses, which use regression analysis with a before-after design to assess the effects of the college readiness and success (CRS) component of the FCCRI for students from a broader range of achievement levels than the regression discontinuity design in the confirmatory analyses. This is important because the likelihood that students will be successful in postsecondary outcomes varies by students' academic achievement. For example, in cohort M1, over 80 percent of FCAT 4 math students who attended college took a transition course or higher, compared with less than a quarter of FCAT 2 students (Figure 20). FCAT 2 students were especially likely to be placed into the lowest level of developmental education (25 percent, compared with only 1 percent for FCAT 4). This suggests that it may be more difficult to help FCAT 2 students become college-ready than other targeted students. There were also differences in college course-taking by FCAT level in English, although these differences were not as large as those in math.

Figure 20. Lowest level of math and English course taken in college for fall-starters in cohort M1, by FCAT level



We hypothesized that the FCCRI may have different impacts on students at different levels of baseline academic achievement. Regression discontinuity analyses can identify effects only for students at the ends of the distribution of academic achievement and cannot uncover these differential impacts. Because of this limitation, we used regression analysis with a before-after design to examine the FCCRI's impacts for a broader range of students. We then compared the results with those from the regression discontinuity analyses. We also explored the extent of differential impacts by students' baseline achievement levels. This analysis does not include M2 because this cohort is not directly comparable to the others due to changes in high school assessments and postsecondary remediation policies. It is also limited to short-term outcomes in the first year following high school graduation due to the changes in the postsecondary remediation policies that occurred in year 2 for cohort M1.

Methodology

Outcome variables

The exploratory impact analysis examined the same short-term outcomes from the confirmatory impact analysis (chapter 4). In addition, it includes a multinomial outcome that allows us to look at the FCCRI's impact on college coursetaking for a wider course-level range.[24] The multinomial outcome captures the level of the first college course a student enrolled in during the first academic year following their cohort's graduation. The multinomial outcome includes five possible outcomes for math and four for English: no course in the given subject, lower-level developmental education, upper-level developmental education, transitional (math only), and degree credit. This allows us to assess the entire course-level spectrum for first year college courses. We also considered the pass rates for the first course taken in math and English, as we would hope to see the pass rates increase or stay the same if students change course levels under the FCCRI.

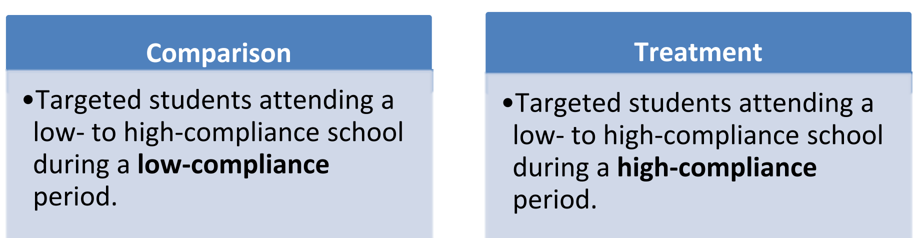
Identification strategy

Selection bias is a concern in this exploratory analysis because students are not randomly assigned to taking college readiness tests and CRS courses. Even in the mandatory period of the FCCRI, compliance was lower than 100 percent, meaning this concern extends to both the voluntary and mandatory cohorts. The RD analysis relied on mandatory assignment and assignment cutoff points to avoid selection bias. In the exploratory analysis we had to develop a different strategy to extend the analysis to a wider range of pretreatment achievement, as we could no longer rely on the cutoff points to form treatment and comparison groups. We used variation over time in school-level FCCRI compliance rates as an exogenous source of assignment to treatment. This yielded an analytical sample where assignment to treatment was conditionally independent from the outcomes of interest.

[24] Binary outcomes are sufficient for analyzing courses at the upper end of the course level spectrum but become less meaningful for courses at the lower end of the spectrum. For example, if we want to examine the FCCRI's impact on enrollment in upper-level developmental education courses, the base group would include students in lower-level developmental education courses and students in degree credit courses; estimates for this outcome would not tell us if students were better or worse off or in more or less challenging courses.

We calculated the school-level compliance rate as the proportion of grade 12 students who were targeted by the FCAT, did not score college-ready on PERT/CPT, and went on to enroll in a CRS course. We defined the treatment group as students who were targeted by the FCAT and attended a high school in grade 12 with at least a 50 percent FCCRI compliance rate (a high-compliance school).[25] The comparison group was limited to students who were targeted by the FCAT and attended a school in grade 12 with less than a 5 percent FCCRI compliance rate (a low-compliance school). We considered high-compliance schools to have implemented the FCCRI, and the low-compliance schools to have not; from this, we obtained a treated group and a comparison group. We limited the analytical sample to students in schools that were categorized as both low- and high-compliance at some point between V2 and M1 (low- to high-compliance schools are in the in-sample group), because we were concerned that low-compliance and high-compliance schools would have differences related to the outcomes.[26] We believe that for almost all of the schools that switch from low- to high-compliance, the switch to being in the treatment group reflects an exogenous change – the FCCRI becoming mandatory – as opposed to some other change in the school. Additionally, the inclusion of school level dummy variables addresses any time invariant, unobserved differences in schools. Essentially, this is a before-after analysis, in which we compared student outcomes before and after schools implemented the FCCRI, and we estimated the intent-to-treat effect for students who were targeted based on their FCAT scores.

Figure 21. Diagram of comparison and treatment groups for the regression analysis



After forming treatment and comparison groups, we conducted an analysis of the sample, presented later in this chapter. We also compared schools that switched from low to high compliance with those that did not to see how generalizable our results were to students attending schools outside of the sample. We also followed WWC (2015) standards and checked the baseline equivalence of the treatment and comparison groups to see if they were similar enough to proceed with the regression. The analysis was designed to meet WWC group design standards with reservations.

[25] The compliance rate was calculated as the percentage of targeted students that had not tested college-ready and took a CRS course.

[26] We included cohorts V2 through M1. V1 was excluded because these students were subject to different high school graduation requirements. M2 was excluded due to changes in high school assessments and postsecondary remediation policies. Schools changed from low to high compliance in different years, so not all of the treatment group was from M1.

Estimation

We used multiple regression analysis to estimate the intent-to-treat effect of the FCCRI. As with the confirmatory analysis, we estimated the FCCRI's effect on high school graduation and seamless college enrollment using all students and estimated its effects on nondevelopmental enrollment and passing using the subsample of seamless college enrollees. For these variables, we used a logit model with maximum likelihood estimation (MLE). For the categorical outcome, we used a multinomial logit model with MLE.

The treatment effect was obtained by estimating β as shown in Equation 1 for individual *i* in school *j*.

$$Y_{ij} = \alpha_i + \beta T_i + \boldsymbol{X_i \theta} + \delta_j + \varepsilon_{ij} \quad (1)$$

Here, the outcome (Y_{ij}) is a function of the treatment status (T_i), a vector of control variables ($\boldsymbol{X_i}$), a school-level fixed effect (δ_j), and an error term clustered by school (ε_{ij}).[27] The set of control variables includes all of the available student-level characteristics that are likely to be related to the outcomes; refer to Table 14 for a summary of the control variables. Functional form was chosen based on scatter plots and the significance of such terms when they were included in the models. All regressors were captured prior to treatment, so the treatment is less likely to have affected them. Intent-to-treat effects were derived from estimated coefficients for the treatment status indicator ($\hat{\beta}$) and are presented as predicted probabilities for ease of interpretation.

Table 14. Control variables included in regression analysis models

Characteristics	Variables Included
Student background characteristics	Race indicators Free and reduced-price lunch (FRPL) status English language learner (ELL) status Native language Special education status Sex
Pretreatment achievement	Grade 10 math FCAT score Grade 10 math FCAT squared Grade 10 math FCAT score interacted with the treatment status (math model only) Grade 10 reading FCAT score Grade 10 reading FCAT score squared Grade 10 reading FCAT score interacted with the treatment status (English model only) Cumulative high school GPA (through grade 10)

[27] We used the students' grade 12 school for both the school-level fixed effect and the clustered error term.

Data and sample

We again relied on data from FLDOE's K–20 Education Data Warehouse. Our sample for this portion of the analysis used all targeted students from cohorts V2 through M1 who met the definition of either the treatment or the comparison group. Models were estimated separately for students targeted in math and English. The sample formation for each of the different model versions is summarized in Table 15. The analytical sample was reduced a bit further because of missing data, and it varied slightly across outcomes for the full sample analysis as some students had a record of enrolling in college but lacked college transcripts.[28] The final full sample size ranged from 140,567 to 147,302 in math and 150,733 to 157,646 in English, and the subsample size was 69,718 in math and 76,772 in English.

Table 15. Sample formation, by targeted subject

	Math				English			
	Full Sample		Subsample		Full Sample		Subsample	
	Student	School	Student	School	Student	School	Student	School
Starting sample (all targeted V2-M1)	361,506	474	171,260	473	250,247	474	122,348	473
Sample remaining after defining treatment/ comparison groups	148,077 (41%)	218 (46%)	69,922 (41%)	217 (46%)	158,453 (63%)	321 (68%)	76,998 (63%)	320 (68%)

Sample analysis

Comparison of characteristics by sample inclusion criteria

To determine the extent to which our sample was representative of the population it was drawn from, we compared student- and school-level characteristics of students in and out of the analytical sample. First, we looked at student-level mean differences in baseline achievement and FRPL status across sample inclusion status (Table 1B, Appendix B).[29] Most of the differences were small in magnitude. The largest difference was in the percentage of FRPL students, where we found that the mean of the out-of-sample group in math was 5.9 percentage points lower than that of the in-sample group. Second, we compared the characteristics of schools that were in and out of the sample based on our restriction of the sample to schools that switched from low to high compliance, and found that most of the mean

[28] We kept this group of students in the sample for the high school diploma and seamless college enrollment outcomes, but they were dropped for the rest of the outcomes.

[29] We present the results only for the full sample, as the subsample results are similar.

differences were small in magnitude. The largest differences were in school locale, as the in-sample group had more suburban/town schools and fewer city schools than the out-of-sample group.

Baseline equivalence of analytical sample

Though the sample of schools is the same for the treatment and comparison groups, the sample of students differs. Because of this, we also used the WWC (2015) standards to verify whether the treatment and comparison groups were equivalent at baseline. Table 2B of appendix B presents the baseline equivalence analysis for the different samples we used for the regression analysis. Our sample meets the WWC's baseline equivalence standards. All of the standardized mean differences are well below the 0.25 standard, and almost all are less than 0.05. The largest standardized mean difference was -0.088, which was for FRPL with the seamless college enrollment subsample. These results suggest that our treatment and comparison groups are equivalent across observed baseline student characteristics. We included all of the available baseline characteristics as control variables in our regression analysis to ensure that we controlled for the small differences between the treatment and comparison groups.

Examination of compliance rates

To better understand what treatment the students in our sample experienced, we examined how college readiness testing and/or CRS course-taking rates varied across treatment status (Table 16). In both subjects, over 80 percent of the treatment group received some type of treatment, and over 80 percent of the comparison group received no treatment at all.

Table 16. Rates of treatment, by subject and treatment status

	Math		English	
	Treatment	Comparison	Treatment	Comparison
No treatment (%)	15	86	16	82
College readiness test only (%)	30	14	35	18
CRS course only (%)	7	0	11	0
Both (%)	48	0	39	0

CRS = college readiness and success. Results are for cohorts V2, V3, and M1. All treatment rates were calculated for the seamless college enrollee subsample.

Estimation Results

All targeted students

We started the analysis by looking at the full sample, or all targeted students that met the definition of either the treatment or the comparison group. We report the estimated marginal effect of the intent to treat on a variety of outcomes in Table 17. Estimates are reported separately for students targeted in math and students targeted in English.

Table 17. Before-after regression estimated marginal effect of treatment for the seamless college enrollee subsample

	Outcome							
Targeted Subject	**High School Diploma**	**Seamless College Enrollment**	**Pass First Course**	**Enroll in Transition-al/Degree Credit**	**Pass Transition-al/Degree Credit**	**Enroll in Degree Credit**	**Pass Degree Credit**	**Complete 3+ College Courses**
Math	0.001 (0.002)	0.002 (0.003)	-0.002 (0.003)	0.032*** (0.003)	0.020*** (0.003)	0.008*** (0.003)	0.005* (0.002)	0.010*** (0.003)
English	0.001 (0.002)	0.002 (0.003)	0.006* (0.003)	N/A	N/A	0.030*** (0.003)	0.024*** (0.003)	0.011*** (0.003)

Numbers reported are the difference in predicted probabilities (the average marginal effect) across intent-to-treat status. Standard errors are shown in parentheses. Results are for cohorts V2, V3, and M1. Models followed a logit specification and included student background characteristics, pretreatment achievement, and a school-level fixed effect as regressors. All students in the sample must meet the definition of either the treatment or comparison group as discussed in the Methodology section. Targeting is determined by the Grade 10 Florida Comprehensive Assessment Test. The sample size for all targeted students is 147,302 in math and 157,646 in English.
*$p < .10$. **$p < .05$. ***$p < .01$.

In the models for all targeted students, we mostly found similar results to those of the regression discontinuity analyses, with little to no effect of the FCCRI on student outcomes.[30] Most of the estimates showed a difference across treatment status that was less than one percentage point. We did, however, find an impact on the likelihood of enrolling in and passing nondevelopmental courses. The treatment group was roughly 3 percentage points more likely to enroll in and 2 percentage points more likely to pass nondevelopmental education courses in both math and English. The positive findings in math were not present in the RD analysis. We believe this difference stems from the designs of the two analyses. The regression sample compared treatment and comparison groups that had dramatically different treatment rates (table 18). The RD analysis on the other hand, had little difference in CRS course-taking rates across the treatment and comparison groups (Figure 16). Further, the regression sample included students from a wider range of baseline achievement as opposed to limiting the sample to students near the FCAT and PERT cutoffs.

Seamless college enrollment subsample analysis

Next, we restricted the sample to students who seamlessly enrolled in college. By looking at this subsample, we no longer dilute the FCCRI's impact on college course enrollment and performance by including students who did not seamlessly enroll in college. Based on the full sample results from

[30] The regression results are most comparable to the RD models that rely on FCAT cutoff scores, so these are the estimates we refer to when comparing the results from chapters 4 and 5.

both the RD analysis and the regression analysis, the FCCRI did not have a detectable impact on college-going; thus, the sample reduction should not bias the estimates.

We began the subsample analysis by re-estimating the outcomes from the full sample analysis (Table 18).[31] Again we report the marginal effect of the intent to treat. The estimates from the seamless college enrollment subsample showed larger differences across treatment status than those of the full sample analysis; however, most of the effects were still small. The biggest difference was in enrollment in transitional or degree credit math courses, where the treatment group was 5.7 percentage points more likely to enroll in such courses than the comparison group. There were also positive effects in English, where the treatment group was 3.5 percentage points more likely to enroll in a degree credit course. In both subjects, the likelihood of passing nondevelopmental courses was also higher in the treatment group than in the comparison group, but the difference was smaller than for the course enrollment outcomes. This indicates that not all of the additional students in the nondevelopmental education courses were passing these courses.

Table 18. Before-after regression estimated marginal effect of treatment for the seamless college enrollee subsample

	Outcome						
Targeted Subject	**Pass First Course**	**Enroll in Transition-al/Degree Credit**	**Pass Transition-al/Degree Credit**	**Enroll in Degree Credit**	**Pass Degree Credit**	**Complete 3+ College Courses**	**First Year College GPA**
Math	-0.011** (0.005)	0.057*** (0.005)	0.035*** (0.005)	0.009** (0.004)	0.006 (0.004)	0.017*** (0.003)	0.024** (0.010)
English	0.000 (0.004)	N/A	N/A	0.043*** (0.004)	0.035*** (0.004)	0.018*** (0.003)	0.025*** (0.009)

Numbers reported are the difference in predicted probabilities (the average marginal effect) across intent-to-treat status. Standard errors are shown in parentheses. Results are for cohorts V2, V3, and M1. Models followed a logit specification and included student background characteristics, pretreatment achievement, and a school-level fixed effect as regressors. All students in the sample must meet the definition of either the treatment or comparison group as discussed in the Method section. Targeting is determined by the Grade 10 Florida Comprehensive Assessment Test. The sample size for the subsample of students who seamlessly enroll in college is 69,718 in math and 76,772 in English.
*$p < .10$. **$p < .05$. ***$p < .01$.

We had two reasons to believe that our analysis up to this point was not capturing the full impact of the FCCRI. First, we thought the average marginal effect could be hiding important variation in the treatment effect across baseline achievement, and there is reason to believe the FCCRI may have had such a differential impact. For example, we thought the initiative might have the largest impact for students in the middle of the FCAT distribution, as students at the lower end may be too far behind

[31] Note the outcomes change slightly from the full sample to the subsample. This is because there was not enough variation in the subsample to include high school diploma and seamless college enrollment, as these outcomes were predetermined by the sample reduction. Additionally, the subsample allowed us to look at first-year college GPA; using the full sample got complicated as it was unclear how to code the GPA outcome for students who did not seamlessly enroll in college.

to become college-ready by the time they enroll in college and students at the upper end may become college-ready by this time without any intervention.

Second, the analysis up to this point had focused on how enrollment in nondevelopmental courses changed without considering what happened to upper- and lower-level developmental course enrollment. To address these concerns, we used a multinomial outcome to examine how the FCCRI impacted enrollment in each course level, and we show how the results varied across baseline achievement levels. Figure 22 and 23 show the results for math and English, respectively. We present the results by plotting the predicted probabilities of enrolling in each course level by treatment status and grade 10 FCAT (baseline achievement level). The difference between the two lines is the marginal effect of the intent to treat.

We find quite a bit of variation in the results across baseline achievement. Students in the upper range of FCAT scores (level 4 in math and level 3 in English) gain the least from the FCCRI. This could be because these students did not need any intervention to become college-ready by the time they graduated high school or because they could take a more advanced course if they did not take a CRS course. In English, it seems the FCCRI was effective only for students at the lower end of the targeted range (FCAT level 2).

The multinomial results indicate that the FCCRI reduced developmental enrollment in math. The impact was especially large for students at the lower end of the targeted range (FCAT level 2), where the treatment group had up to a 12.6 percentage point decrease in the likelihood of enrolling in lower-level developmental education courses. Most of this difference was due to students being more likely to enroll in transitional math; however, there was also an increase in the likelihood of taking no math course in the first year of college. The largest overall positive effect was among treated students in the middle of the targeted range (FCAT level 3) where there was a 10.7 percentage point increase in the likelihood of enrolling in non-developmental (transitional or degree credit) courses.

The change to degree credit course enrollment in math was small across all FCAT levels. This is not surprising, as students were targeted for CRS courses if their test scores would have placed them into a developmental course. Students with placement scores at the transitional level would be exempt from taking a CRS course; thus, the FCCRI is trying to move students from developmental courses to the transitional level.

Developmental education course enrollment was also lower for the treatment group in English, though the differences were smaller. The main change is moving FCAT level 2s from upper-level developmental education to degree credit courses, where FCAT level 2 treatment students had up to a 5.8 percentage point decrease in the likelihood of enrolling in upper-level developmental courses. Over 80 percent of FCAT level 3s in both the treatment and comparison groups enrolled in degree credit courses, which suggests that many of these students became college-ready by the time they enrolled in college regardless of whether they received any intervention.

Figure 22. Math course-level enrollment estimation results, by treatment status and FCAT level

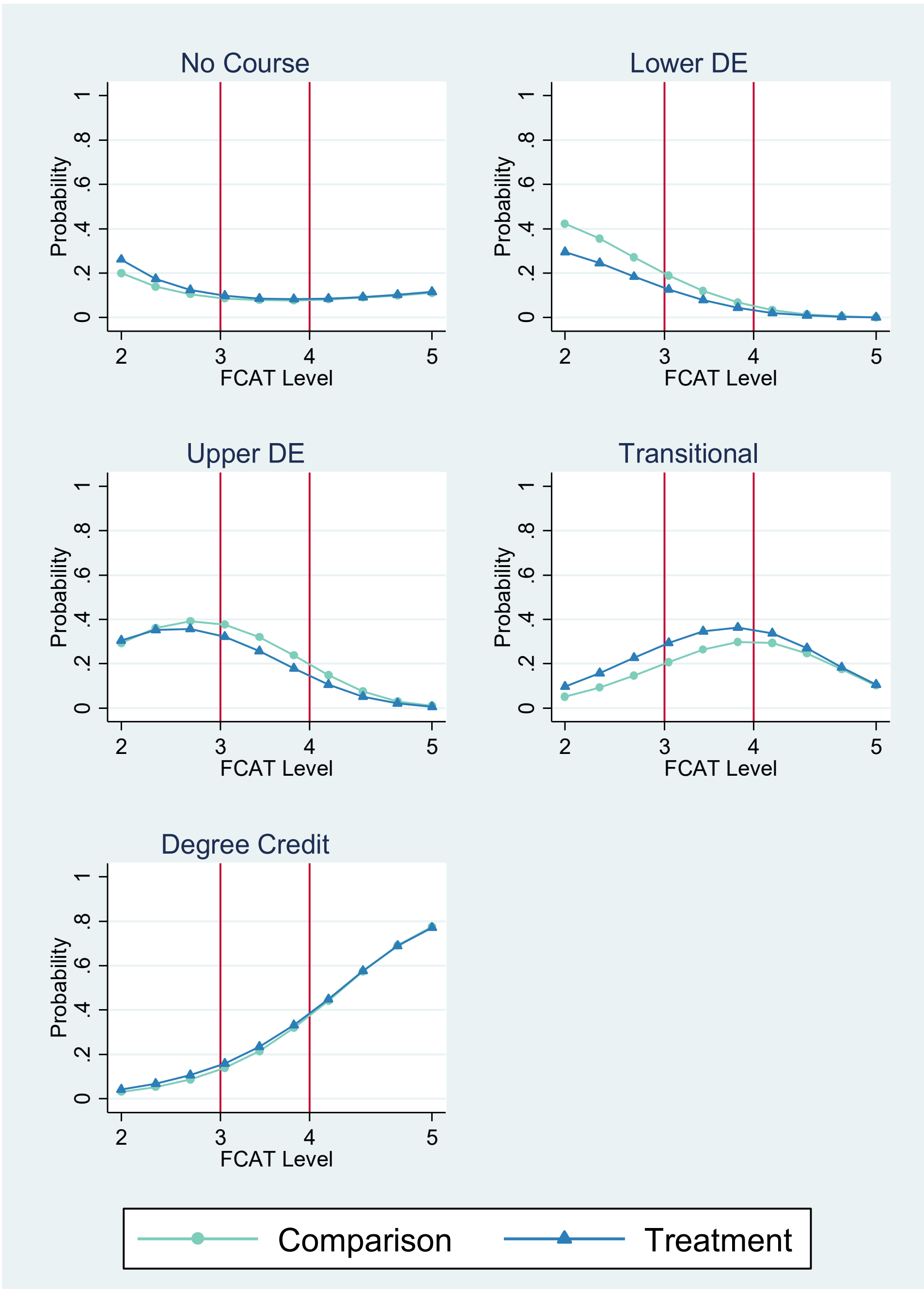


Figure shows the predicted probability of enrolling in each course level, math. FCAT = Florida Comprehensive Assessment Test; DE = developmental education. Dashed lines are drawn between FCAT levels. Results are for the seamless college enrollee subsample in cohorts V2, V3, and M1, with N = 69,718 in math and N = 76,772 in English. Models followed a multinomial logit specification and included student background characteristics, pretreatment achievement, and a school-level fixed effect as regressors.

Figure 23. English course-level enrollment estimation results, by treatment status and FCAT level

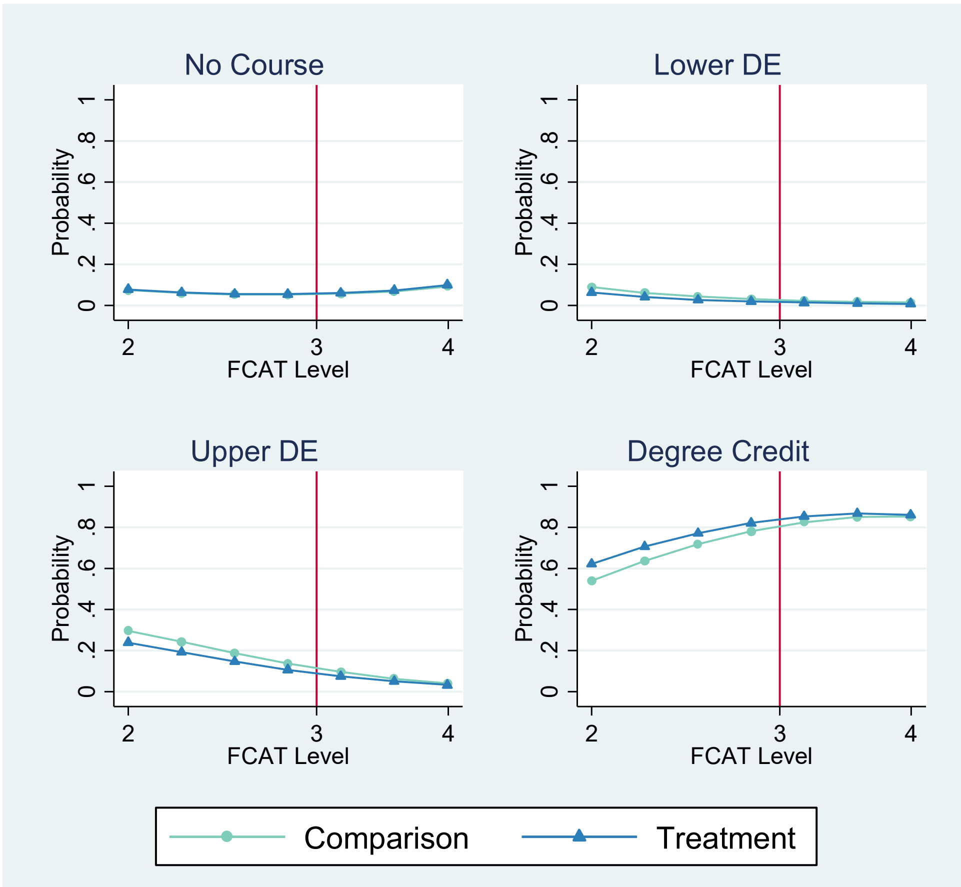


Figure shows the predicted probability of enrolling in each course level, English. FCAT = Florida Comprehensive Assessment Test; DE = developmental education. Dashed lines are drawn between FCAT levels. Results are for the seamless college enrollee subsample in cohorts V2, V3, and M1, with N = 69,718 in math and N = 76,772 in English. Models followed a multinomial logit specification and included student background characteristics, pretreatment achievement, and a school-level fixed effect as regressors.

Although some students moved to a higher course level under the FCCRI, it is important to consider whether the move was appropriate for the students. In other words, we should consider whether the students were moving to a class they were underprepared for. To further examine this issue, we estimated the FCCRI's impact on the likelihood that students would pass the first course that they enrolled in. If students were adequately prepared, the pass rate should increase or stay the same as students enroll in higher-level courses. We found that the pass rate for the treatment group was on average roughly one percentage point lower than that of the comparison group in math and that there was no difference across treatment status for the pass rate in English (Table 18). We also checked for differences in the pass rates across baseline achievement (Figure 24). There was slight variation across baseline achievement in math, with treatment students at the lower end of the targeted range experiencing up to a 3.6 percentage point decrease in the likelihood of passing the first course they took. The decreased likelihood of passing the first math course disappeared toward the upper end of the FCAT distribution. In English, there was no variation across baseline achievement. As a whole, it appears as though most of the students moving to a higher course level were successfully completing the course.

Figure 24. Pass rates for the first course taken in math and English, by treatment status and FCAT level

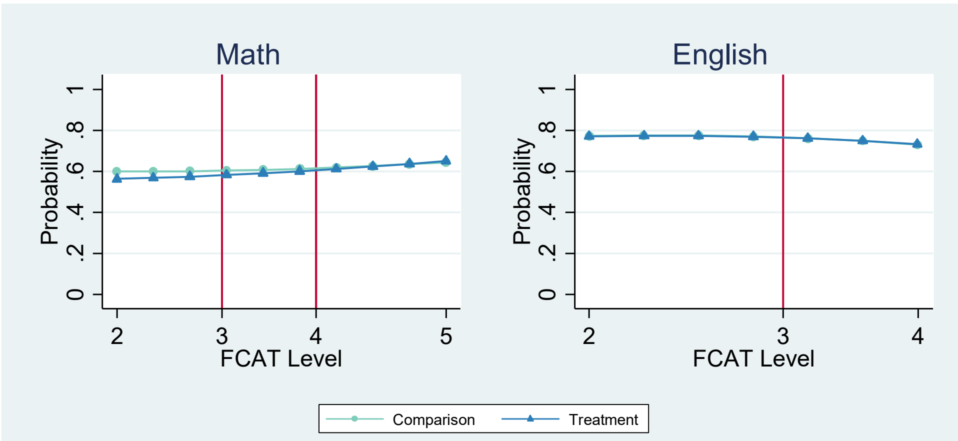


Figure shows the predicted probability of passing the first math and English course taken in college. FCAT = Florida Comprehensive Assessment Test; Red lines are drawn between FCAT levels. Results are for the seamless college enrollee subsample in cohorts V2, V3, and M1, with N = 69,718 in math and N = 76,772 in English. Models followed a logit specification and included student background characteristics, pre-treatment achievement, and a school-level fixed effect as regressors.

The results indicate that the FCCRI demonstrated some success in moving students away from developmental education. Our results imply that the policy should be more narrowly targeted, as students at the upper end of the targeted range of the FCAT distribution saw little to no benefit.

Limitations

There were several limitations to the analysis. First, our results may reveal effects of a different policy shift, thereby making the FCCRI look more effective than it really was. During the first year of college for cohort M1, Senate Bill 1720 made changes to the laws on developmental course-taking. Under this new bill, beginning in the spring semester of 2013 (cohort M1's first year in college), developmental courses were no longer required for recent Florida high school graduates who tested below college-ready on the PERT. With this change, we could see students in this cohort holding off on required developmental courses in the fall because these courses would not be required in the spring. This could make the FCCRI appear to reduce developmental course-taking when in fact the reduction is from this more recent change in policy. We believe the instances of this situation occurring are low, as the new policy was not well known prior to its implementation. Additionally, researchers found little change in developmental education course-taking rates over time in Florida until the fall semester of 2014 (Hu et al., 2016), which corresponds with the second year of college for M1. Through our own analysis, we found developmental course-taking rates during the fall of 2013 to be similar to rates in earlier years. Thus, we believe this policy change is not a big concern for our analysis, but it would become an issue if we were to include later cohorts or a second year of college data.

Second, we estimated the intent-to-treat effect as opposed to the treatment on the treated. Given the way that we defined our sample, over 80 percent of students in the treatment group participated in at least one of the components of the FCCRI, while over 80 percent of the comparison group participated in neither. However, there was still some crossover between the two groups, and only about half of the treatment group and 40 percent of the comparison group participated in both testing and CRS courses. If we estimated the treatment on the treated instead, the estimated impact would be larger.

Third, our results have limited generalizability due to the sample design. Since we limited the sample to schools that switched from low to high compliance when assigning students to the treatment and comparison groups, the results are not generalizable to all schools. The treatment group consists of students at schools where the majority of targeted students who did not test college-ready actually received treatment in the form of taking a college readiness exam and a CRS course. The impact is most likely related to the compliance rate, and thus the results are not generalizable to schools with compliance rates that are lower than 50 percent.

Chapter 6: District Cost-Benefit Analysis

Demand has recently been growing for education policy evaluations to provide information on cost-effectiveness. To provide this information for the FCCRI, we conducted a cost-benefit analysis to complement our evaluation of program effectiveness. This analysis will help education practitioners and policymakers by providing information about the types, amounts, and costs of resources needed to replicate this program, along with the FCCRI's impact on several districts. Similar analyses could be performed in other states if they establish programs akin to the FCCRI.

We compute benefits for short-term student outcomes at program cutoffs using regression discontinuity analysis similar to that in chapter 4, but performed separately for eight districts that participated in our qualitative analysis and provided detailed information on costs. Estimates of the FCCRI's effect on enrolling in and/or passing lower DE, upper DE, or for-credit courses (in Appendix C) let us compute the dollars saved before students at program cutoffs passed a nondevelopmental course. This captures the FCCRI's primary goal of reducing the need for DE. However, one limitation is that the cost-benefit analysis does not consider longer-term impacts, such as on degree completion, or those on students who did not seamlessly enroll in college.

For our cost analysis, we used an approach known as the ingredients method. It requires that the evaluator identify the inputs ("ingredients") used in program implementation, the quantity of each, and the per-unit price of each (Levin & McEwan, 2001). This lets us estimate total program costs and costs per student.

This analysis sought to answer several questions:

1. What were the benefits of the FCCRI, and how do they vary by district?
2. What were the total costs and costs per student of the FCCRI, and how do they vary by district and/or school?
3. How much of the FCCRI's costs were incurred at the state, district, and school levels?
4. How do the costs of PERT testing compare with those of CRS courses?
5. How do the costs of the FCCRI compare with its benefits?

This study focused on recurring annual costs during a single year of the mandatory FCCRI. As the startup period for the FCCRI began in 2007, it was not possible to collect accurate data on startup costs going back that far. We priced ingredients using national prices and did not attempt to adjust for cost-of-living differences between Florida and the rest of the United States. This facilitates estimates for implementing similar programs outside Florida.

The next section, "Cost-benefit methods and data," provides information about how benefits were computed before describing the cost study's sample, data collection methods, and analysis approach. The "Cost-benefit findings" section presents our estimates of both costs and benefits of the FCCRI, including variation by education system level. The final section summarizes our results.

Cost-benefit methods and data

We estimated the benefits of the FCCRI across eight districts using RD analysis and the costs of the FCCRI in nine high schools across six school districts using the ingredients method. Due to the small size of some of the districts that provided cost information, it was not possible to estimate district-level impacts in the same locations. Instead, we estimated benefits from the eight largest districts, where the sample size was sufficiently large for a district-level RD analysis.[32] Two districts were included in the calculation of both costs and benefits. This section explains how this chapter's methodology differs from that in chapter 4 and describes our cost study sample, our data collection process, and the ingredients method. Table 19 lists all districts used at any point in our cost-benefit analyses, the sub-analysis that each is used for, and K-12 characteristics of each district for the 2011/12 school year.

Table 19. Cost-benefit student characteristics by district

District	Study	Students	Asian	Black	White	Hispanic	Other	SWD	ELL	FRPL
Broward	Benefit	258,478	3.5%	39.7%	23.2%	30.6%	3.0%	12.6%	10.6%	57.1%
Dade	Both	350,239	1.1%	22.4%	7.5%	68.3%	0.7%	9.8%	18.3%	71.9%
Duval	Both	125,429	4.2%	43.9%	37.0%	9.9%	4.9%	13.3%	3.8%	52.5%
Escambia	Cost	40,496	2.6%	34.9%	49.6%	5.3%	7.6%	14.2%	1.1%	61.7%
Gadsden	Cost	6,173	0.1%	75.2%	3.3%	19.1%	2.4%	12.2%	6.2%	85.2%
Hillsborough	Benefit	197,041	3.5%	21.4%	35.8%	34.7%	4.6%	14.0%	11.9%	56.8%
Holmes	Cost	3,331	0.6%	3.6%	90.5%	2.9%	2.5%	14.6%	0.1%	65.7%
Orange	Benefit	180,000	4.5%	27.0%	28.8%	36.9%	2.8%	11.0%	13.8%	54.2%
Palm Beach	Benefit	176,901	2.9%	28.6%	33.3%	31.5%	3.6%	15.5%	11.3%	53.5%
Pinellas	Benefit	103,776	4.2%	18.6%	57.4%	15.1%	4.6%	12.9%	5.8%	53.6%
Seminole	Benefit	64,344	4.4%	14.6%	53.3%	23.8%	3.9%	12.8%	4.2%	45.3%
St. Lucie	Cost	39,444	1.7%	30.1%	36.5%	27.6%	4.1%	11.5%	7.7%	62.9%
Full sample	N/A	1,545,652	3.1%	28.3%	28.8%	36.7%	3.2%	12.4%	11.5%	59.0%
Statewide	N/A	2,756,816	2.6%	22.7%	40.2%	30.7%	3.7%	13.1%	9.2%	57.6%

Source: National Center for Education Statistics Common Core of Data (CCD), 2011.

Benefit study methods and data

As before, our estimates use a sharp RD design with a local linear framework. We run RD analyses statewide and by district to determine the impact of being assigned to take the PERT in grade 11 or to enroll in a CRS course in grade 12 for cohort M1. Since RD estimation requires many observations near the treatment cutoff, we restricted our district-level analysis to those with 1,000 or more students

[32] Since CRS course structure and implementation varied even within schools, we cannot directly observe either the nature or quality of FCCRI implementation. While we can observe differences in outcomes, further study would be needed to determine why these differences occurred.

within 20 points of the cutoff being evaluated.[33] Only eight of Florida's 67 LEA districts met that criterion for any combination of cutoff and subject.[34] None met it at the lower FCAT margin in math.

When estimating the probability of enrolling in or passing a given college course level, we add the condition that students have taken a math or English course during their first year of college enrollment. We do, first, because not taking a course cannot be treated as a positive or negative outcome in the same way that passing or failing a course can, as it merely delays a course requirement.[35] Second, this focuses our analysis on when any benefits from the FCCRI would be strongest and when most students would take remedial courses.

We use the cross-validated statewide bandwidths from chapter 4 in our district regressions (outcomes not included in that chapter have bandwidths computed using its methodology) to ensure that district-level variation cannot be attributed to differences in bandwidth.[36] If we instead determine optimal bandwidths for each district and cutoff, our results remain broadly similar, though several estimates gain or lose statistical significance.[37]

The benefit analysis sample contains eight of the 12 largest districts in the state due to sample size requirements. They have a wide variety of students—from 42.6 percent minority in Pinellas to 92.5 percent in Miami-Dade, and from 45.3 percent FRPL in Seminole to 71.9 percent in Miami-Dade. Their

---

[33] Because our bandwidth selector evaluated bandwidths between five and 20 points, some regressions may have sample sizes smaller than 1,000 if a smaller bandwidth was selected.

[34] The FAU Lab School, FSU Lab School, FAMU Lab School, and UF Lab School were included in our data (despite not being full districts), but did not meet the above requirements. The Florida Virtual School and the Florida School for the Deaf and Blind were not included in our data.

[35] Of course, students do not randomly opt out of taking courses in a given subject. This decision is likely correlated with perceived ability and classroom experiences. If students who had bad experiences in CRS courses were more likely to opt out of the corresponding college subject, then our estimates would reflect students who had good experiences in CRS courses and would probably be biased in favor of positive results. However, students at FCCRI cutoffs did not systematically opt out of college math or English during this time frame. Of the 43 district-cutoff combinations we investigated, only four results were statistically significant. Targeted students at the upper FCAT math cutoff in Broward were 5.0 percentage points less likely to take a math class, targeted students at the lower FCAT reading cutoff in Hillsborough were 7.9 percentage points more likely to take English, targeted students at the PERT math cutoff in Palm Beach were 5.1 percentage points less likely to take math, and targeted students at the PERT reading cutoff in Miami-Dade were 4.2 percentage points more likely to take an English.

[36] Because we impose additional sample restrictions, statewide results differ slightly from those in chapter 4.

[37] The most extreme examples of this are:

- The effect on passing a for-credit math class at the upper FCAT margin in Palm beach falls by 4.1 percentage points and becomes statistically insignificant.
- The effect on passing a for-credit math class at the PERT cutoff in Hillsborough falls by 2.8 percentage points and becomes statistically insignificant.
- The effect on enrolling in a for-credit math class at the PERT cutoff in Miami-Dade rises by 2.8 percentage points and becomes statistically significant at the 1 percent level.
- The effect on passing a for-credit math class at the PERT cutoff in Seminole falls by 5.4 percentage points and becomes statistically significant at the 10 percent level.

special education rates and FRPL rates are largely in line with the state average, and their minority and ELL rates are slightly higher than higher than the state average (though not for every district).

Cost study sample

We collected cost data by interviewing program administrators at the state, district, and school levels. Other approaches, such as relying on budget and financial documentation, may not account for all expenses associated with an initiative (such as facilities) or link specific resources to the program. For example, textbook budgets do not identify the courses for which each is used. Similarly, budgets do not link personnel hours—which often account for a large share of program costs—to specific activities. Finally, district budgets may not include items purchased by schools.

Because Florida comprises 67 school districts, we could not collect detailed qualitative data from every district statewide. Therefore, we spoke with staff from districts and schools visited for the implementation evaluation (Mokher et al., 2013, Mokher & Jacobson 2014). We also interviewed FLDOE staff familiar with resources required to develop and administer the PERT and with the curriculum standards for CRS courses.

We originally planned to conduct interviews at the state level, in six districts, and at two schools within each district (the "original sample" for the qualitative analysis). We were unable to obtain school-level cost information from either school from Miami-Dade County or from one school in Holmes County. Therefore, our school-level cost estimates are based on a sample of 9 schools (the "analytic sample").

Districts and schools were chosen for the qualitative analysis to provide a range of district sizes, urbanicities, student demographics, performance levels on state tests, and numbers of students targeted for the PERT. We selected a stratified sample of schools where school size and performance defined the strata.[38] Our goal was to group high schools so that outcomes and implementation problems were likely to be similar within each group but different across groups.

Our sample consisted of two large districts, two medium districts, and two small districts (Table 20). The large districts, Duval and Miami-Dade, enrolled more than 30,000 and 95,000 high school students respectively in 2011/12. The medium districts, Escambia and St. Lucie, each enrolled just over 10,000 high school students. The two small districts, Gadsden and Holmes, each enrolled just over 1,000 high school students. The districts had a wide range of minority enrollment (from 9.5 percent in Holmes to 92.5 percent in Miami-Dade), ELL enrollment (from 0.1 percent in Holmes to 18.3 percent in Miami-Dade), and FRPL enrollment (from 52.5 percent in Duval to 85.2 percent in Gadsden).

[38] *School size* was defined as total high school enrollment. *School performance* was defined as the percentage of students with scores of level 1 or level 2 on the grade 10 FCAT.

Table 20. Characteristics of districts included in the cost study

District	Size Category	Number of High Schools	Total High School Enrollment
Duval	Large	19	30,653
Escambia	Medium	7	10,745
Gadsden	Small	2	1,366
Holmes	Small	4	1,676
Miami-Dade	Large	51	95,469
St. Lucie	Medium	6	11,875

Source: National Center for Education Statistics Common Core of Data (CCD), 2011.

Nine schools from the original sample were Title I schools. The percentage of FRPL students ranged from 37 to 87 percent, and the percentage of minority students ranged from 7 to 100 percent. They also represented the full range of performance on the statewide accountability report card.

Data collection and analysis

The data required to implement the ingredients approach include the list of ingredients, the quantity of each ingredient used, and the per-unit price of each. We therefore report on actual implementation costs, rather than program budgets or planned costs. For example, we report the costs of the actual number of targeted students who took the PERT rather than costs under perfect compliance for all targeted students to participate.

Identifying the ingredients

We began by identifying all ingredients required for PERT testing and CRS courses at the state, district, and school levels, based on previous qualitative data collection and knowledge of program implementation. Resources were organized as personnel, facilities, or materials/equipment. The ingredient list was modified as necessary based on input provided during interviews with state, district, and school stakeholders.

We chose not to include costs of time spent preparing to teach CRS courses. Such costs are difficult to quantify, vary by teacher, and would already have been accounted for if preparation took place during scheduled planning periods. Without CRS courses, teachers would have taught other courses and would have had to dedicate preparation time for them.[39] Similarly, we did not include the time students invested in the PERT or CRS courses, since participation in school activities is mandatory.

Quantifying the amounts of ingredients used

The second step in the analysis was to identify quantities of all ingredients. Quantity estimates for most ingredients were based on interview responses, focusing on ongoing expenses rather than start-up costs. Several researchers interviewed state, district, and school staff using an interview protocol

[39] We do not have data on who taught CRS courses. First-time CRS teachers might have required more preparation time than more experienced CRS teachers or non-CRS teachers. However, this additional preparation time would be better described as a startup cost rather than an ongoing cost. Since our analysis focuses mainly on ongoing costs, we do not include additional preparation time for first-time CRS teachers in our results.

guide to confirm initial ingredient lists and estimate the amount of each ingredient used in the FCCRI. We spoke with FLDOE's PERT coordinator to learn about resources required to administer the PERT and maintain the statewide database of PERT scores. We also spoke with current and former FLDOE staff familiar with the resources used to develop the PERT, the postsecondary readiness competencies on which it was based, and curriculum standards for CRS courses. Because of high turnover, many of these staff no longer work for FLDOE. Interviews with current FLDOE personnel thus provided context for start-up costs but were not designed to determine quantities or costs of ingredients.

We interviewed central office staff within the six study districts to understand the resources used to administer the PERT and support CRS courses. Respondents in each district mentioned PERT testing coordinators and at least one curriculum specialist who supervised the implementation of the CRS courses. At two high schools within each district, we interviewed the staff member who coordinated PERT administration. In total, 21 interviews with district and school staff members were conducted in October through December of 2015, with interviewees reporting on the 2014/15 academic year.

Interview protocols asked about the types and quantities of resources needed to implement the FCCRI. We considered personnel time, materials purchased for CRS courses and PERT administration, the cost of using equipment, and the amortized cost of using facilities such as computer labs for PERT administration, test preparation, or professional development.

We supplemented our interviews with administrative data for the cost of PERT administration. Through grants to districts, the state covered the cost of testing (94 cents per administration) for grade 11 students targeted (and thus required to be tested under the mandatory FCCRI) in each subject area. To estimate this cost to the state, we counted the number of PERT administrations for targeted students in the M1 cohort, using student-level administrative records provided by the Education Data Warehouse.

Pricing the ingredients

The third step in the analysis identified the price of each item and the total cost of the FCCRI to FLDOE, districts, and high schools. When exact prices were given in interviews, we used them; respondents could often recall or look up prices of textbooks or software licenses. In other cases, we used pricing data from the Center for Benefit-Cost Studies of Education (CBCSE) Cost Toolkit©, which listed average national prices for personnel time, materials and equipment, and facilities. The toolkit was most helpful for costs of facilities and computers, which respondents were often unable to estimate. We used the price database to estimate salaries of those implementing the FCCRI to avoid asking respondents for information about their earnings. We based personnel costs on qualifications, years of experience, and estimated number of hours devoted to the FCCRI.

National prices allow us to estimate the cost to implement similar programs in a range of locations. Where educational market rates were not available, the CBSCE Cost Toolkit applied shadow pricing. For example, as the cost of school facilities could not be directly computed, facilities estimates were priced using the cost per square foot of commercial real estate amortized over 30 years. As costs were incurred during 2014/15, prices are reported in 2014 dollars. We used the Toolkit to account for inflation when prices were available only for earlier years.

The final step in the analysis was to calculate per-student cost at the state, district, and school levels. We define a student as served by the FCCRI if he or she was targeted based on grade 10 FCAT performance. We compute costs using students in cohort M1 targeted in either math or reading. Students targeted in both subjects were counted only once.

Cost-benefit findings

We begin by showing the impact of the FCCRI on college course-taking outcomes statewide and for all districts with enough students near FCAT and/or PERT cutoffs. We then translate these impacts into benefits, and compare these benefits against the average cost of the FCCRI.

Benefits of the FCCRI

This section explores whether the FCCRI moved students closer to completion of a for-credit course and translates these results into dollar terms to compute gross savings as a result of the FCCRI. Tables showing the impact on enrollment and pass rates by district may be found in Appendix C. We do not include the lower FCAT margin in math, as no district had 1,000 students within 20 points.

While the statewide results on college coursetaking outcomes at the FCAT math cutoffs were all insignificant, the impact of the FCCRI varied considerably by district. In some districts, we estimate that the FCCRI significantly harmed students' college coursetaking outcomes. For example, in Broward, at the upper FCAT margin, 2.1 percent of targeted math-takers switched from for-credit courses to upper DE. The share that passed upper DE increased by 0.9 percentage points, but the share that passed a for-credit course fell by 7.9 percentage points. Thus, those who stayed in for-credit courses were less likely to pass and many of those who switched to upper DE were unable to pass that. Those in Miami-Dade did not switch courses but were 9 percentage points less likely to pass a for-credit course, suggesting that the FCCRI harmed students who would otherwise have been prepared. Those in Orange were 1.4 percentage points less likely to pass lower DE.[40] In other districts, results were more positive. Math-takers in Hillsborough were 2.5 percentage points less likely to enroll in a for-credit course, but the number passing did not change, suggesting that the FCCRI moved students out of courses they were unprepared for. Those in Palm Beach were 9.7 percentage points more likely to pass for-credit math, suggesting that its version of the FCCRI did not affect course selection but did help students pass for-credit math.

Most results for the FCAT cutoffs in reading are null. We find no discernible impact for Broward at either FCAT margin, or for Hillsborough, Orange, or Palm Beach at the upper FCAT margin. English-takers at both FCAT margins in Miami-Dade appear to be harmed by the FCCRI. At the low margin, they are 4.1 percentage points less likely to pass a lower DE course and are 10.8 percentage points more likely to enroll in a for-credit course without being any more likely to pass it. They also shift towards enrolling in a for-credit course at the upper FCAT margin, but are no more likely to pass it. The FCCRI's statewide impact is statistically insignificant at the upper FCAT margin, but is harmful at

[40] These students may be less likely to pass lower DE because they switched into upper DE or for-credit courses (a good result) or were poorly served by CRS courses (a bad result); initial analysis is inconclusive.

the lower FCAT margin; students who mostly would have passed upper DE courses move to for-credit courses without affecting the number of students passing these courses.

There are some statistically significant differences in the impact of the FCCRI in both subjects for students just below the PERT cutoffs who are assigned to CRS courses. Statewide, students are 1.7 percentage points more likely to enroll in for-credit math courses but no more likely to pass. In English, students are 3.0 percentage points less likely to enroll in for-credit courses and 2.9 percentage points more likely to pass DE courses, again suggesting successful rerouting.

There is also variation in program impacts at the district level for both PERT cutoffs. In some districts, the results are negative. Miami-Dade's FCCRI implementation has no impact in English but appears to harm students who would otherwise be college-ready in math: they are no more likely to enroll in a for-credit class but are 5.2 percentage points less likely to pass one. In other districts, the results are more positive. Math-takers in Hillsborough are no more likely to enroll in a for-credit class but are 9.0 percentage points more likely to pass one; there are no statistically significant results in English. Still other districts have a mix of positive and negative results. In Palm Beach, math-takers' course enrollment is unaffected, but they are much more likely to pass a DE course and much less likely to pass a for-credit course, suggesting its CRS courses may help low-achieving students but harm higher-achieving ones. English-takers from Palm Beach are 5.9 percentage points more likely to pass a DE course, 8.4 percentage points less likely to enroll in a for-credit course, and no less likely to pass a for-credit course, suggesting that students who belong in DE courses are being rerouted to them. In the remaining districts (Broward, Duval, Orange, Pinellas, and Seminole) there are no statistically significant results.[41]

These impacts let us compute the change in the expected probability that students at each FCCRI cutoff pass or fail a lower DE, upper DE, or for-credit course. However, to compute the benefit of the FCCRI, we must assign values to each outcome. As the FCCRI's main goal was to minimize DE costs, we focus narrowly on remaining costs before a student completes a for-credit course, presented in Table 21. We do not consider medium-term outcomes such as completion of subsequent courses or longer-term outcomes such as degree completion rates. However, the impacts on these longer-term outcomes were mostly null (see chapter 4) so the results would likely be similar even if these outcomes were included. The estimates include both institutional costs ($189 per DE hour, $176 per non-DE hour) and student costs ($80 per hour) for each remaining three-hour course. It does not penalize failing a course, and therefore represents the minimum remaining cost for each outcome. Additional analyses (available upon request); in which failing a class incurred a penalty of 50% of that class's cost, had broadly similar patterns.[42]

[41] Duval, Pinellas, and Seminole did not meet sample size requirements in reading.

[42] A student who passes an upper DE class and one who fails a for-credit class both still have to take and pass a for-credit class, so the cost of remaining courses is the same for each. However, the first student is in a better position, since failing a for-credit class may affect motivation and will affect GPA; applying a penalty accounts for this. Some CRS personnel might therefore focus more on keeping students from failing for-credit courses than ensuring that they succeed. While comparing strategies and determining the exact difference in outcomes is beyond the scope of this report, attempting to do so did not affect our qualitative conclusions.

Table 21. Remaining costs before completion of for-credit course.

Outcome	Remaining DE	DE Cost	Remaining FC	FC Cost	Total Cost
Fail LDE	2	$1,614	1	$768	$2,382
Pass LDE	1	$807	1	$768	$1,575
Fail UDE	1	$807	1	$768	$1,575
Pass UDE	0	$0	1	$768	$768
Fail FC	0	$0	1	$768	$768
Pass FC	0	$0	0	$0	$0

Source: Calculation based on statistics from the Florida Department of Education (2015).

Notes: Costs combine institutional cost and student cost. Both types were listed in per-hour units; all courses in this analysis were three hours. We do not assign any extra costs for failing a course (e.g. due to anticipating further retaking or dropout).

Combining our district-level results with course-taking costs lets us compute costs saved prior to completing a for-credit course at each RD cutoff. These benefits are contained in Table 22. We focus on cost savings so that positive values (in bold) represent beneficial outcomes. We omit districts that do not meet sample size requirements at a given cutoff. Each row represents a district, while each column represents a cutoff. We present two sets of results—one using RD point estimates regardless of statistical significance level, and one using shrinkage estimators, which downweight imprecise estimates in order to guard against regression to the mean, based on Weimer (2015).[43]

Table 22. Cost savings per student at RD cutoffs

Unadjusted Point Estimates					
	Math		**Reading**		
	Upper FCAT	**PERT**	**Lower FCAT**	**Upper FCAT**	**PERT**
Broward	-$70.94	**$20.72**	-$26.70	-$27.22	-$12.15
Duval	-	**$63.11**	-	-	-
Hillsborough	-$31.75	**$79.30**	-	**$9.68**	-$26.17
Miami-Dade	-$69.20	-$32.74	**$107.50**	**$2.49**	-$3.09
Orange	-$1.18	**$91.76**	-	-$50.66	-$8.74
Palm Beach	**$81.27**	-$7.41	-	**$28.95**	-$51.95
Pinellas	-$38.93	**$57.78**	-	-	-
Seminole	-	-$53.63	-	-	-
Statewide	-$19.02	**$18.44**	**$37.46**	-$5.15	-$8.25

[43] Shrinkage estimators are $\hat{\beta} = \left[\frac{t^2}{t^2+1}\right]\hat{\beta}_{OLS}$, where $\hat{\beta}_{OLS}$ is the point estimate from the initial equation and t is its associated t-statistic. Variance is computed as $Var(\hat{\beta}) = \left[\frac{t^8+6t^6+9t^4}{(t^2+1)^4}\right]Var(\hat{\beta}_{OLS})$.

Shrinkage Estimators					
	Math		**Reading**		
	Upper FCAT	**PERT**	**Lower FCAT**	**Upper FCAT**	**PERT**
Broward	-$55.26	**$15.51**	-$6.18	-$13.01	-$2.68
Duval	-	**$33.14**	-	-	-
Hillsborough	-$28.69	**$59.08**	-	**$3.31**	-$11.69
Miami-Dade	-$53.18	-$23.41	**$104.32**	**$15.45**	**$5.23**
Orange	-$2.69	**$55.51**	-	-$27.65	**$3.23**
Palm Beach	**$68.89**	-$0.64	-	**$11.74**	-$38.45
Pinellas	-$14.00	**$41.35**	-	-	-
Seminole	-	-$41.49	-	-	-
Statewide	-$10.18	**$14.01**	**$35.88**	-$2.21	-$3.93

Source: Calculation based on values in Table 1C through Table 3C.

Notes: For ease of interpretation, values are framed as expected cost savings; i.e. a positive value (in bold) means fewer costs prior to completing a for-credit course, while a negative value means additional costs. The top panel uses point estimates from RD analysis without adjusting for standard errors, while the bottom panel shrinks point estimates according to Weimer (2015). Blank cells did not meet RD sample requirements.

Annual costs of the FCCRI per student

We estimate the total cost of the FCRI to be about $57 per targeted student. Table 23 breaks this cost down by ingredient category and by system level. About 63 percent of per-student costs were at the school level ($36 of $57), about 33 percent at the district level ($19), and about 3 percent at the state level ($2). Personnel expenses were about $39 per student (68 percent), materials and equipment expenses were about $17 per student (30 percent), and facilities expenses the remainder.

Table 23. Per-student annual cost of the FCCRI, by cost ingredient

Cost ingredient	**PERT testing**				**CRS courses**				**Total cost**
	Level			**Subtotal**	Level			**Subtotal**	
	State	District	School		State	District	School		
Personnel	$1	$3	$33	**$37**	$0[a]	$2	$0[a]	**$2**	**$39**
Materials & equipment	$1	$1	$2	**$4**	$0[a]	$13	$0[a]	**$13**	**$17**
Facilities	$0[a]	$0[a]	$0[a]	**$0[a]**	$0[a]	$0[a]	$0[a]	**$0[a]**	**$0[a]**
Total cost	**$2**	**$4**	**$36**	**$42**	**$0[a]**	**$15**	**$0[a]**	**$15**	**$57**

Source: Calculations from cost data.

Notes: Rounding may result in apparent discrepancies between totals and their components. All costs are reported in 2014 dollars.

a. Estimated cost, rounded to the nearest dollar, is less than $1 per student.

We also distinguished between add-on costs and replacement costs. Add-on costs represent new resources allocated to the FCCRI, while replacement costs would have been used elsewhere absent the FCCRI. Add-on costs are added to existing budgets, while replacement costs are opportunity costs—resources unavailable elsewhere when used for the FCCRI.

Add-on costs

The only true add-on cost of the FCCRI was state spending to reimburse districts and schools for the costs of administering the PERT to students. The cost of testing students required to take the PERT, 94 cents per test administration, was allocated to districts by state grants. These funds ultimately were paid to McCann Associates, which administered the PERT.

Based on our interviews, the total cost to the state of administering the PERT was $709,746. However, this figure includes over $500,000 in additional grants to districts for items such as specialized test forms, 12th grade post-tests, dual enrollment, and ACT or SAT testing not required by the original legislation. Since the legislation that established the FCCRI mandated only that targeted students take the PERT once in grade 11, all other testing was done voluntarily, and these other costs were therefore not technically part of the FCCRI. We therefore calculated the total required state costs of PERT implementation using student-level administrative data, pricing these at 94 cents per administration. We estimated 157,964 required sittings for a total of $148,486. This suggests very low add-on costs of implementing a program such as the FCCRI on a broad scale.

Replacement Costs

Personnel costs accounted for nearly 70 percent of total costs and over 90 percent of school-level costs. They were incurred to support PERT administration and CRS course development and course-taking. However, few new personnel were hired for the FCCRI; for example, schools used existing staff to monitor PERT testing rather than hiring new staff to serve as proctors.

Materials and equipment made up 30 percent of FCCRI costs (approximately $17 per student) and 74 percent of district-level costs. Materials and equipment costs included purchases of textbooks for CRS courses, use of computers for PERT administration, and any additional materials required to prepare students for the PERT.

Facilities made up the smallest replacement cost component, about 1 percent of total costs per students. Facilities costs involved resources such as the use of computer labs to administer the PERT and event spaces for training and professional development.

Replacement cost variation across sites

FCCRI replacement costs varied substantially across districts and schools. To examine variation across sites, we calculated site-specific replacement costs that take into account each site's school-level, district-level, and state-level costs. Table 24 displays site-specific costs for the nine sites for which we have cost data. It also provides information on total program cost, computed by multiplying the per-student cost by the number of targeted students at each site. Per-student costs range from a low of $23 at St. Lucie schools 1 and 2 to a high of $312 per student at Gadsden school 2, while total costs ranged from $4,494 at Holmes school 1 to $27,749 at Duval school 1.

Table 24. Replacement cost per student (including school, district, and state level) and total cost across FCCRI sites

School	Number of targeted students	Replacement cost per student				Total replacement cost
		Personnel	Materials & equipment	Facilities	Total	
Duval School 1	230	$84	$33	$3	$121	$27,749
Duval School 2	155	$54	$32	$0[a]	$86	$13,310
Escambia School 1	113	$93	$21	$0[a]	$115	$12,992
Escambia School 2	399	$31	$20	$0[a]	$51	$20,480
Gadsden School 1	132	$51	$3	$0[a]	$54	$7,129
Gadsden School 2	22	$309	$3	$1	$312	$6,867
Holmes School 1	104	$42	$6	$0[a]	$48	$4,994
St. Lucie School 1	475	$23	$0[a]	$0[a]	$23	$10,825
St. Lucie School 2	380	$23	$0[a]	$0[a]	$23	$8,686

[a] Estimated cost, rounded to the nearest dollar, is less than $1 per student.

Source: Calculations from cost data and administrative data provided by Florida Department of Education.

Factors contributing to cost variation across sites

The two schools with the lowest cost per student were located in Saint Lucie County Public Schools, one of the medium-sized districts. St. Lucie adopted textbooks and developed pacing guides for CRS courses and provided professional development during the first year of implementation (but not subsequent years). PERT administration accounted for the majority of the district's annual costs. The district assessment coordinator wrote grant requests, testing scripts for school proctors, and training materials for school testing coordinators. Along with the testing director, she helped schools prepare to administer the PERT. She also created branching profiles to indicate which students would take the PERT. Both staff members attended conference calls hosted by FLDOE.

Costs at St. Lucie's school 1 consisted mainly of staff time and computer facilities. The testing coordinator was responsible for training proctors and determining PERT targeting. Seven guidance counselors supported her as proctors during PERT administration. In addition, a teacher spent approximately eight hours creating a remedial math packet to help students prepare for the PERT. Finally, an administrative assistant completed tasks associated with the PERT, including sending letters to students' families alerting them to the test's importance.

St. Lucie's school 2 tested all of its juniors. The testing coordinator spent little time on pre- and post-test activities and proctored the PERT with two other teachers. This school was unique in holding a PERT "boot camp" for juniors prior to test administration. Two teachers facilitated the boot camp, which was held on two consecutive Saturdays for three hours each day. The boot camp was voluntary, but the testing coordinator said that participants had a high success rate.

The second-highest costs were in Duval County Public Schools' school 1.[44] Duval provided resources including summer professional development for CRS teachers, textbooks for both CRS courses, and

[44] The highest per-student costs were in Gadsden's school 2. However, it had less than a quarter as many targeted students as any other school. Its costs therefore more likely reflect an unusually small base of targeted students rather than budgetary decisions.

software for PERT preparation. Sixty-five teachers participated in annual summer professional development in math and 35 in English. Since they received professional development credit in lieu of salary, there was no direct cost to the district. Responsibilities for Duval's director of assessment included uploading data to McCann, training school test coordinators, administering the PERT to home schooled students, and creating student login credentials at each school. During each summer, she trained 30 school testing coordinators on how to administer and interpret the PERT. She then held monthly check-ins and quarterly meetings with them.

Duval's school 1 reported many more staff hours than other sites for the PERT, attributable to the five days a month that its graduation coach and three proctors each spent on PERT testing. Prior to the test, the PERT coordinator acquired student login information and informed students about the test through classroom visits and tutoring sessions.

Gadsden County Schools' school 1 came closest to the average per-student cost in our sample. Gadsden was one of the two smallest districts in the analysis, and had few resources to invest in CRS courses and PERT testing. Gadsden did not develop a CRS pacing guide for schools. The supervisor of curriculum and instruction supported teachers in developing curricula but did not offer annual professional development. The district did not purchase textbooks, but PERT preparation software for use in CRS courses represented its most significant investment in the FCCRI. Gadsden's school 1 (and other schools in the two small districts) had few resources to invest in the FCCRI; investments consisted largely of staff time and use of computer lab facilities to administer PERT tests.

Cost comparison: FCCRI testing versus community college placement testing

In this section we compare the costs of PERT testing with the costs of remedial placement testing at community colleges (Rodriguez et al., 2014). Using data from three community colleges, Rodriguez et al. use the ingredients method to identify and price inputs used to offer the placement tests that assign students to either college-level or remedial courses in math and reading/writing. Two of the three colleges (all in the same city) used the same computer-adaptive math and integrated reading/writing exams, purchased from outside vendors and customized to reflect state standards. The third college used a different set of exams, with computer-based standardized math and reading tests purchased from outside vendors, and a paper-and-pencil writing exam developed in-house.

Since Rodriguez et al. (2014) report cost per test, we converted our estimates into their format for comparison. They report an average cost per test of $39 (after converting to 2014 dollars).[45] Our average cost per test, $22, is about 44 percent lower, but our results vary from $13 per test at the lowest-cost school to $164 per test at the highest-cost school. For all three cost components, average PERT costs were lower than community college placement per-test costs.

Cost-Benefit Comparison

There are two main challenges in linking FCCRI impact estimates to cost estimates. The first is that there are too few students at any school to estimate school-level benefits, but most costs of the FCCRI are borne at the school level. We therefore cannot distinguish between a case where the FCCRI has no effect at all and a case where positive and negative effects at different schools offset. The second is

[45] They report estimates with and without the cost of student time. We use those without for comparison.

that the only district with both sets of estimates had no statistically significant benefits. We therefore use a range of cost estimates to discuss the FCCRI's net benefit.

Table 25 shows the net benefit of the FCCRI based on the benefits in Table 22 and costs in Table 24; positive outcomes are bolded for ease of interpretation. After subtracting the average cost of the FCCRI ($57 per student), net benefits are for every district meeting sample size requirements at the upper FCAT and PERT cutoffs in reading. Only Palm Beach had positive net benefits at the upper FCAT cutoff in math, and the remaining districts had large, negative outcomes. At the low FCAT cutoff in reading, one district had a positive net benefit and one had a negative net benefit, making it difficult to discern a pattern. At the PERT cutoff in math, districts were evenly split between positive and negative net benefits, though negative values were larger in magnitude than positive ones. If we instead use shrinkage estimators, nearly every outcome is negative.

Table 25. Net FCCRI benefit per student

Unadjusted Point Estimates					
	Math		Reading		
	Upper FCAT	PERT	Lower FCAT	Upper FCAT	PERT
Broward	-$128	-$36	-$84	-$84	-$69
Duval	-	**$6**	-	-	-
Hillsborough	-$89	**$22**	-	-$45	-$83
Miami-Dade	-$126	-$90	**$50**	-$55	-$60
Orange	-$58	**$35**	-	-$108	-$66
Palm Beach	**$24**	-$64	-	-$26	-$109
Pinellas	-$96	**$1**	-	-	-
Seminole	-	-$111	-	-	-
Statewide	-$76	-$39	-$20	-$62	-$65

Shrinkage Estimators					
	Math		Reading		
	Upper FCAT	PERT	Lower FCAT	Upper FCAT	PERT
Broward	-$112	-$41	-$63	-$70	-$60
Duval	-	-$24	-	-	-
Hillsborough	-$86	**$2**	-	-$54	-$69
Miami-Dade	-$110	-$80	**$47**	-$42	-$52
Orange	-$60	-$1	-	-$85	-$54
Palm Beach	**$12**	-$58	-	-$45	-$95
Pinellas	-$71	-$16	-	-	-
Seminole	-	-$98	-	-	-
Statewide	-$67	-$43	-$21	-$59	-$61

Source: Calculation based on values in Table 21.

Notes: Net benefit = benefit (rounded to the nearest dollar) – cost. Values are framed as expected net cost savings; i.e. positive values (in bold) mean fewer costs prior to completing a for-credit course, while negative values mean additional costs. The top panel uses point estimates from RD analysis without adjusting for standard errors, while the bottom panel shrinks point estimates according to Weimer (2015). Blank cells did not meet sample requirements.

While the values in Table 25 show that the FCCRI is most likely to have yielded a negative net benefit at program cutoffs, they do not preclude the possibility of positive net benefits. Using both unadjusted and shrinkage point estimates and standard errors, however, we can compute this probability under the assumption that errors in our coefficient estimates are independent across regressions. Resulting probabilities are contained in Table 26.

Table 26. Probability of Positive Net Benefits

Unadjusted Point Estimates					
	Math		Reading		
	Upper FCAT	PERT	Lower FCAT	Upper FCAT	PERT
Broward	9.6%	29.8%	30.0%	15.2%	11.4%
Duval	-	**52.1%**	-	-	-
Hillsborough	20.3%	**62.1%**	-	31.0%	13.4%
Miami-Dade	13.9%	4.2%	**63.5%**	25.2%	8.8%
Orange	32.8%	**63.0%**	-	16.1%	20.0%
Palm Beach	**59.0%**	20.8%	-	37.3%	5.4%
Pinellas	23.8%	**50.4%**	-	-	-
Seminole	-	10.2%	-	-	-
Statewide	0.8%	2.7%	36.3%	1.1%	0.0%

Shrinkage Estimators					
	Math		Reading		
	Upper FCAT	PERT	Lower FCAT	Upper FCAT	PERT
Broward	15.3%	25.4%	28.5%	16.6%	1.7%
Duval	-	41.3%	-	-	-
Hillsborough	0.2%	**51.0%**	-	0.0%	14.5%
Miami-Dade	19.1%	7.9%	**62.7%**	23.0%	4.1%
Orange	19.0%	49.3%	-	21.2%	8.6%
Palm Beach	**54.1%**	24.6%	-	27.2%	9.2%
Pinellas	3.3%	42.9%	-	-	-
Seminole	-	15.3%	-	-	-
Statewide	2.2%	0.0%	36.6%	0.7%	0.1%

Source: Calculation based on values in Tables 21 and 1C through 3C.

Notes: Values are the probability of positive net savings, assuming that errors are independent across regressions. The top panel uses unadjusted point estimates and standard deviations, while the bottom panel shrinks both point estimates and standard errors according to Weimer (2015). Blank cells did not meet sample requirements.

Some results from Table 25 are reflected in Table 26—given independent and normally distributed errors, positive predicted net benefits mean that we should expect positive net benefits more than 50% of the time. However, Table 26 contains additional information about the full probability distribution; for example, unadjusted predicted net benefits at both the upper and lower FCAT reading cutoffs in Broward are -$84, but the greater amount of statistical noise at the lower FCAT cutoff makes it twice as likely that net benefits are actually positive there as at the upper FCAT cutoff. Similarly, even though statewide net benefits are not particularly large in magnitude, their large sample sizes (and correspondingly low standard errors) make it highly unlikely that they are actually positive.

Finally, Table 27 shows two extreme scenarios to illustrate the range of possible outcomes at each district and cutoff. Both panels of the table use unadjusted point estimates and standard errors. The top panel shows what would happen if the lower bound of each point estimate's 95 percent confidence interval were used in benefit computation, while the bottom panel uses the upper bound. Since errors are unlikely to be perfectly correlated, these do not represent 95% confidence intervals for net benefits; if errors were independent, they would represent over a 99.999 percent confidence interval.[46] Regardless, actual net benefits are highly unlikely to fall outside the values presented.

Table 27. Net benefits at confidence interval boundaries

Lower Bound for All Estimates					
	Math		Reading		
	Upper FCAT	PERT	Lower FCAT	Upper FCAT	PERT
Broward	-$252	-$177	-$481	-$235	-$197
Duval	-	-$249	-	-	-
Hillsborough	-$226	-$146	-	-$193	-$239
Miami-Dade	-$304	-$197	-$316	-$217	-$158
Orange	-$261	-$143	-	-$291	-$237
Palm Beach	-$117	-$238	-	-$185	-$260
Pinellas	-$274	-$167	-	-	-
Seminole	-	-$246	-	-	-
Statewide	-$113	-$80	-$155	-$108	-$105

Upper Bound for All Estimates					
	Math		Reading		
	Upper FCAT	PERT	Lower FCAT	Upper FCAT	PERT
Broward	-$4	**$105**	**$313**	**$67**	**$58**
Duval	-	**$261**	-	-	-
Hillsborough	**$45**	**$190**	-	**$98**	**$72**
Miami-Dade	**$51**	**$14**	**$416**	**$107**	**$38**
Orange	**$144**	**$213**	-	**$76**	**$106**
Palm Beach	**$165**	**$109**	-	**$129**	**$42**
Pinellas	**$82**	**$169**	-	-	-
Seminole	-	**$24**	-	-	-
Statewide	-$29	**$3**	**$116**	-$17	-$25

Source: Calculation based on values in Tables 21 and 1C through 3C.

Notes: Values are expected cost savings if the true impact of the FCCRI were at the lower bound or at the upper bound of all RD estimates' confidence intervals. Both panels use unadjusted point estimates and standard deviations; net benefits using shrinkage values are generally more negative. Blank cells did not meet sample requirements.

All values at the lower bounds presented in Table 27 are negative and large in magnitude. That is, if we have systematically overestimated the effect of the FCCRI at every stage—if even the null values presented in chapter 4 and appendix C are too high—it will have led students to take lower-level courses and fail those courses more often. Even a relatively cheap program backfiring in such fashion

[46] This crucially assumes that there is no penalty for failing a course—that raising the coefficient on enrollment in a given course level does not make any students worse off. If there were a penalty for failure, then higher enrollment in a given course level could be good or bad, depending on pass rates, and lower bounds would be considerably lower.

would produce large negative net benefits. At the upper bounds, most values are positive—if the FCCRI were actually much more effective than we estimate, it would more than justify its relatively low cost. However, even at upper bounds, some values are negative. Students at the upper FCAT math cutoff in Broward would have benefitted from the FCCRI, but not quite enough to counter $57 in expenditure. Similar patterns hold at three of the statewide cutoffs.

Limitations

Most limitations to RD analysis are discussed in chapter 4. However, this section's sample size requirement mean that we cannot compute benefits for small or medium districts, which may face different challenges than the large districts we include.

There are several additional limitations to the cost-benefit analysis. One is that our interview process may not have identified all ingredients needed to fully implement the FCCRI. Another is the (necessarily) small number of observations, which limits the strength of conclusions that we can draw. A third limitation is that respondents may not always recall or be able to accurately estimate costs.

Schools in our cost sample may not be representative—the true average cost could be higher or lower. As our RD sample contains eight of the 12 largest districts in Florida, cost results may also not be comparable to benefit results. However, even if schools used the lowest-cost implementation strategy reported, only two net benefits would switch from negative to positive (four when using shrinkage estimators). If even the third-highest-cost school is most representative, no cutoff in our RD estimates would have a positive net benefit.

While the FCCRI could be beneficial for many students, we estimate cost-effectiveness at the margins at which it was designed to operate. If our cost analysis is representative of schools statewide, the FCCRI would need to move 7.5 percent of course-takers in either subject from passing an upper DE course to passing a for-credit course to have a positive net benefit. Schools at our low cost estimate would need to move only 3.0 percent of students, while those at our high cost estimate would need to move 15.0 percent.

Finally, our analyses assume that failing a course is no worse than passing the next-highest course. This is unlikely to be true in practice, but without additional data on the harm from failing first-year college courses, we would be unable to accurately account for this difference and would likely introduce new sources of error.

Implications

There are separate implications from each stage of analysis. The benefit analysis reveals very different outcomes by district – some students were much closer to completing a degree, some were no closer, and some were harmed. However, these analyses do not say whether different outcomes were attributable to differences in FCCRI implementation, in quality of implementation, in CRS classroom composition, or other factors.

The cost analysis also has implications for states interested in implementing similar initiatives. A similar program can be implemented for only about $1 per student in new costs. However, implementing a program this way has drawbacks. Most of the FCCRI's costs involved reallocating personnel. This meant that staff time was unavailable for other activities, even if those costs did not show up in a budget. Other states should consider whether their personnel are able to take on additional responsibilities and think about the value of the activities staff will have to give up to focus on a program such as the FCCRI.

While facilities and materials and equipment accounted for fewer costs, there may also have been high opportunity costs from their reallocation. For example, several participants mentioned that their computer facilities were already needed for other computer-based tests, classroom activities, student projects and portfolios, and so forth. Again, policymakers must determine what is lost if computing and other nonpersonnel resources are unavailable for other uses.

Florida's funding approach also has implications for program quality. Florida chose to implement the FCCRI largely as an unfunded mandate, resulting in an uneven distribution of resources. For example, several teachers reported having to develop CRS courses on their own because their district did not provide training, textbooks, computer programs, or other resources (Mokher et al. 2013; Mokher & Jacobson 2014). This disparity led to uneven implementation across districts and schools, some of which may not have had the resources to serve all targeted students adequately.

The FCCRI therefore failed to yield a net benefit either statewide or in eight of the 12 largest districts at program cutoffs. One reason for this may be where cutoffs were located—teachers have stated that CRS courses contained too wide a range of student achievement to adequately differentiate instruction (Mokher et al., 2014). Benefits also might have risen as teachers gained experience running CRS courses. However, given the off-budget costs, Florida may simply not have allocated enough funding for districts and schools to implement the FCCRI effectively. Other states interested in implementing similar programs may want to consider alternative funding approaches to ensure program effectiveness, both overall and across sites.

Chapter 8: Conclusions

This study provides a comprehensive evaluation of the FCCRI's implementation, impacts, and costs from extensive qualitative and quantitative analyses conducted over a six-year period. Below we summarize key findings from each component of the evaluation, as well as discuss implications for Florida and other states interested in implementing similar policies.

Implementation analysis

We found tremendous variation across districts and schools in the level of compliance with state requirements for participation and in how the initiative was implemented. When the FCCRI was voluntary, many schools did not offer college readiness testing or CRS courses and student participation rates were very low. Once the initiative became mandatory, almost all schools offered both components of the FCCRI, although there was still some noncompliance with the assignment of students to college readiness testing and CRS courses.

Overall, the level of difficulty of CRS courses varied widely across schools. CRS teachers noted that they were left largely on their own in developing lesson plans and materials. In most high schools, because no person was responsible for developing curricula, pacing guides, and materials, the CRS courses were not taught in a uniform way. We also found that CRS courses had several goals associated with preparing students for life after high school, including helping students test college-ready, develop the academic skills needed in college, develop soft skills needed in college, develop career-related skills, and decide what they want to do after high school. Districts and teachers placed different levels of emphasis on each of these goals as they implemented the CRS courses. In some schools the CRS courses were very similar to the courses taken by students prior to the FCCRI.

We also examined changes over time in the types of courses taken in grade 12 for targeted students at each FCAT level. Higher-performing targeted students were more likely to take CRS courses at the expense of honors-level and other advanced courses, which may negatively influence postsecondary outcomes. Yet lower-performing targeted students may be taking CRS courses in lieu of other regular courses such as liberal arts math. These other courses may not prepare students as well for college-level work, so the CRS courses may be a better option for these students.

Confirmatory impact analysis of the FCCRI

The confirmatory impact analysis used regression discontinuity (RD) methodology to examine the effect of being assigned to take college readiness assessments, being assigned to take college readiness coursework, and being assigned into remedial college coursework (based on grade 12 test scores) for students in cohorts M1 and M2 near FCCRI program cutoffs.

Assignment to participate in college readiness testing or coursework had no detectable effect on high school graduation or college enrollment in either cohort. It had no detectable effect on non-developmental enrollment or passing in math in cohort M1; there was too little data from M2 to determine an effect. In English, students at the low FCAT margin in M1 were 4.3 percentage points more likely to enroll in non-developmental coursework but were no more likely to pass; there was no impact for the highest-performing targeted students. In cohort M2, there was no effect at the low FCAT margin, but students at the higher margin were 2.8 percentage points more likely to pass.

The FCCRI may have had limited effects at FCAT cutoffs because these cutoffs capture a wide range of student ability. FCCRI targeting begins below the score required for high school graduation and ends only where large majorities of students were already testing college-ready. Students unprepared to graduate from high school likely require more than a single course to become college-ready, while those who are already college ready by definition do not need CRS courses.

The FCCRI may have been more beneficial near the college readiness benchmark in grade 11 for M1. Students below college-ready had scores corresponding to developmental courses, while those above college-ready were exempt. Thus, if their performance remained the same between grade 11 and college enrollment, we would expect those just below college-ready to be much more likely to be placed into developmental courses. Yet in math, students just below college-ready were equally likely to enroll in for-credit courses as those who were barely college-ready. In English, students just below college-ready were 2.3 percentage points less likely to enroll in for-credit courses (but no less likely to pass) than those who were barely college-ready. This suggests that students targeted for CRS courses performed comparably in college to students who were already college-ready in grade 11. For comparison, students below college-ready on the grade 12 PERT were 31 percentage points less likely to enroll in and 18 percentage points less likely to pass non-developmental courses in math and 16.8 percentage points less likely to enroll in and 13.5 percentage points less likely to pass non-developmental courses in English than those just above college-ready. In cohort M2, the FCCRI may have improved placement in college-level courses, as students at grade 11 PERT cutoffs were less likely to enroll in non-developmental courses but were no less likely to pass them. This suggests that the FCCRI was more effective at preventing students from failing than at helping them succeed in nondevelopmental courses.

Results for students assigned to CRS courses are striking when considering counterfactuals. On the grade 11 PERT, students who barely scored college-ready were more likely than those who did not to enroll in honors and other advanced courses during grade 12, while those just below this cutoff were much more likely to enroll in standard-level CRS courses. This suggests that students assigned to CRS courses are being compared with a high standard; either they fared well by demonstrating similar performance to peers in rigorous high school courses or it is possible that their peers may have enrolled in inappropriately challenging courses or ones not aligned with college content.

We found little long-term impact of the FCCRI after two to three years following high school graduation for students near the FCAT cutoffs. The only effect at any of the FCAT cutoffs for cohort M1 was a slight increase in degree receipt within two years for students near the low reading cutoff. In M2, seamless enrollees at the low reading cutoff were slightly more likely to persist to a second year of school and students overall were more likely to be enrolled in a two-year institution.

There were more statistically significant results for long-term outcomes at the grade 11 PERT cutoffs, though they were also small in magnitude. In cohort M1, seamless enrollees who were barely targeted for CRS math courses were slightly less likely to transfer from a two-year institution to a four-year one in their second year. Seamless enrollees barely assigned to CRS English courses were slightly more likely to have passed a non-developmental English course within two years and to persist to a third year of college. Overall, students at this cutoff were slightly more likely to be enrolled at a four-year institution within three years. In cohort M2, students narrowly assigned to CRS math courses were slightly less likely to transfer in year 2 and those assigned to CRS English courses were less likely to have enrolled in a four-year school in that time frame.

Exploratory impact analysis of the FCCRI

The confirmatory impact analyses assessed the FCCRI's impact only on students near the cutoff points of the targeted achievement levels for participation in the FCCRI. However, student performance on postsecondary outcomes varied considerably by pretreatment achievement levels. This raises concerns that estimates of the FCCRI's impact might be different if a wider range of students were included in the analyses. For this reason, we conducted exploratory impact analyses to study the same set of outcomes as in the confirmatory impact analyses, while including students from the entire distribution of targeted FCAT levels.

The results from the exploratory impact analysis were mostly similar to those of the confirmatory impact analysis. Among all targeted students, the regression analysis failed to find a positive effect from treatment on the probabilities of receiving a high school diploma or equivalent, seamless college enrollment, or passing three or more for-credit courses. However, the regression results did point toward a positive impact on nondevelopmental course enrollment and pass rates, as the treatment group was associated with an increased likelihood of both enrolling in and passing nondevelopmental courses in both math and English. The RD analysis did not find positive effects on math course enrollment or pass rates. We believe the difference arises from the definition of the treatment group in the regression analysis. The treatment group included only students who attended schools with high FCCRI compliance rates; thus, the regression sample compared treatment and comparison groups that had dramatically different treatment rates. The RD analysis on the other hand, showed little difference in CRS course-taking rates across the treatment and comparison groups.

We also examined how the results for the college course-taking outcomes changed when we constrained the sample to students who seamlessly enrolled in college. The results were similar to the results from the full sample, but the estimates were larger because the impact was not diluted by students who did not seamlessly enroll in college.

As a final line of analysis, we examined how the FCCRI impacted course level enrollment across the full range of college course levels, as opposed to just nondevelopmental courses, by using a multinomial outcome. We also coupled this analysis with a breakdown of the results by baseline achievement level to see if any important differences emerged. Treated students at the upper end of the targeted range of the FCAT distribution saw little to no impact from the FCCRI in both math and English. Differences emerged in the middle and lower range of baseline achievement. We found that treatment students were less likely to enroll in both lower- and upper-level developmental education courses in math and English. In math, this effect was the largest for students at the lower end of the targeted range. FCAT level 2 students experienced more of a decline in lower-level developmental math course enrollment, and FCAT level 3 students experienced more of a decline in upper-level developmental math course enrollment. In English, very few targeted students were enrolled in lower-level developmental courses, so most of the impact was found in upper-level developmental course enrollment. The results indicate that students moved away from developmental courses toward transitional math and degree credit English courses.

Cost-benefit analysis

We estimate the annual cost of the FCCRI (not including startup costs) to be approximately $57 per student in 2014. With an estimated 128,988 students participating in 2011, the total cost of all resources devoted to the FCCRI was approximately $7 million (in 2014 dollars). Since most of these resources were repurposed from other uses, the actual on-budget cost of the FCCRI (consisting of state reimbursements to schools and districts for administering the PERT) was only about $150,000 per year, or just over $1 per student.

About 63 percent of the total cost of the program was incurred at the school level, with district-level costs constituting about 33 percent and state-level costs about 3 percent. Personnel costs, incurred mainly at the school level, were the most expensive component of the program, at about 69 percent of the total. Materials and equipment made up about 30 percent of program costs, and were incurred mainly at the district level. Facilities appear to have made up only a very small proportion of FCCRI implementation costs.

Per-student costs varied across schools and districts. Across schools, estimated total costs (incorporating state- and district-level costs) ranged from $23 to $312 per student. Schools and districts that allocated more resources to the program tended to dedicate these resources to CRS course materials such as planning guides, textbooks, and computer software; providing teachers with professional development; and PERT administration. Schools with relatively few students tested tended to have higher per-student costs and may have had difficulty achieving economies of scale.

CRS costs represented about 26 percent of total program costs, while PERT testing costs represented about 74 percent. Comparison with costs of remedial placement testing at three community colleges located outside Florida (Rodriguez et al., 2014) reveals that FCCRI testing costs were 44 percent lower on average, but with a much larger range.

The FCCRI produced a range of different benefits at the district level—which we define as the amount of money saved prior to passing a for-credit course—based on the districts in which students were enrolled and their preparation levels. Seven of eight districts studied produced positive gross benefits at one or more cutoffs, but only one did at the upper FCAT cutoff in math and none did at the PERT cutoff in reading (two did if adjusting RD estimates to account for large standard errors).

When comparing expected benefits against average costs, net benefits were negative for all districts meeting sample size requirements at the upper FCAT and PERT cutoffs in reading and for all but one at the upper FCAT cutoff in math. Only one of these estimates would switch to positive if we instead used the lowest per-student costs from our cost sample (and none would if using shrinkage estimators); if we used even the third-highest costs, all net benefits at all cutoffs would be negative. Thus, while a program such as the FCCRI can be implemented for an on-budget cost of only $1 per student if substantial resources are repurposed from other uses, additional on-budget costs could be necessary to see widespread net benefits.

Policy implications

Implications for program improvement in Florida

The findings from the evaluation of the FCCRI have led to collaboration between researchers and practitioners to improve the initiative's effectiveness. In this first year of the evaluation, we found that teachers were supportive of FCCRI's goals and had positive views of its effectiveness, despite identifying two important impediments: lack of information about the skills tested on the PERT, which made it difficult to structure high school CRS courses to help students graduate college-ready, and lack of curricular and instructional materials for those CRS courses (texts, exercises, practice tests, pacing guides), which made it difficult to develop effective lesson plans.

The research team found that the materials existed but were not readily accessible to the teachers who needed them. Realizing this, and in cooperation with Florida state and college officials, the research team launched an "FCCRI group" for CRS teachers on the social media site Edmodo, loading the group page with links and downloadable documents collected by the research team. Now the group site contains information on the PERT from FLDOE and PERT test preparation materials from the state's colleges; it also provides instructional resources from state college developmental education and gateway courses, which the high school teachers can use in preparing their CRS courses. The site also offers a virtual community for CRS teachers to collaborate with one another. Within a month of the FCCRI group site's availability, more than 300 educators in districts across Florida had joined the Edmodo group. Nearly three-quarters of the state's 67 county-based districts have at least one teacher who has joined. The rapid growth of the Edmodo group indicates a strong demand for instructional resources for CRS courses. It also suggests that social media can be an effective medium for reaching out to teachers, which state and local policymakers overseeing educational initiatives elsewhere may want to consider.

In year two, we collected additional feedback about continuing needs, which led to collaboration between researchers and practitioners that resulted in a series of professional development forums for K–12 and postsecondary educators. Presentations and activities addressed several topics at these forums, including an overview of FCCRI and recent legislative changes, bridging the high school–college transition through collaboration, strategies for helping students plan for college, college pathways and programs, student skills needed for success in specific college courses, and student support services at state colleges.

Overall, participants at the summer 2015 events found all of the sessions presented at the forums to be useful and relevant to their teaching and counseling responsibilities. In their open-ended survey responses, participants also noted numerous specific examples of how they plan to use information acquired from the sessions that they found most useful. These include sharing information on postsecondary opportunities other than traditional four-year baccalaureate degree programs with students, sharing information on the earnings potential for multiple types of college credentials with students, sharing scholarship and financial aid opportunities with students, and working with school staff to develop action strategies that can be implemented to improve college and career readiness. Participants also provided feedback on additional ways that the state, districts, and schools can improve college and career readiness for high school students.

The FCCRI professional development forums may serve as a model for other districts interested in bringing together K–12 and postsecondary educators to improve students' college readiness. Our final report on the forums was disseminated to all districts statewide, and all presentations and handouts from the forums were made available on the FCCRI Edmodo site so that they could be used by others.

Implications for future policy decisions in Florida

This evaluation is unlikely to have much influence on policy decisions about the FCCRI. Legislation passed in April 2015 under House Bill 7069 eliminated the requirements for common placement testing and postsecondary preparatory instruction. Participation in both of these components became voluntary at both the student and school levels beginning with the 2015/16 school year. This policy change appears to have been motivated by political pressure to reduce the amount of standardized testing, rather than in response to the implementation or effectiveness of the FCCRI. However, it appears that it may be worthwhile for high schools to find an effective method to assign students to the CRS courses, which still remain certified by FLDOE, even if the PERT is not used in this process.

Policy implications for other states

This evaluation has important implications for researchers and policymakers to consider for similar programs in other states. When the FCCRI first began in 2007, it was a rather innovative policy, but initiatives like this have gained increasing popularity over time. A national scan conducted in 2012 found that 29 states offer transitional math and English courses during the senior year of high school for students who have not previously met college readiness benchmarks (Barnett et al., 2013).

First, the findings from Florida suggest that initiatives like the FCCRI may not be effective for improving outcomes such as high school graduation rates or college enrollment. In our original theory of action, we hypothesized that the FCCRI may provide a signal to help students to see that they are capable of obtaining the skills needed for college-level work, which may encourage them to complete high school and continue to postsecondary education. Yet this does not appear to be the case, as we find that the FCCRI had no effect on these outcomes for students at any level of performance on the FCAT or PERT. These findings from the impact analyses also correspond with feedback that we received from educators that the FCCRI seems to be most effective for students who want to attend college but are not quite college-ready, and least effective for students who are disengaged from school and lack realistic postsecondary goals (Mokher & Jacobson, 2014). It may not be practical to limit CRS courses to students who indicate they are college-bound in grade 11, particularly since high school students' college intent is often uncertain and subject to change. However, the findings do suggest that states looking for ways to improve high school graduation and college enrollment rates may want to consider alternative policies.

Second, there is some evidence the FCCRI may meaningfully reduce the need for developmental education for some students, particularly those who enroll in college and have certain academic performance levels. This suggests that the FCCRI was successful in its goal of reducing the need for remediation among students who enroll in state colleges, and other states may want to consider this type of initiative if they have a similar goal.

Third, our study has important implications about the types of students who should be targeted for participation in this type of policy. We found that the FCCRI had differential impacts based on students' prior academic achievement. The initiative was most beneficial to students who were not so far behind that it was unfeasible to catch up in a single year, but not so advanced that they were already college-ready. States should collect feedback from educators and look at their own data to identify the types of students who are most likely to benefit from this type of initiative, and use this information to inform decisions about student eligibility criteria.

Fourth, our cost-benefit analysis demonstrates that programs such as the FCCRI can be implemented at a low cost (about $1 per student) if substantial resources are repurposed from other uses. Other states should consider whether their personnel have the capacity to add activities to their existing responsibilities and assess the value of the alternative activities they will have to give up to reallocate staff time. There may also be high opportunity costs associated with reallocating nonpersonnel resources as well, even though facilities and materials and equipment represented a relatively small percentage of the FCCRI's dollar cost. An additional consideration is that Florida chose to implement the FCCRI largely as an unfunded mandate. Because of this choice, resource allocation was very uneven across districts and schools. Other states interested in implementing a similar program would not necessarily have to take the same approach to funding and may want to consider alternative approaches to increase the likelihood of program effectiveness, both overall and across districts and schools.

Lastly, our study sheds light on the types of challenges states may face in implementing similar initiatives, as well as suggestions for ways to make statewide college readiness testing and CRS courses more effective. In particular, states should focus on making sure that information is being communicated clearly and educators have the resources they need. In Florida, many teachers were unsure about the type of content included on the new college readiness test, and they also lacked curricular materials for CRS courses, which led them to be uncertain about whether they were covering key topics. High school educators may benefit from working with staff at local colleges, including teachers in their own schools who have experience teaching in adjunct positions at colleges. Although collaboration between K–12 and postsecondary education tends to be limited, those who had participated in collaborative activities found them very beneficial, and those who had not indicated a strong desire for greater collaboration.

Appendix A: Technical Appendix for the Confirmatory Impact Analysis

Descriptive statistics of student characteristics

Tables 1A and 2A use all students with valid FCAT scores in math or reading, respectively, while Tables 3A and 4A use all students with valid grade 11 PERT scores who had been targeted for the FCCRI and all students with valid grade 12 PERT scores who had been targeted for the FCCRI, failed the PERT in grade 11, and taken a CRS course. For each assessment, the first column contains characteristics for the full sample in the given subject. The first column in Tables 1A and 2A covers all FCAT-takers with valid scores regardless of whether they were targeted; the first column for the grade 11 PERT in Tables 1A and 2A contains all targeted students with valid FCAT and PERT scores; and the first column for the grade 12 PERT in Tables 3A and 4A contains all targeted students with valid FCAT scores who took and failed the grade 11 PERT, took a CRS course, and have valid grade 12 PERT scores. Subsequent pairs of columns show characteristics just above and just below various score cutoffs.[47]

[47] Some FCAT-takers may be simultaneously just above one cutoff and just below another.

Table 1A. Summary of descriptive statistics for math, by FCAT performance, cohort M1

Variable	FCAT Overall	+/- 10 FCAT Points from Cutoff For… Level 2 Below	Level 2 Above	Graduation Below	Graduation Above	Level 3 Below	Level 3 Above	Level 4 Below	Level 4 Above	Level 5 Below	Level 5 Above
Female	0.5104	0.5767	0.5885	0.5882	0.5742	0.5668	0.5470	0.5213	0.4993	0.4598	0.4335
FRPL	0.5849	0.8062	0.7995	0.7957	0.7555	0.7463	0.6953	0.6093	0.5467	0.3960	0.3449
Black	0.2052	0.4286	0.3943	0.3843	0.3446	0.3243	0.2693	0.1965	0.1538	0.0784	0.0607
Hispanic	0.2661	0.2959	0.2944	0.2954	0.2913	0.2887	0.2990	0.2821	0.2641	0.2246	0.2081
Asian	0.0300	0.0128	0.0132	0.0146	0.0154	0.0154	0.0171	0.0228	0.0255	0.0481	0.0577
Other Minority	0.0306	0.0218	0.0211	0.0233	0.0274	0.0303	0.0308	0.0299	0.0340	0.0327	0.0366
Separate ELL	0.0534	0.1369	0.1204	0.1166	0.0902	0.0826	0.0580	0.0400	0.0287	0.0137	0.0090
Integrated ELL	0.0041	0.0090	0.0069	0.0066	0.0058	0.0052	0.0053	0.0028	0.0030	0.0015	0.0010
Former ELL	0.1368	0.1391	0.1556	0.1523	0.1582	0.1598	0.1684	0.1514	0.1380	0.1071	0.0983
Mentally Disabled	0.0862	0.1968	0.1524	0.1405	0.1312	0.1169	0.0914	0.0671	0.0527	0.0292	0.0201
Otherwise Disabled	0.0231	0.0428	0.0479	0.0431	0.0315	0.0306	0.0216	0.0180	0.0151	0.0099	0.0075
Gifted	0.0869	0.0016	0.0017	0.0018	0.0040	0.0055	0.0112	0.0277	0.0557	0.2075	0.3038
Cumulative GPA	2.78 (0.67)	2.23 (0.53)	2.32 (0.54)	2.34 (0.55)	2.38 (0.54)	2.41 (0.55)	2.51 (0.55)	2.70 (0.56)	2.84 (0.56)	3.23 (0.53)	3.37 (0.51)
Took Math PERT	0.6416	0.3871	0.8240	0.8290	0.8501	0.8492	0.8491	0.8155	0.7614	0.5522	0.0964
Took Math CRS	0.3866	0.4483	0.5965	0.6129	0.6311	0.6279	0.6245	0.5250	0.3988	0.0942	0.0394
Overall N	145749	3,761	4,783	5,017	9,335	11,038	14,427	18,540	18,842	11,158	7,031
Math PERT Score	103.6 (14.2)	87.8 (10.8)	89.9 (10.7)	90.7 (10.6)	91.9 (10.7)	93.6 (10.6)	97.2 (10.5)	103.2 (10.2)	107.9 (10.2)	120.9 (10.7)	124.8 (11.4)
Passed Math PERT	0.2644	0.0076	0.0129	0.0147	0.0183	0.0305	0.0607	0.1739	0.3249	0.8142	0.9041
PERT-Taker N	93,513	1,456	3,941	4,159	7,936	9,374	12,250	15,120	14,347	6,162	678

Table 2A. Summary of descriptive statistics for math, by PERT performance in grade 11 and retest for targeted CRS participants, cohort M1

Variable	Initial PERT Test, Targeted Students: Overall	+/- 10 Points of Passing: Below	+/- 10 Points of Passing: Above	PERT Retest, Targeted CRS Participants: Overall	+/- 10 Points of Passing: Below	+/- 10 Points of Passing: Above
Female	0.5118	0.5343	0.5364	0.5184	0.5263	0.5276
FRPL	0.6296	0.6142	0.5372	0.6891	0.6733	0.6428
African-American	0.2074	0.1979	0.1643	0.2265	0.2294	0.1948
Hispanic	0.2941	0.2909	0.2789	0.3107	0.2920	0.2895
Asian	0.0240	0.0232	0.0341	0.0163	0.0167	0.0223
Other Minority	0.0301	0.0308	0.0319	0.0289	0.0291	0.0338
Separate ELL	0.0492	0.0466	0.0379	0.0557	0.0523	0.0522
Integrated ELL	0.0037	0.0033	0.0033	0.0039	0.0032	0.0047
Former ELL	0.1611	0.1566	0.1485	0.1738	0.1649	0.1406
Mentally Disabled	0.0757	0.0598	0.0418	0.0955	0.0912	0.0750
Otherwise Disabled	0.0193	0.0150	0.0111	0.0213	0.0195	0.0170
Gifted	0.0478	0.0427	0.0887	0.0179	0.0177	0.0231
Overall N	89,176	24,184	16,196	29,253	11,325	6,535

Table 3A. Summary of descriptive statistics for reading, by FCAT performance, cohort M1

Variable	Overall	+/- 10 FCAT Points from Cutoff For... Level 2 Below	Level 2 Above	Graduation Below	Graduation Above	Level 3 Below	Level 3 Above	Level 4 Below	Level 4 Above	Level 5 Below	Level 5 Above
Female	0.5104	0.5240	0.5033	0.4961	0.5016	0.5000	0.4972	0.4938	0.5144	0.5162	0.5083
FRPL	0.5852	0.6888	0.6602	0.6473	0.6395	0.5736	0.5381	0.4814	0.4534	0.4281	0.3806
Black	0.2056	0.2713	0.2482	0.2359	0.2278	0.1755	0.1587	0.1180	0.1079	0.0962	0.0850
Hispanic	0.2660	0.2842	0.2813	0.2760	0.2776	0.2702	0.2660	0.2527	0.2464	0.2403	0.2227
Asian	0.0300	0.0209	0.0218	0.0202	0.0227	0.0252	0.0253	0.0320	0.0392	0.0375	0.0456
Other Minority	0.0306	0.0293	0.0287	0.0308	0.0330	0.0326	0.0339	0.0319	0.0329	0.0330	0.0328
Separate ELL	0.0536	0.0553	0.0441	0.0417	0.0389	0.0278	0.0222	0.0146	0.0114	0.0097	0.0093
Integrated ELL	0.0041	0.0030	0.0055	0.0055	0.0035	0.0033	0.0032	0.0029	0.0019	0.0018	0.0008
Former ELL	0.1367	0.1566	0.1600	0.1562	0.1572	0.1495	0.1438	0.1321	0.1248	0.1124	0.0987
Mentally Disabled	0.0865	0.1008	0.0844	0.0794	0.0823	0.0585	0.0496	0.0381	0.0311	0.0265	0.0220
Otherwise Disabled	0.0232	0.0269	0.0235	0.0206	0.0228	0.0203	0.0174	0.0135	0.0109	0.0120	0.0114
Gifted	0.0867	0.0113	0.0199	0.0210	0.0205	0.0466	0.0586	0.1059	0.1418	0.1682	0.2103
Cumulative GPA	2.78 (0.67)	2.51 (0.57)	2.58 (0.57)	2.60 (0.57)	2.63 (0.56)	2.77 (0.57)	2.85 (0.57)	2.99 (0.58)	3.08 (0.56)	3.15 (0.56)	3.23 (0.54)
Took Reading PERT	0.4457	0.4300	0.7693	0.7628	0.8116	0.7618	0.7335	0.6518	0.1183	0.1124	0.1010
Took Reading CRS	0.3200	0.4885	0.5221	0.5032	0.5154	0.3819	0.2995	0.1920	0.1121	0.0914	0.0720
Overall N	145,922	7,093	7,888	6,938	11,574	11,935	10,414	8,822	8,258	7,311	5,178
Reading PERT Score	101.5 (16.3)	94.1 (13.5)	96.6 (13.8)	97.4 (14.1)	97.3 (14.2)	102.7 (14.4)	105.9 (14.3)	110.7 (14.5)	114.8 (14.1)	116.6 (15.5)	119.9 (13.3)
Passed Reading PERT	0.4561	0.2456	0.3042	0.3316	0.3231	0.4945	0.5871	0.7200	0.8127	0.8175	0.9120
PERT-Taker N	65,033	3,050	6,068	5,292	9,394	9,092	7,639	5,750	977	822	523

Table 4A. Summary of descriptive statistics for reading, by PERT performance in grade 11 and retest for targeted CRS participants, cohort M1

Variable	Initial PERT Test, Targeted Students: Overall	+/- 10 Points of Passing: Below	+/- 10 Points of Passing: Above	PERT Re-Test, Targeted CRS Participants: Overall	+/- 10 Points of Passing: Below	+/- 10 Points of Passing: Above
Female	0.4883	0.5067	0.5236	0.4169	0.4584	0.3965
FRPL	0.6052	0.6206	0.5862	0.6681	0.6717	0.6429
African-American	0.1853	0.2022	0.1727	0.2079	0.2119	0.1724
Hispanic	0.2945	0.3070	0.2936	0.3135	0.3256	0.2921
Asian	0.0251	0.0251	0.0271	0.0199	0.0229	0.0202
Other Minority	0.0315	0.0294	0.0318	0.0274	0.0273	0.0263
Separate ELL	0.0330	0.0362	0.0254	0.0471	0.0457	0.0370
Integrated ELL	0.0034	0.0034	0.0024	0.0035	0.0026	0.0025
Former ELL	0.1677	0.1812	0.1687	0.1853	0.1981	0.1707
Mentally Disabled	0.0663	0.0666	0.0515	0.1014	0.0999	0.0868
Otherwise Disabled	0.0182	0.0176	0.0156	0.0228	0.0199	0.0234
Gifted	0.0435	0.0346	0.0493	0.0173	0.0141	0.0243
Overall N	51,279	14,138	13,969	10,044	3,412	2,431

Table 5A. Summary of descriptive statistics for math, by FCAT performance, cohort M2

Variable	FCAT Overall	+/- 10 FCAT Points from Cutoff For... Level 2 Below	Level 2 Above	Level 3 Below	Level 3 Above	Level 4 Below	Level 4 Above	Level 5 Below	Level 5 Above
Female	0.4657	0.4561	0.4745	0.5077	0.4836	0.4435	0.4792	0.4715	0.3582
FRPL	0.7570	0.7944	0.7963	0.7720	0.7137	0.6638	0.5708	0.5699	0.5672
African-American	0.3057	0.3499	0.3537	0.2936	0.2503	0.1812	0.1333	0.1451	0.1642
Hispanic	0.2574	0.2595	0.2584	0.2725	0.2295	0.2449	0.2208	0.2124	0.1791
Asian	0.0137	0.0058	0.0146	0.0120	0.0120	0.0130	0.0292	0.0415	0.1194
Other Minority	0.0321	0.0274	0.0282	0.0321	0.0366	0.0377	0.0417	0.0518	0.0299
Separate ELL	0.1009	0.1119	0.1116	0.0805	0.0781	0.0913	0.0667	0.0777	0.1343
Integrated ELL	0.0085	0.0066	0.0081	0.0057	0.0093	0.0130	0.0042	0.0052	0.0000
Former ELL	0.1176	0.1177	0.1186	0.1226	0.1000	0.0986	0.0958	0.1140	0.1493
Mentally Disabled	0.2034	0.2653	0.2232	0.1595	0.1366	0.0696	0.0833	0.0622	0.0299
Otherwise Disabled	0.0560	0.0779	0.0498	0.0517	0.0393	0.0275	0.0500	0.0466	0.0448
Gifted	0.0092	0.0025	0.0054	0.0053	0.0060	0.0217	0.0500	0.0725	0.1045
Cumulative GPA	2.19 (0.62)	1.98 (0.55)	2.07 (0.55)	2.23 (0.55)	2.36 (0.59)	2.63 (0.62)	2.85 (0.60)	2.94 (0.59)	3.21 (0.57)
Took Math CRS	0.4196	0.2919	0.4431	0.4962	0.5060	0.4478	0.3208	0.2694	0.0746
Overall N	11,657	1,206	1,846	2,088	1,830	690	240	193	67
Math PERT Score	95.9 (11.2)	88.3 (10.0)	91.1 (8.7)	95.5 (7.6)	99.5 (7.6)	107.6 (8.5)	115.4 (10.8)	118.2 (10.8)	122.1 (11.3)
Passed Math PERT	0.0563	0.0029	0.0025	0.0110	0.0342	0.2897	0.5743	0.6754	0.7647
PERT-Taker N	6,505	339	1,193	1,454	1,285	435	148	114	17

Table 6A. Summary of descriptive statistics for math, by PERT performance in grade 11 and retest for targeted CRS participants, cohort M2

Variable	Initial PERT Test, Targeted Students: Overall	+/- 10 Points of Passing: Below	+/- 10 Points of Passing: Above	PERT Retest, Targeted CRS Participants: Overall	+/- 10 Points of Passing: Below	+/- 10 Points of Passing: Above
Female	0.5179	0.5299	0.5380	0.5241	0.5308	0.5207
FRPL	0.6467	0.6345	0.5566	0.6979	0.6969	0.6426
African-American	0.2071	0.2011	0.1491	0.2427	0.2555	0.2136
Hispanic	0.2667	0.2615	0.2414	0.2882	0.2931	0.2870
Asian	0.0232	0.0204	0.0326	0.0206	0.0200	0.0300
Other Minority	0.0324	0.0340	0.0364	0.0315	0.0332	0.0307
Separate ELL	0.0637	0.0536	0.0400	0.0807	0.0698	0.0704
Integrated ELL	0.0047	0.0042	0.0026	0.0061	0.0061	0.0074
Former ELL	0.1280	0.1278	0.1140	0.1453	0.1568	0.1467
Mentally Disabled	0.0967	0.0778	0.0484	0.1138	0.0906	0.0695
Otherwise Disabled	0.0238	0.0175	0.0154	0.0233	0.0211	0.0173
Gifted	0.0427	0.0301	0.0672	0.0195	0.0184	0.0394
Overall N	57,671	16,751	11,002	27,857	17,183	10,894

Table 7A. Summary of descriptive statistics for reading, by FCAT performance, cohort M2

Variable	Overall	+/- 10 FCAT Points from Cutoff For... Level 2 Below	Level 2 Above	Level 3 Below	Level 3 Above	Level 4 Below	Level 4 Above	Level 5 Below	Level 5 Above
Female	0.5101	0.5322	0.5413	0.5348	0.5072	0.5060	0.4873	0.4856	0.4901
FRPL	0.6012	0.7954	0.7366	0.6679	0.5696	0.5606	0.4682	0.4247	0.3507
African-American	0.2021	0.3380	0.2800	0.2241	0.1690	0.1648	0.1143	0.0989	0.0672
Hispanic	0.2751	0.3189	0.3019	0.2966	0.2691	0.2657	0.2425	0.2299	0.2107
Asian	0.0303	0.0192	0.0206	0.0231	0.0276	0.0287	0.0360	0.0420	0.0519
Other Minority	0.0303	0.0246	0.0288	0.0293	0.0332	0.0332	0.0349	0.0335	0.0351
Separate ELL	0.0571	0.1171	0.0667	0.0447	0.0237	0.0218	0.0092	0.0068	0.0032
Integrated ELL	0.0045	0.0082	0.0045	0.0036	0.0037	0.0035	0.0021	0.0020	0.0008
Former ELL	0.1396	0.1687	0.1769	0.1714	0.1452	0.1434	0.1163	0.1039	0.0891
Mentally Disabled	0.0896	0.1580	0.0997	0.0736	0.0573	0.0561	0.0393	0.0336	0.0260
Otherwise Disabled	0.0246	0.0386	0.0250	0.0222	0.0183	0.0177	0.0147	0.0143	0.0124
Gifted	0.0854	0.0027	0.0087	0.0199	0.0493	0.0534	0.1281	0.1804	0.2920
Cumulative GPA	2.77 (0.69)	2.33 (0.57)	2.50 (0.58)	2.65 (0.58)	2.82 (0.59)	2.84 (0.59)	3.07 (0.58)	3.17 (0.57)	3.36 (0.54)
Took Reading PERT	0.4632	0.4170	0.7793	0.7305	0.6943	0.6832	0.1668	0.1408	0.1063
Took Reading CRS	0.3217	0.5086	0.5598	0.4621	0.3225	0.3034	0.1330	0.0976	0.0558
Overall N	147,571	13,667	21,836	24,050	31,437	31,077	22,812	17,759	9,107
Reading PERT Score	98.0 (14.2)	86.6 (11.1)	92.5 (10.9)	97.3 (11.1)	102.9 (11.1)	103.6 (11.1)	110.9 (11.6)	114.0 (11.6)	119.9 (11.5)
Passed Reading PERT	0.3490	0.0497	0.1366	0.2776	0.4985	0.5262	0.7763	0.8513	0.9514
PERT-Taker N	68,358	5,699	17,016	17,569	21,827	21,232	3,804	2,501	968

Table 8A. Summary of descriptive statistics for reading, by PERT performance in grade 11 and retest for targeted CRS participants, cohort M2

Variable	Initial PERT Test, Targeted Students: Overall	+/- 10 Points of Passing: Below	+/- 10 Points of Passing: Above	PERT Re-Test, Targeted CRS Participants: Overall	+/- 10 Points of Passing: Below	+/- 10 Points of Passing: Above
Female	0.5118	0.5429	0.5219	0.4610	0.4951	0.4688
FRPL	0.6583	0.6798	0.6025	0.7110	0.7179	0.6707
African-American	0.2138	0.2272	0.1727	0.2191	0.2224	0.1759
Hispanic	0.3062	0.3241	0.2879	0.3290	0.3322	0.3248
Asian	0.0234	0.0258	0.0254	0.0197	0.0205	0.0192
Other Minority	0.0295	0.0297	0.0310	0.0286	0.0288	0.0342
Separate ELL	0.0447	0.0476	0.0232	0.0549	0.0552	0.0470
Integrated ELL	0.0036	0.0038	0.0034	0.0038	0.0028	0.0049
Former ELL	0.1807	0.2020	0.1665	0.1913	0.1936	0.1906
Mentally Disabled	0.0758	0.0711	0.0545	0.0997	0.0937	0.0695
Otherwise Disabled	0.0212	0.0192	0.0173	0.0232	0.0205	0.0226
Gifted	0.0281	0.0236	0.0408	0.0103	0.0108	0.0139
Overall N	52,625	18,470	13,496	12,500	4,726	2,660

Outcome variables for confirmatory and exploratory impact analyses

Table 9A defines all the outcomes variables that are used in both the confirmatory and the exploratory impact analyses.

Table 9A. Outcomes and descriptions for the confirmatory and exploratory impact analyses

Short-Term Outcomes	Chapters	Description
High school diploma or equivalent	4, 5	Binary variable equal to one if the student received a high school diploma or equivalent.
Seamless college enrollment	4, 5	Binary variable equal to one if the student enrolled in college the fall semester following their cohort's high school graduation.
First math/English course level	5	Categorical outcome that captures the level of the first course a student takes in college: no course, lower level DE, upper level DE, transitional (math only), or degree credit. If no course was taken during the first year, dual enrollment and advanced placement courses are considered. Outcomes were created separately for math and English.
Pass first math/English course	5	Binary variable equal to one if the student passed their first college course. If no course was taken during the first year, dual enrollment and advanced placement courses are considered. Outcomes created separately for math and English.
Nondevelopmental enroll	4, 5	Binary variable equal to one if the student enrolled in a nondevelopmental math/English course first. Outcomes created separately for math and English.
Nondevelopmental pass	4, 5	Binary variable equal to one if the student enrolled in a nondevelopmental math/English course first and passed. Outcomes created separately for math and English.

Long-Term Outcomes	Chapters	Description
Persist	4	Binary variable equal to one if a seamless enrollee was still enrolled in a postsecondary institution during the second or third fall semester.
Transfer two-to-four	4	Binary variable equal to one if a seamless enrollee initially enrolled only in a two-year school later enrolled in a four-year school
Non-DE enrollment	4	Defined as above
Non-DE pass	4	Defined as above
Any degree	4	Binary variable equal to one if a seamless enrollee completed any degree or credential within two years or three year
Enrolled two-year	4	Binary variable equal to one if a student enrolled in a two-year school within two years or three years
Enrolled four-year	4	Binary variable equal to one if a student enrolled in a four-year school within two years or three years

McCrary density test results examining the integrity of the running variable

Table 10A provides the results of McCrary density tests, which examine the integrity of the running variable, for cohort M1. The outcome variable is the number of students at each possible FCAT score. Discontinuities are computed using quartic functions on either side of the cutoff and using all points between the next lowest and next highest cutoff (using a minimum of 250 and a maximum of 425 when computing discontinuities at PL2 and PL5 respectively). We omit FCAT scores attained by fewer than 25 students, which reflect mechanical scoring issues rather than conscious manipulation of the running variable (36 non-attriting students with scores between 250 and 425 fit this description). For the grade 11 and grade 12 PERT assessments, bandwidths were selected using the range to the left and to the right of the cutoff that produced the highest adjusted R^2 under quartic regression. Standard errors are listed in parentheses. Figure 1.A shows these results graphically.

Table 10A. McCrary density test results, cohort M1

FCAT Score	Math	Reading (Modified)
Level 2	1.2 (14.1)	-38.8 (42.6)
Graduation	317.1*** (59.7)	430.0*** (38.7)
Level 3	25.3 (69.1)	72.9 (63.0)
Level 4	98.1* (57.2)	-60.9* (35.4)
Level 5	-39.4 (37.0)	-37.1 (68.2)

Grade 11 PERT, Targeted Students Only	Math	Reading
Pass	-231.9 (195.4)	345.9** (140.6)

PERT Retest, Targeted CRS-Takers Only	Math	Reading
Pass	-52.9 (128.7)	5.2 (154.5)

* = 10% significance Level, ** = 5% significance Level, *** = 1% significance Level.

In both Table 10A and Figure 1A, there are large discontinuities at the FCAT graduation cutoffs in both subjects. These discontinuities are due to policies that specifically encourage bunching above the graduation cutoff. RD results at this cutoff would be invalid, but it is far enough away from the cutoffs we are analyzing that it should not affect our estimates. At the grade 11 PERT cutoff in reading, there is a statistically significant estimate in Table 10A, but the figure does not show a large discontinuity at the FCAT graduation cutoffs. These results might be due to an abnormally low density at the score just below the cutoff and an abnormally high one at the cutoff—statistical noise could be overfit using a recentered quartic regression. Nevertheless, results using the grade 11 PERT's reading cutoff should be taken with caution.

Figure 1A. McCrary Density Tests, cohort M1

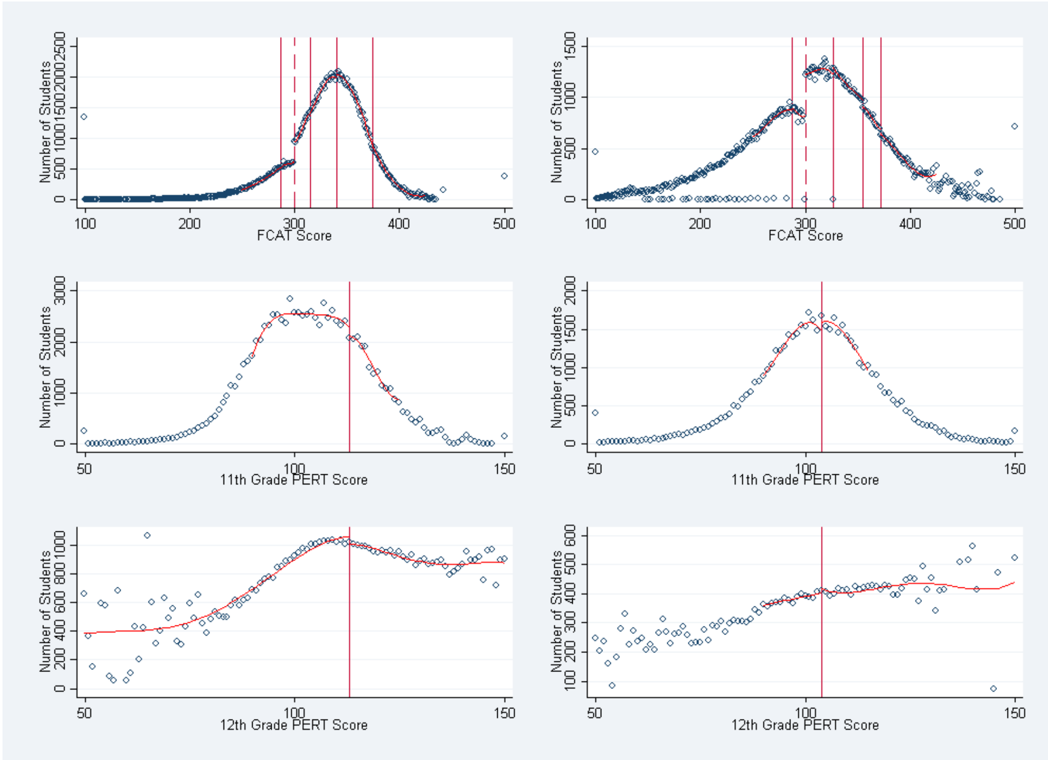


Table 11A and Figure 2A contain similar results for cohort M2. It is unclear why there is a negative discontinuity at the graduation cutoff for the Algebra 1 EOC. However, since relatively few students take the Algebra 1 EOC, it may be susceptible to statistical noise. Overfitting may explain why we find a statistically significant result at the PL5 cutoff for the EOC, as this result is not visible in Figure 2A (which does not recenter scores). However, since there are so few students at this cutoff, we do not use it in our analyses. The only other statistically significant effect is at the graduation cutoff in reading.

Table 11A. McCrary density test results, cohort M2

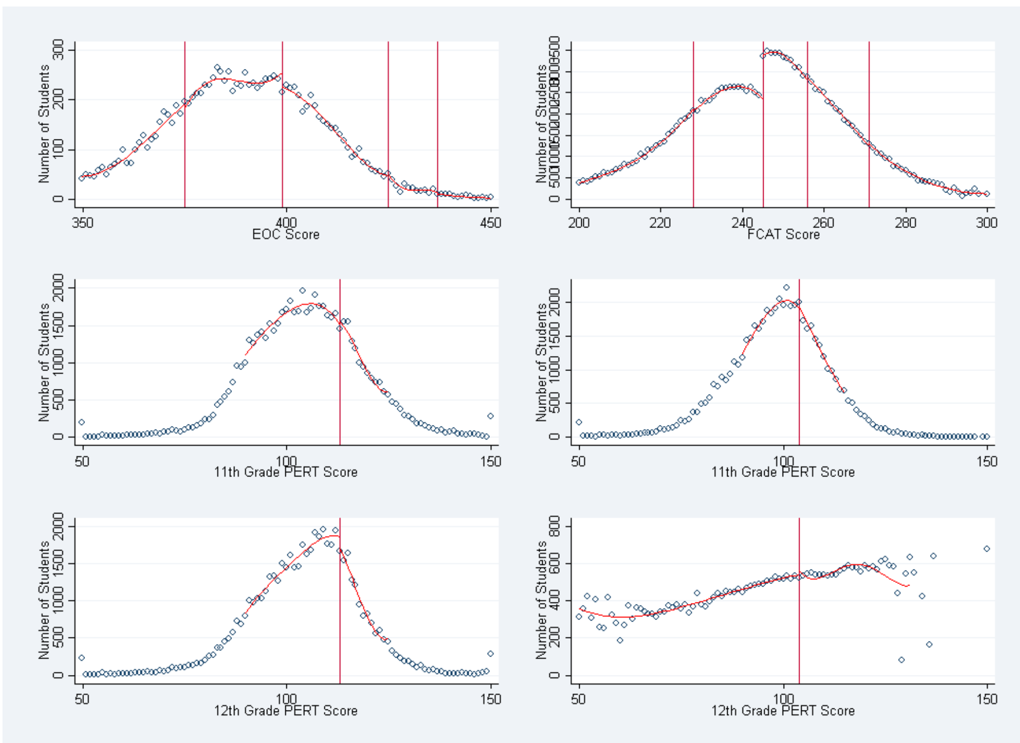
FCAT/EOC Score	Math	Reading
Level 2	15.5 (10.1)	-18.9 (61.3)
Level 3/ Graduation	-41.6** (17.3)	996.6*** (93.9)
Level 4	12.5 (15.1)	136.0 (119.8)
Level 5	-24.4*** (7.8)	57.6 (95.4)

Grade 11 PERT, Targeted Students Only	Math	Reading
Pass	-39.8 (140.7)	254.5 (180.2)

PERT Retest, Targeted CRS-Takers Only	Math	Reading
Pass	-205.2 (154.8)	20.8 (69.3)

* = 10% significance Level, ** = 5% significance Level, *** = 1% significance Level.

Figure 2A. McCrary Density Tests, cohort M2



Attrition at the RD cutoffs

WWC guidelines state that attrition must be both computed within a bandwidth of RD cutoffs and directly estimated at the cutoffs themselves. Results of estimation at RD cutoffs are presented in Table 12A. Local linear estimation was used to separately compute the amount of attrition at each cutoff from below and from above. Average attrition and differential attrition are the average value and the difference of these estimates, respectively. These estimates are similar in magnitude to those computed in the body of the report.

Table 12A: Attrition within interval, cohort M1

Cohort M1	Math		Reading	
	Average	Differential	Average	Differential
FCAT Low Margin	8.2%	1.0%	6.0%	2.5%
FCAT Upper Margin	2.4%	0.3%	3.4%	0.4%
Grade 11 PERT	1.8%	0.2%	2.1%	0.6%
Grade 12 PERT	0.4%	0.1%	0.4%	0.1%

Cohort M2	Math		Reading	
	Average	Differential	Average	Differential
FCAT/EOC Low Margin	8.1%	1.8%	7.1%	1.9%
FCAT/EOC Upper Margin	9.8%	0.5%	2.8%	0.6%
Grade 11 PERT	2.6%	0.3%	2.2%	0.5%
Grade 12 PERT	1.4%	0.4%	0.5%	0.3%

Baseline equivalence

Tables 13A through 20A present results from examining baseline equivalence for all groups in the analytic sample. Tables 13A and 14A present FCAT results for cohort M1 in math and reading respectively, Tables 15A and 16A present PERT results for cohort M1 in math and reading respectively, Tables 17A and 18A present FCAT results for cohort M2 in math and reading respectively, and Tables 19A and 20A present PERT estimates for cohort M2 in math and reading respectively. The outcome variable in each cell is the percentage of students fitting the covariate indicated for the row. Point estimates are listed in percentage points rather than as percentages or elasticities. Standard errors are in parentheses. * = significant at the 10% level, ** = significant at the 5% level, *** = significant at the 1% level. Bandwidths (in FCAT or PERT points) are in brackets.

Of the 184 estimates for cohort M1, 24 are effect sizes greater than 0.05 standard deviations. Of these, only 18 are also statistically significant, and 13 of these are when restricting the sample to seamless college enrollees. Cohort M2 has 11 estimates with effect sizes greater than 0.25 standard deviations; however, these either are at the upper Algebra 1 EOC margin (which we do not include in our estimates) or for seamless college enrollees at the lower EOC margin (which has too little density to support estimates for college outcomes). As a result, the 30 estimates for these cutoffs and individuals are not relevant to our analyses. Of the remaining 154 estimates, 18 have effect sizes greater than 0.05 standard deviations, and 11 are statistically significant.

Of the statistically significant estimates, only one is in math, and five are when restricting the sample to seamless college enrollees.

Table 13A. Continuity of covariates across FCAT math cutoffs, cohort M1.

Covariate	All Students		Seamless College Enrollees	
	Lower Margin	Upper Margin	Lower Margin	Upper Margin
Female	0.0111 (0.0132) Effect = 0.0223	-0.0141 (0.0104) Effect = 0.0282	-0.0081 (0.0191) Effect = 0.0163	-0.0109 (0.0139) Effect = 0.0219
Asian	0.0019 (0.0025) Effect = 0.0116	-0.0003 (0.0057) Effect = 0.0019	0.0080 (0.0068) Effect = 0.0422	-0.0003 (0.0072) Effect = 0.0018
Black	-0.0068 (0.0121) Effect = 0.0168	-0.0132** (0.0050) Effect = 0.0324	0.0064 (0.0293) Effect = 0.0165	-0.0213*** (0.0070) Effect = 0.0553
Hispanic	-0.0106 (0.0069) Effect = 0.0240	-0.0046 (0.0078) Effect = 0.0104	-0.0123 (0.0365) Effect = 0.0280	-0.0039 (0.0080) Effect = 0.0090
Other Minority	-0.0002 (0.0038) Effect = 0.0011	0.0013 (0.0042) Effect = 0.0075	0.0083 (0.0083) Effect = 0.0478	-0.0027 (0.0063) Effect = 0.0158
ELL	0.0018 (0.0053) Effect = 0.0045	-0.0023 (0.0035) Effect = 0.0058	0.0041 (0.0155) Effect = 0.0108	-0.0042 (0.0058) Effect = 0.0111
Special Education	0.0015 (0.0016) Effect = 0.0045	0.0004 (0.0005) Effect = 0.0011	0.0017 (0.0014) Effect = 0.0064	0.0001 (0.0003) Effect = 0.0005
FRPL	0.0140 (0.0093) Effect = 0.0285	0.0091 (0.0090) Effect = 0.0186	0.0272 (0.0239) Effect = 0.0543	0.0132 (0.0133) Effect = 0.0264
Cumulative GPA	0.0325 (0.0195) Effect = 0.0483	-0.0065 (0.0127) Effect = 0.0096	0.0570* (0.0324) Effect = 0.0956	0.0030 (0.0106) Effect = 0.0051
FCAT Reading	-0.7933 (1.7119) Effect = 0.0125	-0.7921 (0.9504) Effect = 0.0125	-4.2206 (2.6829) Effect = 0.0787	-0.5469 (0.9629) Effect = 0.0102

Table 14A. Continuity of covariates across FCAT reading cutoffs, cohort M1.

Covariate	All Students		Seamless College Enrollees	
	Lower Margin	Upper Margin	Lower Margin	Upper Margin
Female	-0.0049 (0.0099) Effect = 0.0099	0.0182* (0.0104) Effect = 0.0365	-0.0008 (0.0132) Effect = 0.0016	0.0145 (0.0135) Effect = 0.0293
Asian	0.0022 (0.0026) Effect = 0.0130	-0.0029 (0.0043) Effect = 0.0172	-0.0006 (0.0056) Effect = 0.0033	-0.0012 (0.0060) Effect = 0.0064
Black	0.0004 (0.0058) Effect = 0.0010	0.0062 (0.0071) Effect = 0.0151	0.0104 (0.0119) Effect = 0.0271	0.0046 (0.0068) Effect = 0.0120

Covariate	All Students		Seamless College Enrollees	
	Lower Margin	Upper Margin	Lower Margin	Upper Margin
Hispanic	-0.0043 (0.0066) Effect = 0.0098	-0.0017 (0.0064) Effect = 0.0038	-0.0261** (0.0101) Effect = 0.0595	-0.0094 (0.0074) Effect = 0.0215
Other Minority	-0.0042 (0.0039) Effect = 0.0242	0.0009 (0.0030) Effect = 0.0054	0.0020 (0.0061) Effect = 0.0116	-0.0031 (0.0044) Effect = 0.0177
ELL	0.0045 (0.0045) Effect = 0.0112	-0.0079* (0.0046) Effect = 0.0196	0.0025 (0.0088) Effect = 0.0065	-0.0072 (0.0069) Effect = 0.0190
Special Education	-0.0005 (0.0010) Effect = 0.0016	0.0002 (0.0004) Effect = 0.0007	-0.0011 (0.0007) Effect = 0.0043	0.0002 (0.0006) Effect = 0.0006
FRPL	0.0079 (0.0108) Effect = 0.0161	0.0023 (0.0091) Effect = 0.0047	0.0147 (0.0136) Effect = 0.0294	0.0080 (0.0111) Effect = 0.0160
Cumulative GPA	0.0223 (0.0148) Effect = 0.0331	0.0060 (0.0107) Effect = 0.0090	0.0255 (0.0154) Effect = 0.0426	-0.0062 (0.0132) Effect = 0.0104
FCAT Math	1.0245** (0.4718) Effect = 0.0239	-0.8117* (0.4356) Effect = 0.0189	0.7536 (0.7082) Effect = 0.0231	-0.2910 (0.4237) Effect = 0.0089

Table 15A. Continuity of covariates across PERT math cutoffs, cohort M1.

Covariate	All Students		Seamless College Enrollees	
	Grade 11	Grade 12	Grade 11	Grade 12
Female	-0.0080 (0.0098) Effect = 0.0159	0.0096 (0.0097) Effect = 0.0191	-0.0061 (0.0142) Effect = 0.0123	0.0428*** (0.0160) Effect = 0.0867
Asian	-0.0019 (0.0019) Effect = 0.0119	-0.0047 (0.0031) Effect = 0.0370	0.0017 (0.0031) Effect = 0.0099	-0.0070* (0.0041) Effect = 0.0508
Black	0.0059 (0.0056) Effect = 0.0148	-0.0253*** (0.0079) Effect = 0.0604	0.0092 (0.0061) Effect = 0.0233	-0.0318*** (0.0100) Effect = 0.0741
Hispanic	-0.0078 (0.0049) Effect = 0.0175	0.0218** (0.0102) Effect = 0.0475	-0.0092 (0.0060) Effect = 0.0207	0.0251* (0.0127) Effect = 0.0540
Other Minority	-0.0026 (0.0031) Effect = 0.0149	0.0051 (0.0051) Effect = 0.0302	-0.0011 (0.0036) Effect = 0.0064	0.0040 (0.0070) Effect = 0.0239
ELL	0.0034 (0.0038) Effect = 0.0088	-0.0034 (0.0065) Effect = 0.0080	0.0039 (0.0050) Effect = 0.0100	-0.0004 (0.0090) Effect = 0.0010
Special Education	0.0002 (0.0004) Effect = 0.0007	-0.0004 (0.0005) Effect = 0.0013	0.0005 (0.0005) Effect = 0.0020	-0.0001 (0.0009) Effect = 0.0003

Covariate	All Students		Seamless College Enrollees	
	Grade 11	Grade 12	Grade 11	Grade 12
FRPL	-0.0079 (0.0070) Effect = 0.0161	-0.0024 (0.0073) Effect = 0.0052	-0.0139* (0.0081) Effect = 0.0279	-0.0073 (0.0086) Effect = 0.0152
Cumulative GPA	-0.0167* (0.0091) Effect = 0.0271	-0.0176 (0.0110) Effect = 0.0340	-0.0174* (0.0089) Effect = 0.0310	-0.0229* (0.0115) Effect = 0.0455
FCAT Math	0.5107* (0.2891) Effect = 0.0236	0.9017** (0.4438) Effect = 0.0514	0.7091* (0.3669) Effect = 0.0345	1.2169*** (0.4506) Effect = 0.0710
FCAT Reading	1.4301** (0.6082) Effect = 0.0284	2.5264*** (0.7665) Effect = 0.0587	1.7888** (0.6819) Effect = 0.0388	1.7835* (0.9469) Effect = 0.0439
Grade 11 PERT – Math	-	0.0428 (0.2780) Effect = 0.0044	-	0.1372 (0.3993) Effect = 0.0151
Grade 11 PERT – Reading	-0.1929 (0.3004) Effect = 0.0121	0.4668 (0.4168) Effect = 0.0307	-0.0794 (0.2551) Effect = 0.0053	0.2231 (0.5899) Effect = 0.0153
Grade 12 PERT - Reading	-	-0.0579 (0.5171) Effect = 0.0041	-	-0.1870 (0.7057) Effect = 0.0147

Table 16A. Continuity of covariates across PERT reading cutoffs, cohort M1.

Covariate	All Students		Seamless College Enrollees	
	Grade 11	Grade 12	Grade 11	Grade 12
Female	-0.0057 (0.0083) Effect = 0.0114	-0.0136 (0.0172) Effect = 0.0276	-0.0133 (0.0102) Effect = 0.0267	-0.0109 (0.0321) Effect = 0.0218
Asian	0.0002 (0.0027) Effect = 0.0011	-0.0037 (0.0047) Effect = 0.0267	0.0010 (0.0042) Effect = 0.0059	-0.0192* (0.0108) Effect = 0.1274
Black	-0.0190*** (0.0060) Effect = 0.0496	0.0028 (0.0151) Effect = 0.0069	-0.0188*** (0.0067) Effect = 0.0488	0.0392* (0.0219) Effect = 0.0937
Hispanic	-0.0081 (0.0050) Effect = 0.0183	-0.0012 (0.0130) Effect = 0.0027	-0.0095* (0.0054) Effect = 0.0212	0.0009 (0.0208) Effect = 0.0018
Other Minority	0.0021 (0.0030) Effect = 0.0120	-0.0073 (0.0080) Effect = 0.0445	0.0050 (0.0035) Effect = 0.0283	-0.0128 (0.0116) Effect = 0.0793
ELL	-0.0065* (0.0036) Effect = 0.0171	0.0006 (0.0072) Effect = 0.0014	-0.0119** (0.0052) Effect = 0.0311	0.0052 (0.0113) Effect = 0.0119
Special Education	0.0000 (0.0004) Effect = 0.0000	-0.0001 (0.0010) Effect = 0.0004	0.0000 (0.0006) Effect = 0.0001	0.0018 (0.0016) Effect = 0.0057

Covariate	All Students		Seamless College Enrollees	
	Grade 11	Grade 12	Grade 11	Grade 12
FRPL	0.0055 (0.0051) Effect = 0.0112	0.0551*** (0.0175) Effect = 0.1163	0.0124 (0.0099) Effect = 0.0247	0.0450** (0.0213) Effect = 0.0929
Cumulative GPA	-0.0006 (0.0094) Effect = 0.0010	-0.0140 (0.0160) Effect = 0.0268	-0.0195 (0.0149) Effect = 0.0363	-0.0219 (0.0303) Effect = 0.0430
FCAT Math	-0.7244** (0.3286) Effect = 0.0293	-0.3730 (0.7690) Effect = 0.0162	-0.7328* (0.3896) Effect = 0.0319	-1.6818 (1.0420) Effect = 0.0749
FCAT Reading	0.0954 (0.3174) Effect = 0.0052	0.1499 (0.9030) Effect = 0.0092	0.1241 (0.4425) Effect = 0.0068	0.0118 (1.1778) Effect = 0.0007
Grade 11 PERT – Math	-0.3867* (0.2069) Effect = 0.0295	1.3244*** (0.4498) Effect = 0.1102	-0.7556** (0.3123) Effect = 0.0605	1.2896* (0.6460) Effect = 0.1086
Grade 11 PERT – Reading	-	-0.0783 (0.3489) Effect = 0.0073	-	0.4670 (0.4146) Effect = 0.0453
Gade 12 PERT – Math	-	0.6037 (0.4271) Effect = 0.0532	-	0.7868 (0.7028) Effect = 0.0700

Table 17A. Continuity of covariates across Algebra 1 EOC cutoffs, cohort M2.

Covariate	All Students		Seamless College Enrollees	
	Lower Margin	Upper Margin	Lower Margin	Upper Margin
Female	-0.0059 (0.0264) Effect = 0.0118	-0.2175*** (0.0814) Effect = 0.4350	-0.0023 (0.0454) Effect = 0.0046	-0.4537*** (0.1360) Effect = 0.9144
Asian	0.0116*** (0.0043) Effect = 0.0689	0.0799 (0.0482) Effect = 0.4741	0.0087 (0.0122) Effect = 0.0457	0.1554** (0.0663) Effect = 0.8150
Black	0.0496 (0.0357) Effect = 0.1225	-0.0291 (0.0793) Effect = 0.0717	0.0883* (0.0493) Effect = 0.2313	-0.0489 (0.0566) Effect = 0.1282
Hispanic	-0.0353 (0.0291) Effect = 0.0789	-0.0905 (0.0806) Effect = 0.2021	-0.0189 (0.0328) Effect = 0.0426	-0.1286 (0.0925) Effect = 0.2899
Other Minority	0.0008 (0.0095) Effect = 0.0048	-0.0630 (0.0399) Effect = 0.3688	-0.0437** (0.0181) Effect = 0.2547	-0.1057 (0.0741) Effect = 0.6164
ELL	-0.0061 (0.0109) Effect = 0.0150	0.0453 (0.0446) Effect = 0.1122	-0.0046 (0.0326) Effect = 0.0118	0.0605 (0.0730) Effect = 0.1564
Special Education	-0.0131* (0.0066) Effect = 0.0390	0.0019 (0.0022) Effect = 0.0055	-0.0084 (0.0064) Effect = 0.0314	0.0000 (0.0000) Effect = 0.0000

Covariate	All Students		Seamless College Enrollees	
	Lower Margin	Upper Margin	Lower Margin	Upper Margin
FRPL	-0.0212 (0.0234) Effect = 0.0435	0.0501 (0.0797) Effect = 0.1031	-0.0468 (0.0532) Effect = 0.0936	0.1042 (0.1169) Effect = 0.2086
Cumulative GPA	-0.0191 (0.0308) Effect = 0.0276	0.2031 (0.1780) Effect = 0.2944	-0.0756 (0.0656) Effect = 0.1241	0.3417* (0.1912) Effect = 0.5609
FCAT Reading	-0.4709 (1.0205) Effect = 0.0212	4.6940 (3.3414) Effect = 0.2115	0.1024 (1.7469) Effect = 0.0054	5.0329 (4.6322) Effect = 0.2633

Table 18A. Continuity of covariates across FCAT reading cutoffs, cohort M2.

Covariate	All Students		Seamless College Enrollees	
	Lower Margin	Upper Margin	Lower Margin	Upper Margin
Female	0.0010 (0.0077) Effect = 0.0021	-0.0126* (0.0070) Effect = 0.0252	0.0215** (0.0094) Effect = 0.0434	-0.0189*** (0.0071) Effect = 0.0380
Asian	-0.0007 (0.0019) Effect = 0.0043	-0.0013 (0.0022) Effect = 0.0076	0.0019 (0.0045) Effect = 0.0098	-0.0005 (0.0033) Effect = 0.0027
Black	0.0077 (0.0077) Effect = 0.0191	-0.0128** (0.0048) Effect = 0.0315	-0.0027 (0.0128) Effect = 0.0072	-0.0139** (0.0069) Effect = 0.0364
Hispanic	-0.0022 (0.0057) Effect = 0.0050	0.0042 (0.0045) Effect = 0.0093	-0.0037 (0.0086) Effect = 0.0084	0.0044 (0.0081) Effect = 0.0100
Other Minority	0.0012 (0.0029) Effect = 0.0068	-0.0007 (0.0022) Effect = 0.0040	-0.0017 (0.0041) Effect = 0.0100	-0.0012 (0.0021) Effect = 0.0068
ELL	0.0109** (0.0043) Effect = 0.0269	-0.0022 (0.0030) Effect = 0.0055	0.0169** (0.0080) Effect = 0.0438	0.0032 (0.0037) Effect = 0.0083
Special Education	0.0009 (0.0010) Effect = 0.0028	0.0000 (0.0006) Effect = 0.0001	0.0014 (0.0014) Effect = 0.0053	-0.0011 (0.0007) Effect = 0.0040
FRPL	0.0138* (0.0069) Effect = 0.0284	0.0060 (0.0051) Effect = 0.0123	0.0259** (0.0115) Effect = 0.0518	0.0148** (0.0068) Effect = 0.0295
Cumulative GPA	-0.0049 (0.0102) Effect = 0.0070	0.0278*** (0.0062) Effect = 0.0402	-0.0272* (0.0154) Effect = 0.0446	0.0269*** (0.0070) Effect = 0.0442
Algebra 1 EOC	1.6175* (0.9067) Effect = 0.0597	2.3558 (1.6983) Effect = 0.0870	0.1174 (1.7996) Effect = 0.0047	-1.8412 (2.5707) Effect = 0.0741

Table 19A. Continuity of covariates across PERT math cutoffs, cohort M2.

Covariate	All Students		Seamless College Enrollees	
	Grade 11	Grade 12	Grade 11	Grade 12
Female	0.0162** (0.0079) Effect = 0.0323	-0.0067 (0.0091) Effect = 0.0134	0.0147 (0.0121) Effect = 0.0296	0.0041 (0.0116) Effect = 0.0083
Asian	-0.0019 (0.0029) Effect = 0.0112	-0.0039 (0.0024) Effect = 0.0232	0.0007 (0.0045) Effect = 0.0038	-0.0011 (0.0023) Effect = 0.0058
Black	-0.0112* (0.0063) Effect = 0.0276	-0.0027 (0.0098) Effect = 0.0066	-0.0150* (0.0078) Effect = 0.0393	-0.0145 (0.0122) Effect = 0.0381
Hispanic	0.0021 (0.0054) Effect = 0.0048	0.0162** (0.0071) Effect = 0.0361	-0.0020 (0.0088) Effect = 0.0045	0.0202* (0.0106) Effect = 0.0456
Other Minority	0.0031 (0.0031) Effect = 0.0179	-0.0019 (0.0040) Effect = 0.0110	-0.0018 (0.0046) Effect = 0.0106	-0.0019 (0.0040) Effect = 0.0110
ELL	-0.0055 (0.0048) Effect = 0.0135	0.0091** (0.0042) Effect = 0.0234	0.0003 (0.0059) Effect = 0.0008	0.0091** (0.0042) Effect = 0.0234
Special Education	-0.0004 (0.0005) Effect = 0.0011	-0.0006 (0.0009) Effect = 0.0021	0.0007 (0.0007) Effect = 0.0026	-0.0006 (0.0009) Effect = 0.0021
FRPL	-0.0001 (0.0059) Effect = 0.0002	-0.0027 (0.0091) Effect = 0.0055	0.0073 (0.0092) Effect = 0.0146	-0.0027 (0.0091) Effect = 0.0055
Cumulative GPA	-0.0160* (0.0086) Effect = 0.0232	-0.0152 (0.0127) Effect = 0.0250	-0.0111 (0.0119) Effect = 0.0182	-0.0152 (0.0127) Effect = 0.0250
Algebra 1 EOC	0.9385 (1.9985) Effect = 0.0378	-0.5911 (1.1957) Effect = 0.0238	0.9385 (1.9985) Effect = 0.0378	-0.5911 (1.1957) Effect = 0.0238
Reading FCAT	0.3808 (0.2484) Effect = 0.0172	0.0204 (0.2627) Effect = 0.0011	0.2050 (0.3287) Effect = 0.0107	0.0204 (0.2627) Effect = 0.0011
Grade 11 PERT – Math	-	0.1370 (0.2653) Effect = 0.0104	-	0.1370 (0.2653) Effect = 0.0104
Grade 11 PERT – Reading	-0.0668 (0.2479) Effect = 0.0046	0.4090 (0.3712) Effect = 0.0308	-0.4636* (0.2728) Effect = 0.0349	0.4090 (0.3712) Effect = 0.0308
Gade 12 PERT - Reading	-	0.2746 (0.5964) Effect = 0.0197	-	0.2746 (0.5964) Effect = 0.0197

Table 20A. Continuity of covariates across PERT reading cutoffs, cohort M2.

Covariate	All Students		Seamless College Enrollees	
	Grade 11	Grade 12	Grade 11	Grade 12
Female	-0.0268*** (0.0065) Effect = 0.0536	-0.0224 (0.0189) Effect = 0.0450	-0.0229*** (0.0068) Effect = 0.0466	-0.0335 (0.0223) Effect = 0.0670
Asian	-0.0005 (0.0032) Effect = 0.0034	0.0060 (0.0048) Effect = 0.0431	-0.0022 (0.0046) Effect = 0.0128	0.0047 (0.0094) Effect = 0.0302
Black	-0.0206*** (0.0046) Effect = 0.0508	0.0019 (0.0112) Effect = 0.0045	-0.0265*** (0.0079) Effect = 0.0652	0.0164 (0.0152) Effect = 0.0386
Hispanic	0.0001 (0.0058) Effect = 0.0002	0.0003 (0.0082) Effect = 0.0006	0.0034 (0.0082) Effect = 0.0075	0.0042 (0.0171) Effect = 0.0088
Other Minority	-0.0027 (0.0022) Effect = 0.0157	0.0025 (0.0055) Effect = 0.0148	0.0024 (0.0037) Effect = 0.0145	-0.0004 (0.0078) Effect = 0.0023
ELL	-0.0048 (0.0046) Effect = 0.0121	0.0126 (0.0084) Effect = 0.0291	-0.0047 (0.0038) Effect = 0.0116	0.0153 (0.0125) Effect = 0.0341
Special Education	0.0006 (0.0003) Effect = 0.0019	-0.0020 (0.0015) Effect = 0.0062	-0.0001 (0.0006) Effect = 0.0004	-0.0002 (0.0009) Effect = 0.0005
FRPL	-0.0068 (0.0069) Effect = 0.0141	-0.0255* (0.0132) Effect = 0.0564	-0.0058 (0.0101) Effect = 0.0117	-0.0581*** (0.0206) Effect = 0.1244
Cumulative GPA	-0.0035 (0.0066) Effect = 0.0058	-0.0172 (0.0251) Effect = 0.0321	0.0080 (0.0112) Effect = 0.0144	-0.0206 (0.0341) Effect = 0.0391
Algebra 1 EOC	-3.0112** (1.5052) Effect = 0.1316	-2.1173 (2.3756) Effect = 0.0952	-4.9473* (2.7039) Effect = 0.2339	-0.5586 (3.6952) Effect = 0.0258
Reading FCAT	0.2039* (0.1074) Effect = 0.0257	0.2188 (0.2106) Effect = 0.0294	0.0886 (0.1784) Effect = 0.0115	-0.1078 (0.2979) Effect = 0.0145
Grade 11 PERT – Math	0.1514 (0.2628) Effect = 0.0127	-0.0925 (0.3826) Effect = 0.0086	0.2479 (0.3673) Effect = 0.0215	0.9252** (0.4290) Effect = 0.0859
Grade 11 PERT – Reading	-	-0.3749 (0.3338) Effect = 0.0410	-	0.4296 (0.5076) Effect = 0.0496
Gade 12 PERT - Math	-	-0.1289 (0.4700) Effect = 0.0113	-	0.5961 (0.5628) Effect = 0.0541

Continuity of the Relationship Between the Outcome and the Forcing Variable

This appendix section provides a graphical analysis displaying the relationship between each of the outcomes used in the analysis and the forcing variables.

Figure 3A. Relationship between year 1 outcomes and forcing variables in math, cohort M1

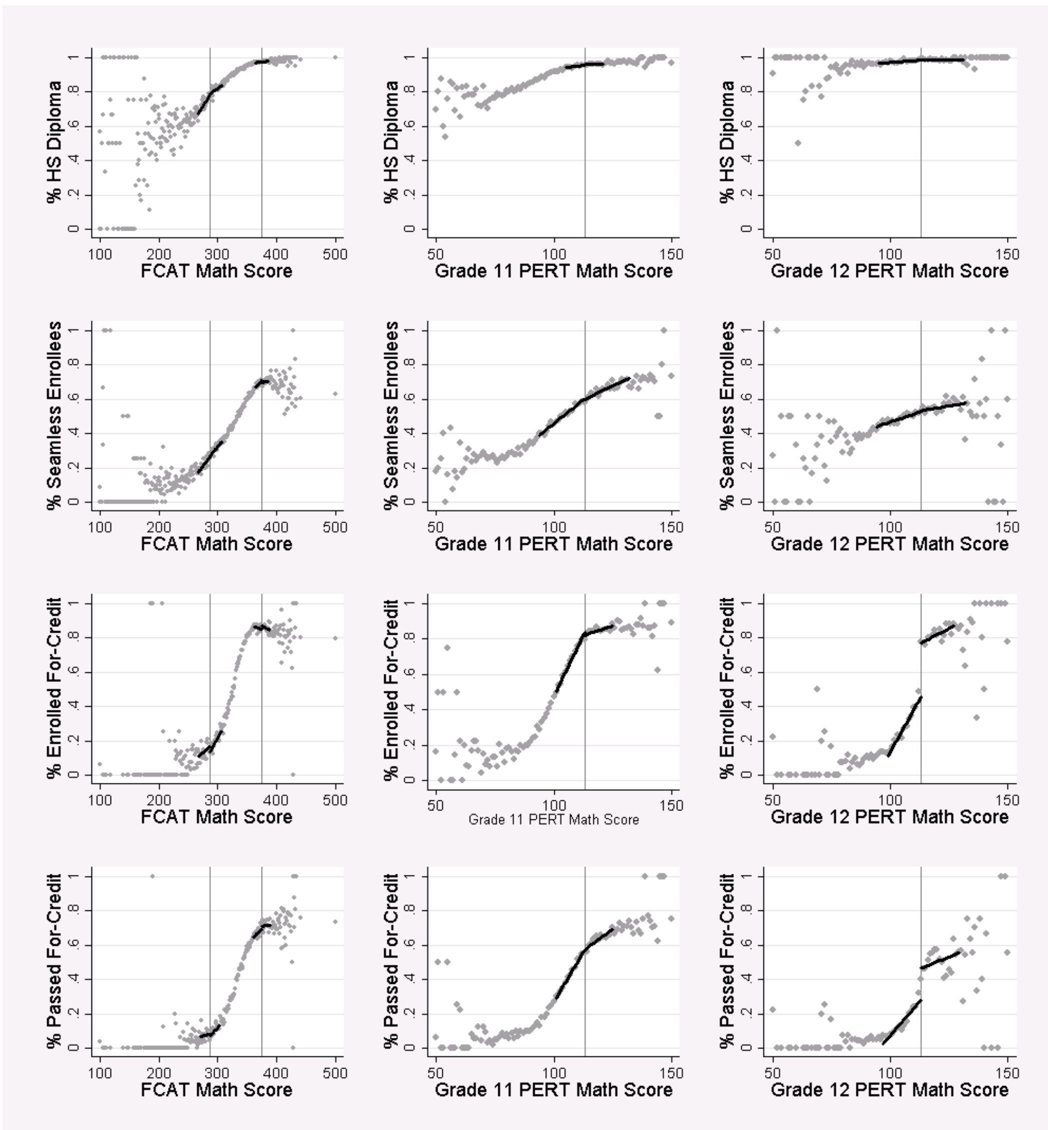


Figure 4A. Relationship between year 2 outcomes and forcing variables in math, cohort M1

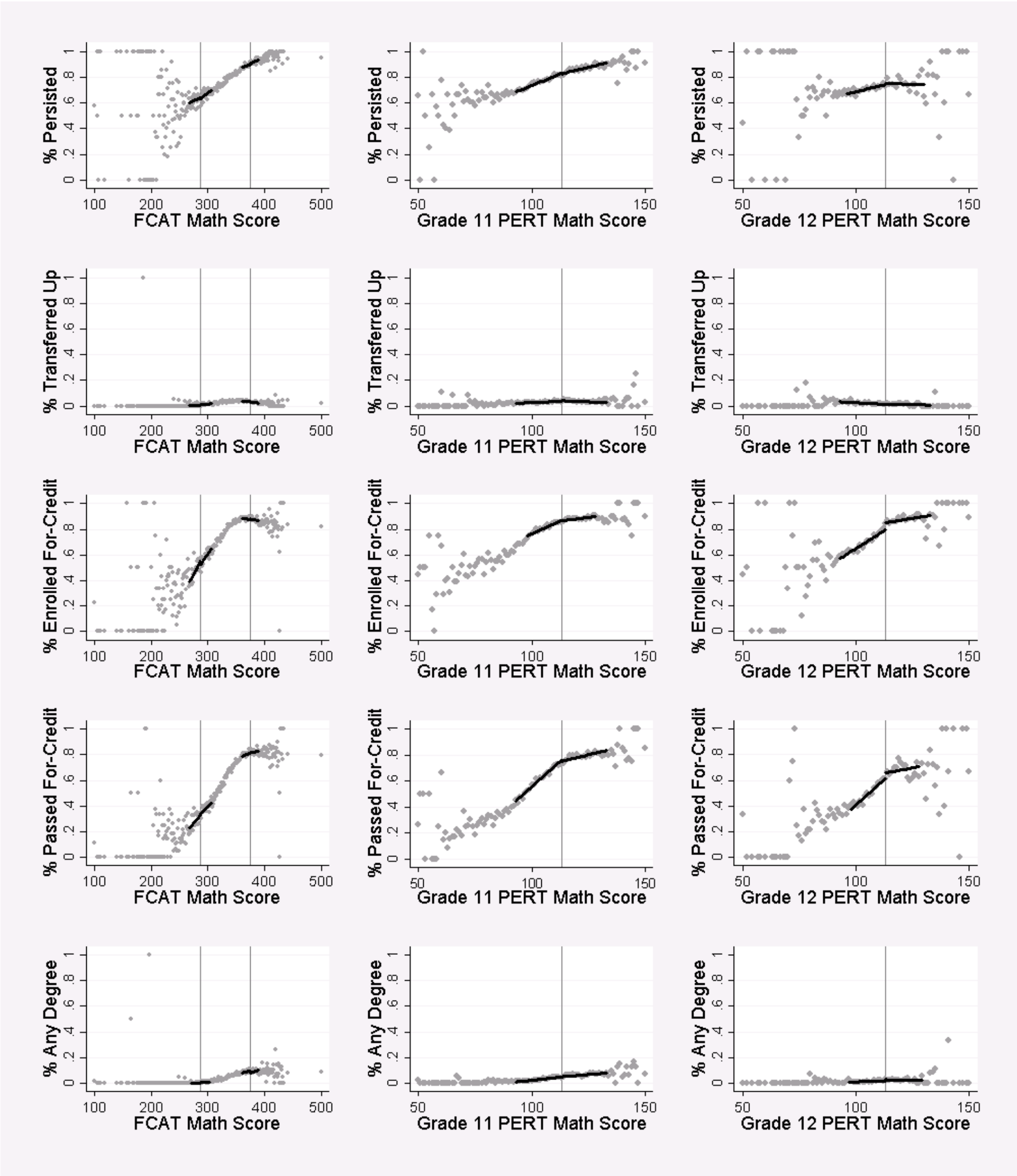


Figure 5A. Relationship between year 3 outcomes and forcing variables in math, cohort M1

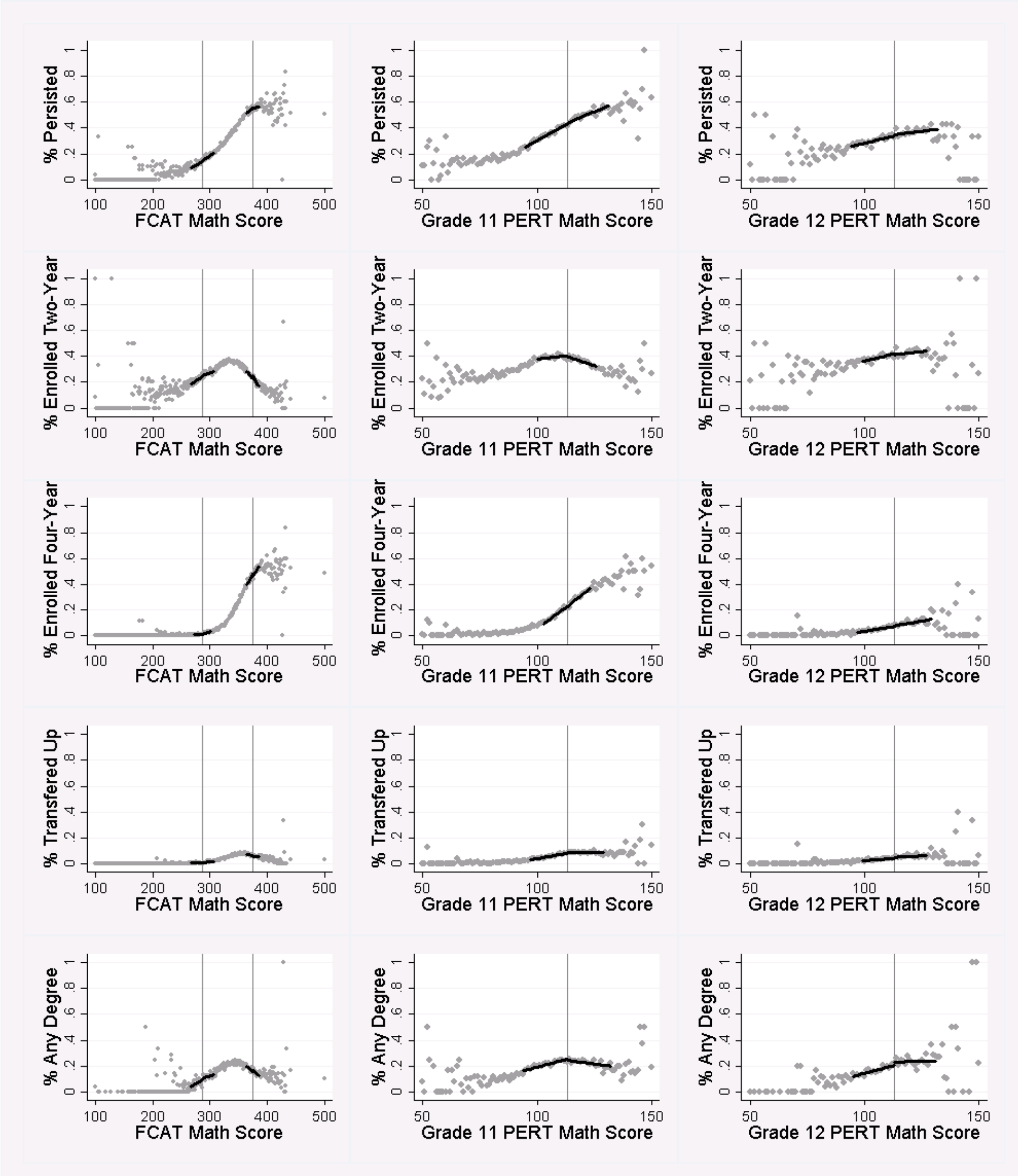


Figure 6A. Relationship between year 1 outcomes and forcing variables in reading, cohort M1

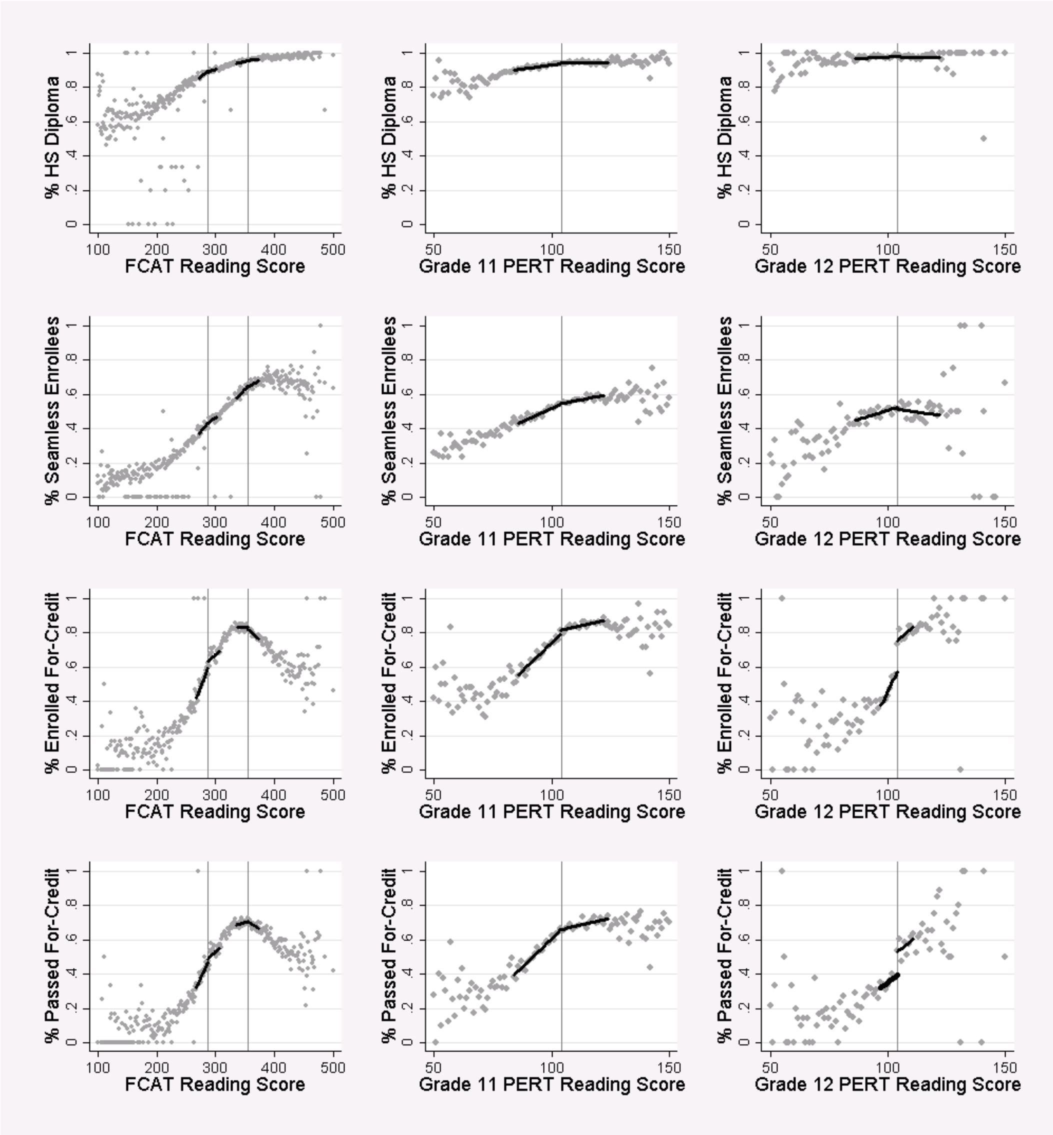


Figure 7A. Relationship between year 2 outcomes and forcing variables in reading, cohort M1

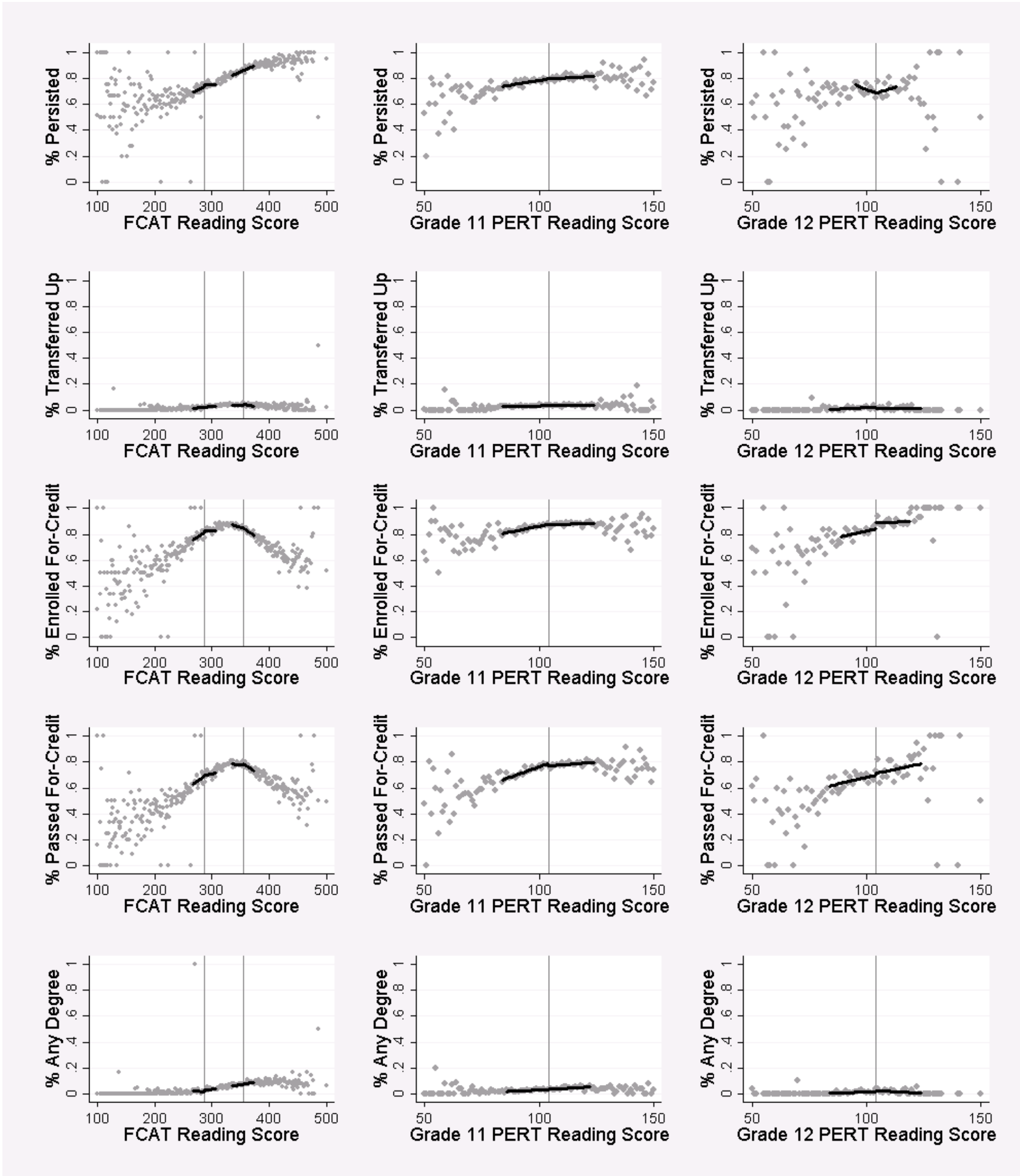


Figure 8A. Relationship between year 3 outcomes and forcing variables in reading, cohort M1

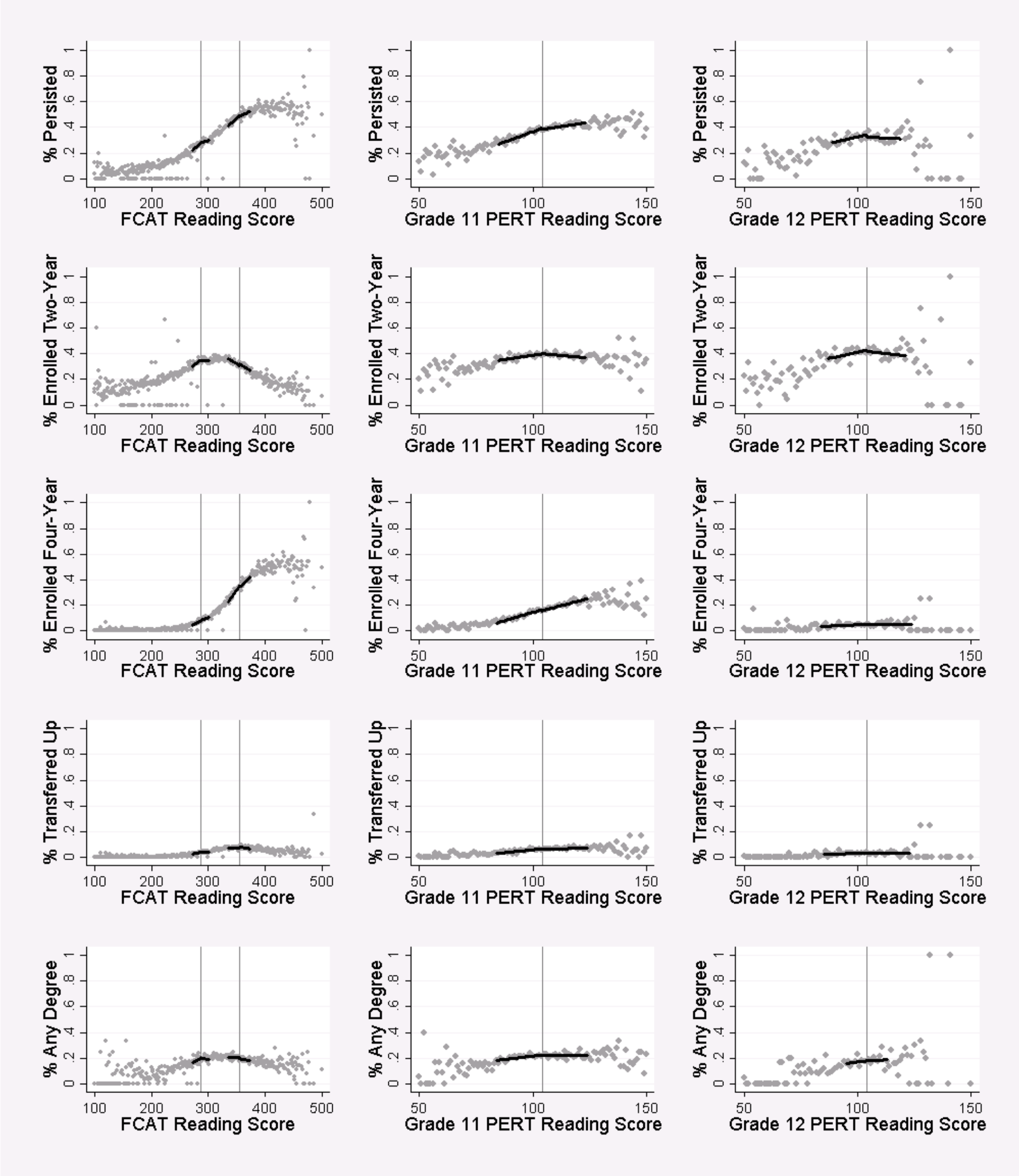


Figure 9A. Relationship between year 1 outcomes and forcing variables in math, cohort M2

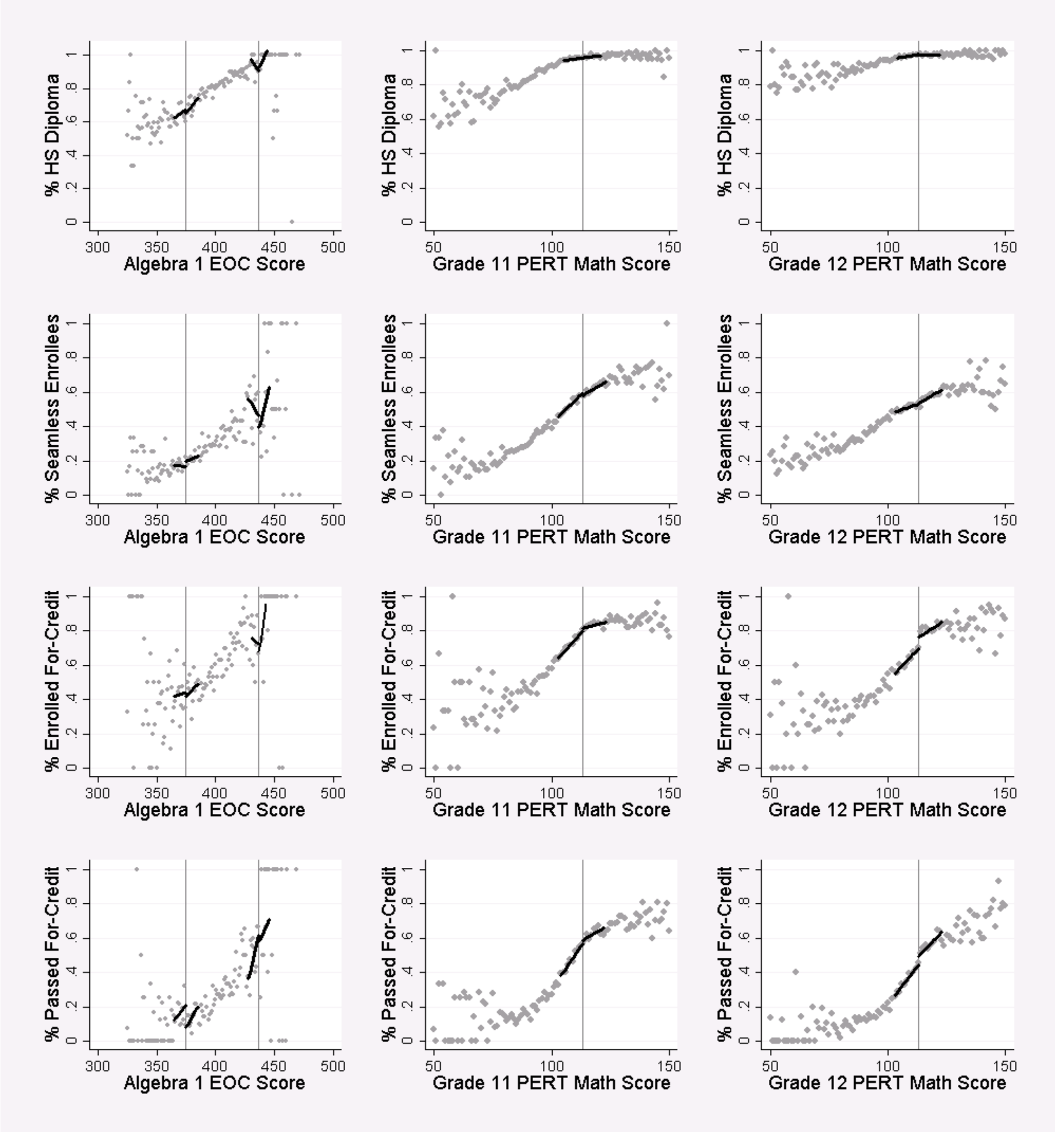


Figure 10A. Relationship between year 2 outcomes and forcing variables in math, cohort M2

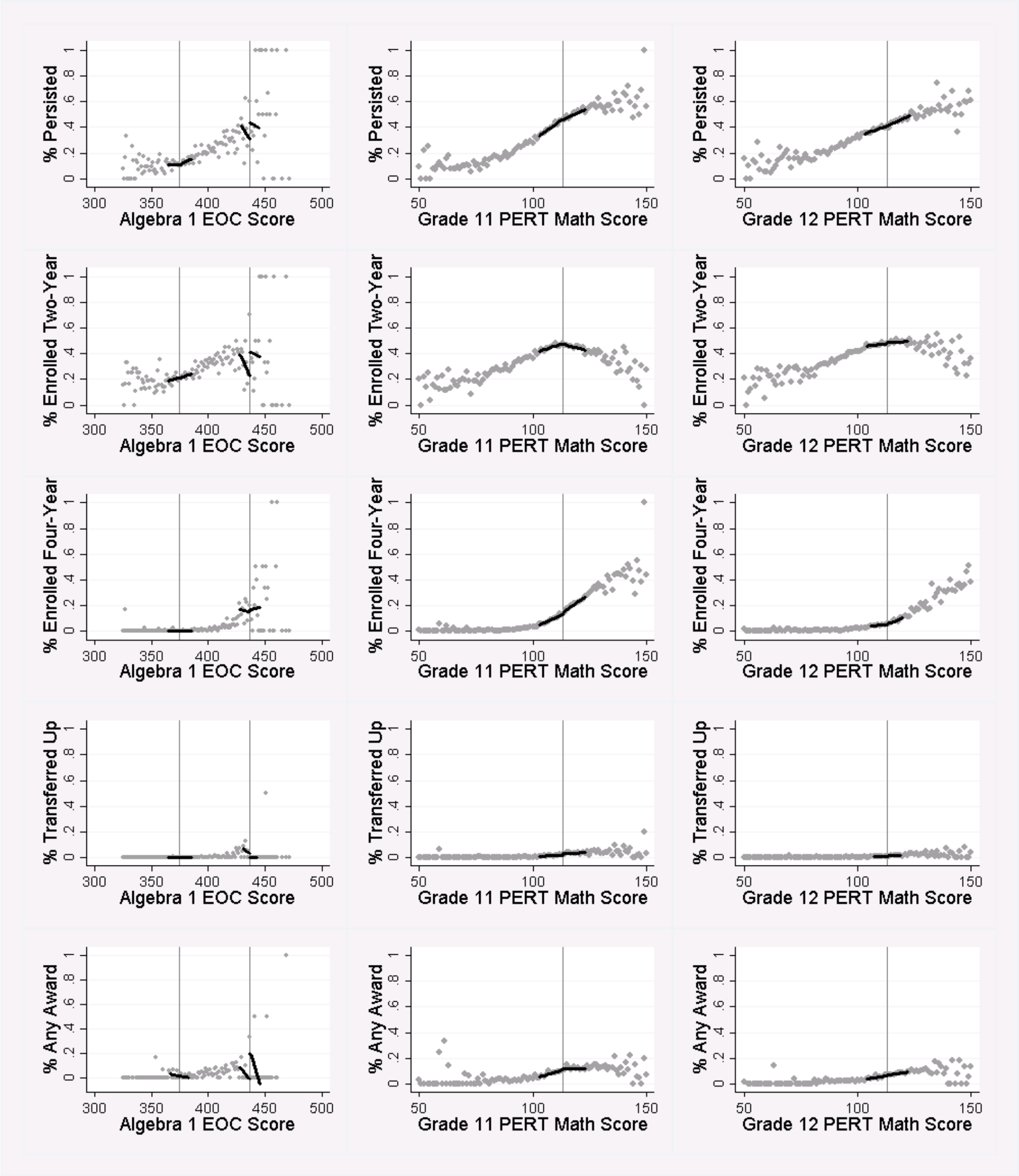


Figure 11A. Relationship between year 1 outcomes and forcing variables in reading, cohort M2

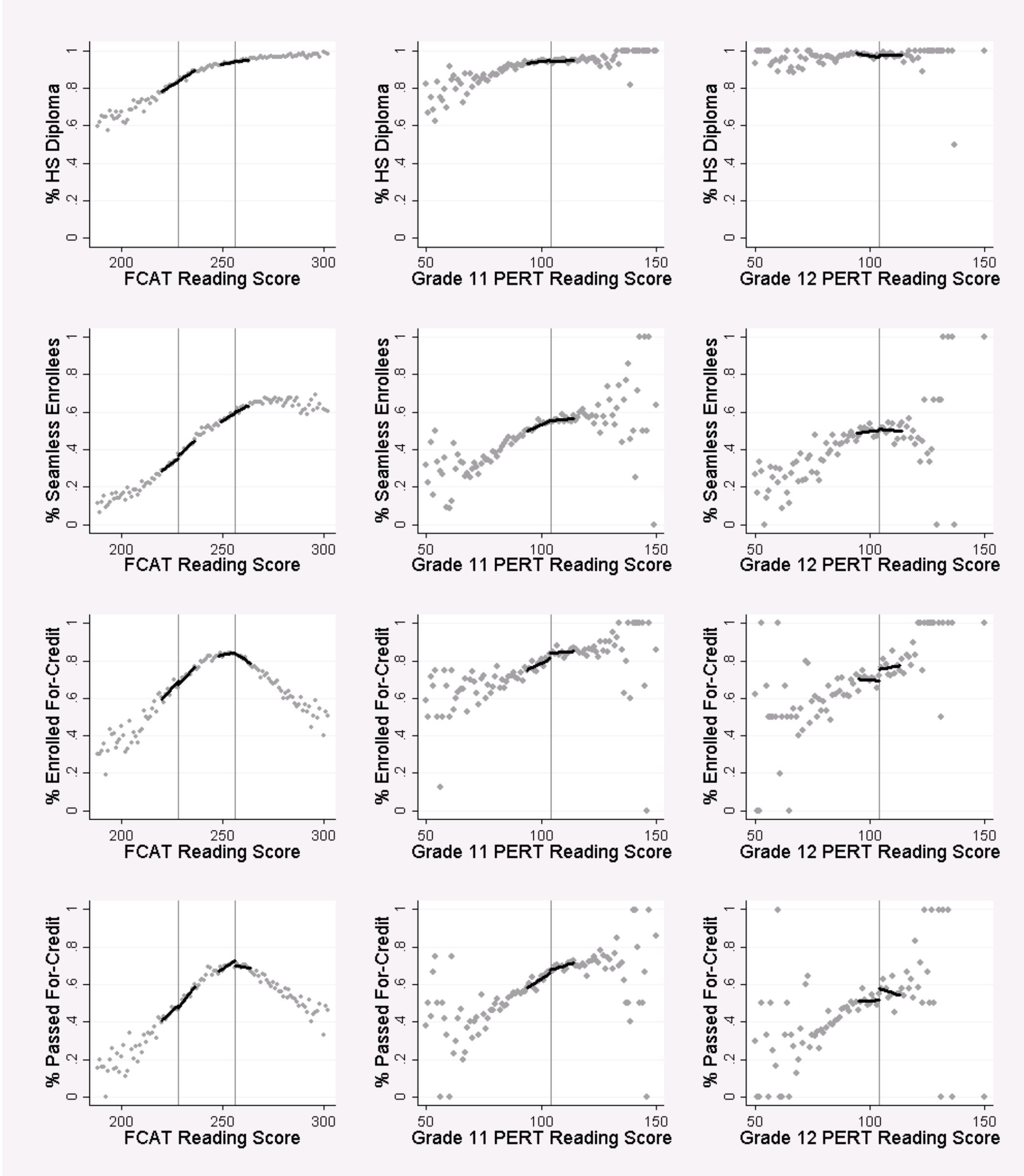
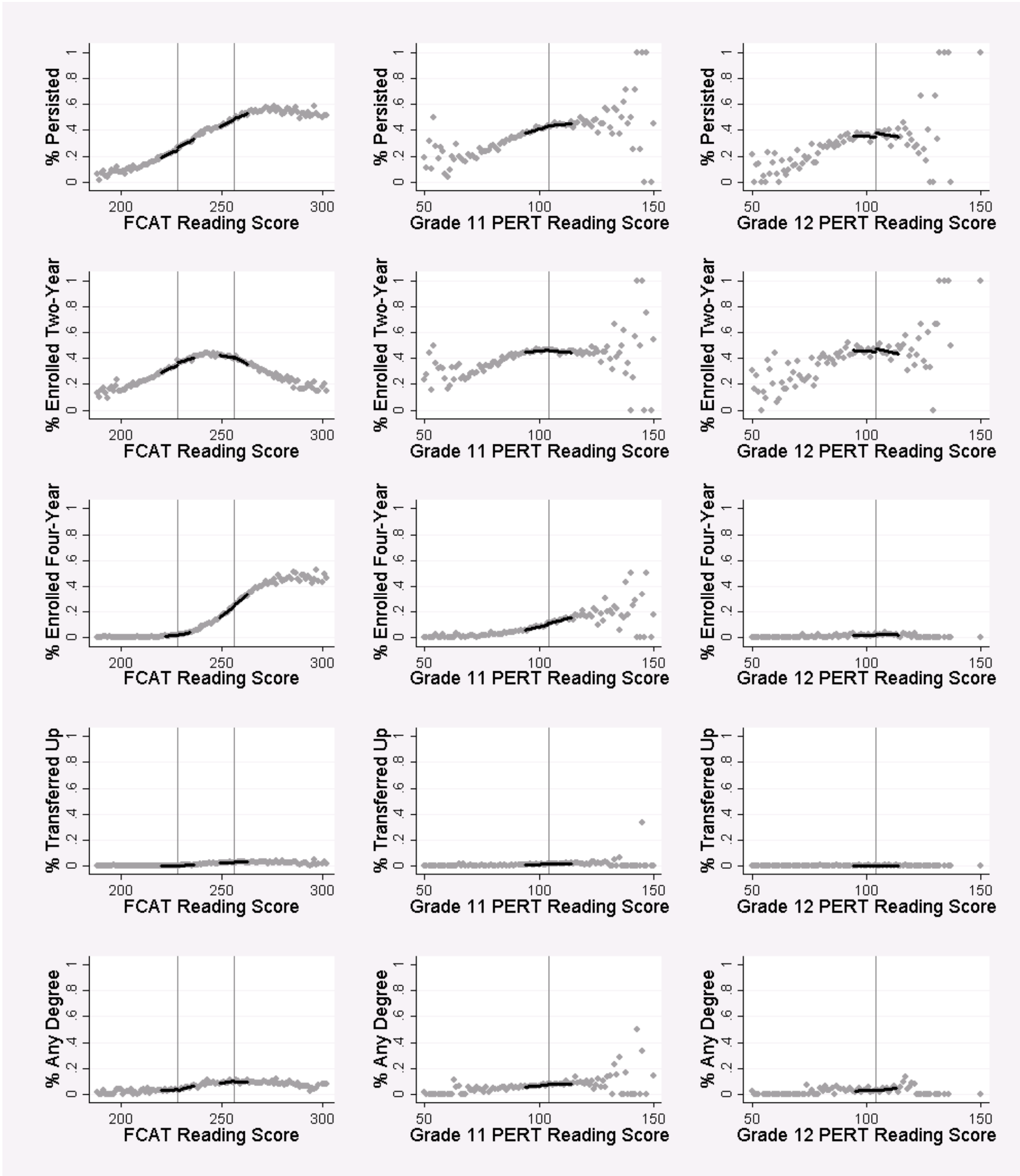


Figure 12A. Relationship between year 2 outcomes and forcing variables in reading, cohort M2



Computing bandwidths through cross-validation

This appendix section explains the cross-validation procedure used in our RD estimates and explains how this process was incorporated into a STATA program usable in a variety of contexts.

Cross-validation procedure

The use of cross-validation as a method of optimal bandwidth selection in RD studies is presented in Imbens and Lemieux (2008, pp. 627–629) and expanded upon in Lee and Lemieux (2010, p. 321). The process uses a similar principle as the RD analysis itself, but applied to a broader range of values. RD analysis, at its crux, examines the difference between the predicted value of an outcome variable at the cutoff value and the actual (average) value at that cutoff. Cross-validation applies this logic to each point within a particular range of the cutoff.[48] Specifically, for each point X_i to the left of the cutoff, the outcome variable Y is regressed on values to the left of X_i but within the bandwidth being considered $(X_i - bw \leq X < X_i)$.[49] This is repeated for all values i within the above predetermined range of the cutoff. A similar process is used for points X_i to the right of the RD cutoff, reversing the inequality above. Predicted values $\widehat{Y}_i$ are then used to compute cross-validation statistics

$$CV_Y(h) = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \widehat{Y}_i(X_i)\right)^2 \quad (1)$$

This is repeated for all bandwidth candidates h. The optimal bandwidth is then the one than satisfies

$$h_{CV}^{opt} = \arg\min_h\{CV_Y(h)\} \quad (2)$$

The rough structure of the process is then as follows:

1. Select a bandwidth h
    a. Select a value X_i of the running variable.
        i. For each data point, estimate $Y = \alpha_0 + f(X) + \varepsilon$
            1. Use domain $(X_i - bw \leq X < X_i)$ if $X_i < c$
            2. Use domain $(X_i < X \leq X_i + bw)$ if $X_i \geq c$

[48] Imbens and Lemieux discuss the range of values that should be used for this process. However, when there are multiple cutoffs for different treatments in the distribution of the running variable, discretion may be used in determining the values to be used.

[49] This specifically does not take into account the value of Y corresponding to observation X_i.

ii. Predict $\widehat{Y}_i$

iii. Repeat as necessary for all values of i within some range of c

b. Compute $CV_Y(h) = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \widehat{Y}_i(X_i)\right)^2$

c. Repeat for all bandwidths under consideration

2. Compare $CV_Y(h)$ for all values of h
3. The value of h associated with the smallest value of $CV_Y(h)$ is the optimal bandwidth

User-written programs cv_nocov and cv_cov

We wrote a general-purpose STATA program, cv_nocov, designed to quickly run cross-validation in any setting. Another program, cv_cov, runs analogously but includes additional user-specified covariates in the analysis; as the two programs are identical except for the inclusion of covariates, the discussion below focuses on cv_nocov. Since these programs are designed to be usable in any context, they require a specific set of information to run accurately. Users must specify the outcome and running variables, the bandwidths under consideration, the cutoff value, and the maximum distance from the cutoff under consideration. Users may optionally specify whether they prefer to use a quadratic function of the running variable in the cross-validation process.[50] The program displays the cross-validation statistics of all considered bandwidths, specifies which bandwidth is optimal, and returns scalars for the optimal bandwidth and whether a quadratic specification was used. Its structure is presented here, and code is available upon request:

1. User inputs:
    a. Output variable and running variable (in that order).
    b. Any "if" conditions (optional, default is none).
    c. The cutoff value for treatment.
    d. Minimum and maximum candidate bandwidths, along with the increment to be used between the two.
    e. The maximum distance from the cutoff to run regressions.
    f. Whether to use a quadratic function of the running variable (optional, default is no).
2. Program runs:

[50] RD may be run using a flexible function of the running variable; Lee and Lemieux recommend that the function be chosen using the Akaike Information Criteria.

a. Computes recentered value of running variable around cutoff (if this has already been done, a value of 0 should be used).
b. Keeps observations outlined in “if” statement.
c. Loops over bandwidth candidates.
    i. Generates empty $\hat{Y}$ value to be filled in.
    ii. Loops over values of recentered running variable (strictly) left of the cutoff and within max distance of cutoff.
        1. Regresses output variable on re-centered running variable (strictly) less than loop value and (weakly) within loop bandwidth.
        2. Fill in $\hat{Y}$values corresponding to the loop value.
    iii. Loops over values of recentered running variable (weakly) right of the cutoff and within max distance of the cutoff.
        1. Regresses output variable on recentered running variable (strictly) greater than loop value and (weakly) within loop bandwidth.
        2. Fill in $\hat{Y}$values corresponding to the loop value.
    iv. Generates CV variable equal to the squared difference between $\hat{Y}$ and the outcome variable.
    v. Computes mean value of CV, stores this value in a local macro.
d. Restores original data set.

3. Program output:
    a. Displays all candidate bandwidths and their mean CV values.
    b. Displays optimal bandwidth corresponding to minimum CV value.
    c. Stores optimal bandwidth value in scalar r(opt).
    d. Stores whether quadratic specification used in scalar r(quad) (1 = yes, 0 = no).

Use in FCCRI project

The programs cv_nocov and cv_cov are used in all components of the FCCRI study involving RD analysis. The programs are easily adaptable to a large range of analyses; while every combination of outcome variable, running variable, and cutoff will have its own optimal bandwidth, the programs allow users to specify these.

Bandwidth selection in RD analysis involves a tradeoff between bias and statistical imprecision. When using a very small bandwidth around the cutoff for treatment, bias will be very small—linear approximations of non-linear functions are most accurate over the smallest possible range—but using a small number of data points makes precise estimation difficult. Using a large bandwidth incorporates data points far from the cutoff, which lowers standard errors but may degrade the linear approximation. Optimal bandwidth estimators try to resolve this conflict.

A brief explanation of the syntax of cv_nocov follows. The program repeats an algorithm for every combination of outcome variable, running variable, and cutoff under consideration, all of which are specified by the user. The program has the following format:

cv_nocov PERT_MATH_ INV_G11 SCORE_MA_10_PRE, c(287) bwmin(5) step(1) bwmax(20) mdist(25)

This would be interpreted as follows:

- cv_nocov is the program name, and the name of the command.
- PERT_MATH_ INV_G11 indicates whether a student took the math section of the PERT during grade 11. It is the outcome variable in our sample RD analysis.
- SCORE_MA_10_PRE is the student's grade 10 FCAT assessment score in math. It is the running variable, prior to any norming or recentering.[51]
- c(287) specifies that the cutoff for RD treatment occurs when the running variable reaches a value of 287. In this case, 287 is the cutoff between proficiency level 1 and 2 in math, the lowest score at which students are required to take the math section of the PERT.
- bwmin(5) and bwmax(20) specify that the smallest bandwidth under consideration will be 5 FCAT points, while the largest will be 20 FCAT points.
- step(1) specifies that evaluated bandwidths increase in increments of 1. Thus, the above code would consider bandwidths of 5, 6, 7, and so on up to 20 FCAT points. Discretely valued running variables will likely use step(1).
- mdist(25) specifies that points up to 25 FCAT points away from the cutoff be considered in the cross-validation process.

[51] A centered version of the running variable could be used here. In that event, c(287) would be changed to c(0). In the program cv_cov, the user would specify additional covariates after the running variable.

- Had the user included the word “quadratic” anywhere after the comma, cross-validation would use a quadratic function of SCORE_MA_10_PRE rather than a linear function.

The program returns two values that may be accessed by the user:

- r(opt) returns the optimal bandwidth among those considered.
- r(quad) indicates whether the quadratic option was used (1 = yes, 0 = no).

The program does not run the actual RD analysis; this is left to the user.

Appendix B: Technical Appendix for the Exploratory Impact Analysis

This appendix includes supplemental tables to support the exploratory impact analysis. Table 1B provides mean comparisons of student- and school-level variables by sample inclusion status. Table 2B summarizes the comparison of baseline equivalence of the treatment and comparison groups for each subject and sample.

Table 1B. Mean comparisons of student- and school-level variables by sample inclusion status

	Math			English		
School-Level Comparison	**In Sample**	**Out of Sample**	**Std Mean Difference**	**In Sample**	**Out of Sample**	**Std Mean Difference**
Avg Cumulative GPA	2.7	2.7	-0.038	2.7	2.7	-0.015
Avg grade 10 math FCAT	2027	2034	-0.076	2034	2039	-0.055
Avg grade 10 English FCAT	2035	2049	-0.063	2036	2041	-0.047
% FRPL	61.5	55.6	0.119	58.4	55.4	0.062
School-Level Comparison						
Targeted in math (%)	86.3	84.4	-0.28	86.0	83.6	-0.36
Targeted in English (%)	55.1	53.9	-0.15	54.9	53.5	-0.16
ELL (%)	6.9	6.1	-0.12	6.3	6.7	0.06
Special education (%)	12.2	11.0	-0.23	12.2	10.1	-0.41
Total Enrollment	1654	1620	-0.04	1653	1601	-0.06
Student/teacher ratio	17	16	-0.20	16	17	0.08
FRPL (%)	42.8	38.2	-0.25	41.5	37.8	-0.20
Female (%)	48.5	48.9	0.09	48.5	49.1	0.12
Asian (%)	1.9	2.4	0.26	1.9	2.6	0.30
Hispanic (%)	23.6	18.5	-0.24	20.6	21.3	0.03
Black (%)	22.7	21.7	-0.05	21.8	22.8	0.05
White (%)	49.1	54.4	0.19	52.8	50.1	-0.10
Rural (%)	25.2	28.1	0.07	28.0	24.2	-0.09
Suburb/Town (%)	55.0	44.9	-0.02	50.5	47.7	-0.06
City (%)	19.7	27	0.17	21.5	28.1	0.16
Attendance (daily average %)	93.2	93.8	0.28	93.3	94.0	0.27

School-Level Comparison	Math In Sample	Math Out of Sample	Math Std Mean Difference	English In Sample	English Out of Sample	English Std Mean Difference
Dropout rate (%)	1.7	1.4	-0.20	1.7	1.2	-0.34
Attending 4-year college (%)	15.7	18.5	0.25	15.7	20.4	0.43
Attending 2-year college (%)	35.2	32.9	-0.24	34.9	32.0	-0.30
N	218	256		321	153	

GPA = grade point average; FCAT = Florida Comprehensive Assessment Test; FRPL = free and reduced-price lunch ELL = English language learner. Results are for cohorts V2, V3, and M1. Schools are included in the sample based on whether or not they switched from low- to high-compliance. Only 5 schools are always low compliance in math and English and only 1 school is always high compliance in math (none in English).

Table 2B. Baseline equivalence of treatment and comparison groups, by targeted subject and sample

Variable	Math: Full	Math: Seamless College Enrollment	English: Full	English: Seamless College Enrollment
Grade 10 cum. GPA	0.005	0.019	0.007	0.012
Math FCAT	0.003	-0.005	0.004	0.000
Reading FCAT	0.016	-0.012	0.034	0.031
FRPL	-0.080	-0.088	-0.079	-0.086
ELL	0.013	-0.002	0.033	0.023
Special education	-0.052	-0.062	-0.068	-0.066
Female	-0.002	-0.009	0.001	-0.016
Asian	-0.004	-0.012	0.004	-0.006
Black	-0.007	-0.008	0.023	0.013
Hispanic	-0.008	-0.002	0.014	0.022
White	0.008	0.004	-0.031	-0.033
Native English	0.026	0.027	-0.006	-0.006

GPA = grade point average; FCAT = Florida Comprehensive Assessment Test; FRPL = free and reduced-price lunch ELL = English language learner. Table reports the Hedges' g for both samples. Results are for cohorts V2, V3, and M1.

Appendix C: Technical Appendix for the Cost Analysis

Table 1C. FCCRI impact by district - math FCAT, cohort M1

Upper FCAT Margin					
Level Outcome	Lower DE Pass	Upper DE Enroll	 Pass	For-Credit Enroll	 Pass
Broward	-0.0017 (0.0035) [N=1,157]	0.0205 * (0.0111) [N=797]	0.0090 ** (0.0035) [N=1,626]	-0.0205 ** (0.0103) [N=1,400]	-0.0785 * (0.0403) [N=1,626]
Hillsborough	-	0.0077 (0.0094) [N=591]	-0.0042 (0.0058) [N=1,193]	-0.0248 * (0.0143) [N=1,029]	0.0071 (0.0435) [N=1,193]
Miami-Dade	-0.0018 (0.0084) [N=1,108]	-0.0028 (0.0163) [N=759]	0.0062 (0.0101) [N=1,526]	-0.0008 (0.0154) [N=1,323]	-0.0901 * (0.0469) [N=1,526]
Orange	-0.0141 * (0.0085) [N=715]	0.0051 (0.0110) [N=476]	0.0139 (0.0112) [N=986]	0.0090 (0.0228) [N=859]	-0.0256 (0.0519) [N=986]
Palm Beach	-	0.0064 (0.0112) [N=623]	0.0116 (0.0084) [N=1,257]	-0.0049 (0.0134) [N=1,057]	0.0972 ** (0.0431) [N=1,257]
Pinellas	-	0.0059 (0.0110) [N=427]	-0.0071 (0.0113) [N=830]	-0.0170 (0.0182) [N=714]	-0.0137 (0.0545) [N=830]
Statewide	-0.0013 (0.0012) [N=11,125]	0.0030 (0.0034) [N=7,734]	0.0016 (0.0023) [N=15,569]	-0.0032 (0.0038) [N=13,370]	-0.0152 (0.0128) [N=15,569]
Bandwidth	10	7	14	12	14

Note. Results are for cohort M1 only. Estimates reflect the impact of assignment to treatment—this means the impact of being just above the lower FCAT cutoff, just below the upper FCAT cutoff, and just below the grade 11 PERT cutoff. Point estimates are listed in percentage points, rather than percentages or elasticities. Standard errors (clustered by district) in parentheses, sample Ns in brackets. Districts listed are those with at least 1000 respondents within 20 test points of two or more cutoffs. Results for other districts available upon request. * = significant at the 10% level, ** = significant at the 5% level, *** = significant at the 1% level.

Table 2C. FCCRI impact by district - FCAT reading, cohort M1

Lower FCAT Margin					
Level	Lower DE	Upper DE		For-Credit	
Outcome	Pass	Enroll	Pass	Enroll	Pass
Broward	0.0217 (0.0212) [N=1,493]	0.0073 (0.0443) [N=1,493]	-0.0089 (0.0421) [N=1,416]	-0.0058 (0.0462) [N=1,493]	-0.0437 (0.0490) [N=1,493]
Miami-Dade	-0.0410 ** (0.0199) [N=1,904]	-0.0624 (0.0407) [N=1,904]	-0.0261 (0.0412) [N=1,783]	0.1078 *** (0.0417) [N=1904]	0.0495 (0.0440) [N=1,904]
Statewide	-0.0071 (0.0073) [N=12,314]	-0.0340 ** (0.0153) [N=12,314]	-0.0248 * (0.0147) [N=11,592]	0.0445 *** (0.0159) [N=12,314]	0.0245 (0.0172) [N=12,314]
Bandwidth	20	20	19	20	20

Upper FCAT Margin					
Level	Lower DE	Upper DE		For-Credit	
Outcome	Pass	Enroll	Pass	Enroll	Pass
Broward	0.0041 (0.0047) [N=1,908]	0.0047 (0.0151) [N=1,630]	0.0081 (0.0127) [N=1,630]	-0.0128 (0.0158) [N=1,814]	-0.0263 (0.0311) [N=2,021]
Hillsborough	-0.0018 (0.0046) [N=1,184]	-0.0000 (0.0125) [N=1,015]	0.0016 (0.0113) [N=1,015]	0.0080 (0.0130) [N=1,130]	-0.0040 (0.0381) [N=1,242]
Miami-Dade	-0.0071 (0.0067) [N=2,203]	-0.0273 * (0.0160) [N=1,874]	-0.0231 (0.0156) [N=1,874]	0.0371 ** (0.0168) [N=2,096]	-0.0143 (0.0300) [N=2,331]
Orange	-0.0025 (0.0065) [N=1,055]	-0.0086 (0.0158) [N=877]	-0.0175 (0.0141) [N=877]	0.0038 (0.0187) [N=1,005]	-0.0439 (0.0423) [N=1,117]
Palm Beach	0.0043 (0.0071) [N=1,497]	-0.0194 (0.0144) [N=1,277]	-0.0095 (0.0122) [N=1,277]	0.0182 (0.0162) [N=1,413]	0.0253 (0.0329) [N=1,583]
Statewide	0.0010 (0.0016) [N=17,889]	-0.0048 (0.0042) [N=15,181]	-0.0029 (0.0039) [N=15,181]	0.0043 (0.0045) [N=16,979]	-0.0087 (0.0105) [N=18,822]
Bandwidth	19	16	16	18	20

Note. Results are for cohort M1 only. Estimates reflect the impact of assignment to treatment—this means the impact of being just above the lower FCAT cutoff, just below the upper FCAT cutoff, and just below the grade 11 PERT cutoff. Point estimates are listed in percentage points, rather than percentages or elasticities. Standard errors (clustered by district) in parentheses, sample Ns in brackets. Districts listed are those with at least 1000 respondents within 20 test points of two or more cutoffs. Results for other districts available upon request. * = significant at the 10% level, ** = significant at the 5% level, *** = significant at the 1% level.

Table 3C. FCCRI impact by district - PERT (math and reading), cohort M1

	Math			Reading		
Level	Any DE	For-Credit		Any DE	For-Credit	
Outcome	Pass	Enroll	Pass	Pass	Enroll	Pass
Broward	0.0261 (0.0244) [N=1,367]	0.0305 (0.0258) [N=2,208]	-0.0325 (0.0393) [N=2,208]	-0.0118 (0.0231) [N=2,188]	0.0113 (0.0258) [N=2,133]	-0.0153 (0.0318) [N=2,239]
Duval	-0.0202 (0.0485) [N=449]	0.0373 (0.0468) [N=731]	0.0642 (0.0663) [N=731]		-	
Hillsborough	-0.0055 (0.0317) [N=1,020]	0.0185 (0.0303) [N=1,593]	0.0896 ** (0.0445) [N=1,593]	0.0129 (0.0262) [N=1,291]	-0.0085 (0.0293) [N=1,259]	-0.0387 (0.0432) [N=1,318]
Miami-Dade	-0.0134 (0.0185) [N=2,317]	0.0226 (0.0193) [N=3,788]	-0.0523 * (0.0299) [N=3,788]	0.0281 (0.0182) [N=3,509]	-0.0207 (0.0190) [N=3,420]	-0.0118 (0.0248) [N=3,581]
Orange	0.0093 (0.0342) [N=932]	0.0494 (0.0330) [N=1,486]	0.0578 (0.0455) [N=1,486]	0.0263 (0.0301) [N=1,342]	-0.0199 (0.0352) [N=1,305]	-0.0181 (0.0434) [N=1,363]
Palm Beach	0.0838 ** (0.0340) [N=1,048]	-0.0039 (0.0312) [N=1,647]	-0.0936 ** (0.0444) [N=1,647]	0.0593 ** (0.0278) [N=1,558]	-0.0837 *** (0.0302) [N=1,524]	-0.0420 (0.0375) [N=1,601]
Pinellas	-0.0037 (0.0282) [N=779]	0.0025 (0.0290) [N=1,197]	0.0765 (0.0495) [N=1,197]		-	
Seminole	0.0172 (0.0157) [N=615]	-0.0000 (0.0169) [N=899]	-0.0879 (0.0539) [N=899]		-	
Statewide	0.0067 (0.0075) [N=14,972]	0.0172 ** (0.0076) [N=23,449]	-0.0011 (0.0115) [N=23,449]	0.0289 *** (0.0072) [N=20,393]	-0.0297 *** (0.0079) [N=19,900]	-0.0099 (0.0107) [N=20,857]
Bandwidth	7	12	12	19	18	20

Note. Results are for cohort M1 only. Estimates reflect the impact of assignment to treatment—this means the impact of being just above the lower FCAT cutoff, just below the upper FCAT cutoff, and just below the grade 11 PERT cutoff. Point estimates are listed in percentage points, rather than percentages or elasticities. Standard errors (clustered by district) in parentheses, sample Ns in brackets. Districts listed are those with at least 1000 respondents within 20 test points of two or more cutoffs. Results for other districts available upon request. * = significant at the 10% level, ** = significant at the 5% level, *** = significant at the 1% level.

References

Achieve and the Education Trust. (2008). *Making college and career readiness the mission for high schools: A guide for state policymakers.* Retrieved from http://www.achieve.org/measuresthatmatter

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college.* Washington, D.C.: U.S. Department of Education. Retrieved from www.ed.gov/rschstat/research/pubs/toolboxrevisit/index.html

Adelman, C. (1999). *Answers in the toolbox: Academic intensity, attendance patterns, and bachelor's degree attainment.* U.S. Department of Education. Washington, DC: Office of Educational Research and Improvement. Retrieved from http://www.postsecconnect.org/files/AnswersintheToolbox.pdf

Anderson, R., & Chang, B. (2011). Mathematics course-taking in rural high schools. *Journal of Research in Rural Education, 26*(1), 1–10.

Attewell, P., & Domina, T. (2008) Raising the bar: Curricular intensity and academic performance. *Educational Evaluation and Policy Analysis, 30*(1). Retrieved from http://www.jstor.org/stable/30128052

Bailey, T., Jeong, D. W., & Cho, S. W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review, 29*(2), 255–270.

Barnett, E. A., Fay, M. P., Trimble, M. J., & Pheatt, L. (2013). *Reshaping the college transition: Early college readiness assessments and transition curricula in four states.* New York, NY: Columbia University, Teachers College, Community College Research Center.

Betts, J. R., & Grogger, J. (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review, 22*(4), 343–352.

Bilsky, J. (2011). Crossing boundaries: Postsecondary readiness assessment in Florida. *Change: The Magazine of Higher Learning, 43*(4), 20–24.

Bilsky, J., & Tappen, M. J. (2010, March 29). *High school students completing Math/Reading/Writing for College Success course and college placement determinations.* Memorandum. Tallahassee, FL: Florida Department of Education.

Bilsky, J., & Tappen, M. J. (2009, March 20). *Senate Bill 1908: High school reading for college success – placement and exit requirements.* Memorandum. Tallahassee, FL: Florida Department of Education.

Boser, U., & Burd, S. (2009). *Bridging the gap: How to strengthen the PK-16 pipeline to improve college readiness.* New America Foundation, Washington, DC.

Burdman, P. (2011). *Testing ground: How Florida schools and colleges are using a new assessment to increase college readiness. Achieving the Dream, Jobs for the Future, and MDC.* Retrieved from http://www.jff.org/publications/education/testing-ground-how-florida-schools-and-c/1307

Center for Community College Student Engagement. (2016). *Expectations meet reality: The underprepared student and community colleges.* Austin, TX: The University of Texas at Austin, College of Education, Department of Educational Administration, Program in Higher Education Leadership.

Conley, D. T. (2007). *Redefining college readiness, Volume 3*. Eugene, OR: Educational Policy Improvement Center.

Daugherty, L, Martorell, P., & McFarlin, I. (2014). Percent Plans, Automatic Admissions, and College Outcomes. *IZA Journal of Labor Economics* 3(10).

DiNardo, J., & Lee, D. (2004). Economic Impacts of New Unionization on Private Sector Employers: 1984–2001. *The Quarterly Journal of Economics 119*(4), 1383–1441.

Fla. Stat. § 1008.30 (2011)

Fla. Stat. § 1008.30 (2008)

Florida College System. (2012a). *How do Florida College System first-time in college degree seeking students perform in developmental Education?* Retrieved from https://www.floridacollegesystem.com/sites/www/Uploads/Publications/TAPPs/q3.pdf

Florida College System. (2012b). *Postsecondary Education Readiness Test (PERT) Frequently asked questions.*

Florida Department of Education. (2015). *Assessment Investigation: February 18, 2015.* Retrieved from http://www.fldoe.org/core/fileparse.php/12003/urlt/CommAssessmentInvestigationReport.pdf

Florida Department of Education (2013). *Performance on college placement tests: Florida's public high school graduates.* Retrieved from http://www.fldoe.org/articulation/perfCPT/default.asp

Florida Department of Education (2010). *College and Career Readiness Initiative: College placement testing results for high school students.* Retrieved from http://www.fldoe.org/gr/li/CCRI_EPRA-QA.pdf

Florida Department of Education. (2008). *College and Career Readiness Initiative Expanded Postsecondary Readiness Assessment: Frequently asked questions.* Memo prepared by the Florida Department of Education. Tallahassee, FL.

Florida Department of Education (2005). *Postsecondary student success begins with high school preparation. Data Trend report 33.* Retrieved from http://www.fldoe.org/cc/OSAS/DataTrendsResearch/DT33.pdf

Florida Office of Program Policy Analysis and Government Accountability. (2007). *Half of college students needing remediation drop out; remediation completers do almost as well as other students (Report 07-31).* Retrieved from http://www.oppaga.state.fl.us/reports/pdf/0731rpt.pdf

Furchtgott-Roth, D., Jacobson, L., & Mokher, C.G. (2009). *Strengthening community colleges' influence on economic mobility.* Report prepared for the Pew Charitable Trusts by Hudson Institute and CNA. http://www.economicmobility.org

Horn, L., & Kojaku, L.K. (2001). *High school academic curriculum and the persistence path through college (NCES 2001–163).* U.S. Department of Education. Washing-ton, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2001163

Hu, S., Park, T., Woods, C., Tandberg, D., Richard, K, & Hankerson, D. (2016). *Investigating developmental and college-level course enrollment and passing before and after Florida's developmental education reform.* Report prepared for the National Center for Education Evaluation and Regional Assistance.

Hurwitz, M., Smith, J., Niu, S., & Howell, J. (2015) The Maine question: How is 4-year college enrollment affected by mandatory college entrance exams? Educational *Evaluation and Policy Analysis 37*(1), 138–159. doi: 10.3102/0162373714521866

Hyman, J. (2017). ACT for all: The effect of mandatory college entrance exams on postsecondary attainment and choice. *Education Finance and Policy, 12*(3), 281-311.

Imbens, G.W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*(2), 615–635.

Jacob, B.A. (2001). Getting tough? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis, 23*(2), 99–121.

Jacobson, L., & Mokher, C.G. (2009). *Rural-urban differences in the outcomes associated with career and technical education in Florida.* Report prepared for the United States

Department of Education's National Assessment of Career and Technical Education by CNA.

Johnson, J., & Strange, M. (2008). *Why rural matters 2007: The realities of rural education growth.* A report by the Rural School and Community Trust Policy Program. Arlington, VA.

Kirst, M. W. (2005). *Separation of K-12 and postsecondary education governance and policymaking: Evolution and impact (Report #17):* The Bridge Project, Stanford University.

Kirst, M. W. (2001). *Overcoming the senior slump: New education policies.* Washington, DC: The Institute for Educational Leadership. (ERIC Document Reproduction No. ED455720).

Kirst, M. W., & Bracco, K. R. (2004). *Bridging the great divide: How the K-12 and postsecondary split hurts students, and what can be done about it.* In M. W. Kirst & A. Venezia (Eds.), *From high school to college: Improving opportunities for success in postsecondary education* (pp. 1–30). San Francisco, CA: Jossey-Bass.

Kurlaender, M., Jackson, J., Grodsky, E., & Howell, J. (2016). *Ready or Not? California's Early Assessment Program and the Transition to College.* Paper session presented at the meeting of American Educational Research Association, Washington, DC.

Lee, S. (2016). *Florida High School Student Performance on the Postsecondary Education Readiness Test (PERT): Results for 11th-Graders in 2011/12*. Report prepared for the Florida Department of Education. Arlington, VA: CNA.

Lee, D.S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*(2), 281–335.

Leeds, D. M., & DesJardins. S.L. (2015). The effect of merit aid on enrollment: A regression discontinuity analysis of Iowa's National Scholars Award. *Research in Higher Education 56*(5), 471–495.

Levin, H.(2007). What are the mechanisms of high-poverty disadvantages? On the relationship between poverty and curriculum. *North Carolina Law Review, 85*(1381), 1–38.

Levin, H. M., & McEwan, P. J. 2001. *Cost-effectiveness analysis: methods and applications.* 2nd ed. Thousand Oaks, CA: Sage

Lillard, D. R., & DeCicca, P. P. (2001). Higher standards, more dropouts? Evidence within and across time. *Economics of Education Review, 20*(5), 459–473.

Martorell, P., & McFarlin, I. (2011). Help or hindrance? The effects of college remediation on academic and labor market outcomes. *The Review of Economics and Statistics, 93*(2), 436–454.

McCormick, N., & Lucas M. (2011). Exploring mathematics college readiness in the United States. *Current Issues in Education, 14*(1).

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Mokher, C. G., & Jacobson, L. (2014). *Assessment of the Florida College and Career Readiness Initiative: Year 2 Report*. Report prepared for the Florida Department of Education. Arlington, VA: CNA.

Mokher, C. G., Jacobson, L., Rosenbaum, J., & LaLonde, R. (2013). *Assessment of the Florida College and Career Readiness Initiative: Year 1 Report.* Report prepared for the Florida Department of Education. Arlington, VA: CNA.

Mokher, C. G., Leeds, D. M., Harris, J., & Geraghty, T. (2016). *Assessment of the Florida College and Career Readiness Initiative: 2016 Interim Report.* Report prepared for the Florida Department of Education. Arlington, VA: CNA.

National Commission on the High School Senior Year. (2001, January). *The lost opportunity of senior year: Finding a better way.* Princeton, NJ: The Woodrow Wilson National Fellowship Foundation. (ERIC Document Reproduction No. ED 453 604).

Peterson, K. (2003). *Overcoming senior slump: The community college role.* (ERIC Digest (EDO-JC-03-01). Los Angeles: ERIC Clearinghouse for Community Colleges. Retrieved from http://www.ericdigests.org/2004-1/slump.htm

Picciano, A. G., & Seaman, J. (2009). *K–12 online learning: A 2008 follow-up of the survey of U.S. school district administrators.* Needham, MA: The Sloan Consortium.

Rosenbaum, J. (2001). *Beyond college for all.* New York: Russell Sage Foundation.

State Impact Florida. (2012). *13th grade: What Florida colleges are doing to help more students complete remedial courses.* Retrieved from https://stateimpact.npr.org/florida/2012/12/18/13th-grade-what-florida-colleges-are-doing-to-help-more-students-complete-remedial-courses/

Stewart, P. (2011). *Mathematics and English language arts college success, readiness, and prep course enrollments*. Memorandum. Tallahassee, FL: Florida Department of Education.

Strong American Schools. (2008). *Diploma to nowhere*. Retrieved from http://www.deltacostproject.org/resources/pdf/DiplomaToNowhere.pdf

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far?. In *Handbook of the Economics of Education* (Vol. 3, pp. 249-277). Elsevier.

Tierney, W. G., Bailey, T., Constantine, J., Finkelstein, N., & Hurd, N. F. (2009). *Helping students navigate the path to college: What high schools can do* (NCEE #2009-4066). IES Practice Guide. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression discontinuity approach. *International Economic Review 43*(4), 1249–1287.

Weimer, D.L. (2015). The thin reed: Accommodating weak evidence for critical parameters in cost-benefit analysis. *Risk Analysis 35*(6), 1101-1113.

What Works Clearinghouse. (2017). *Standards Handbook Version 4.0.* Retrieved from: https://ies.ed.gov/ncee/wwc/Handbooks

What Works Clearinghouse. (2015). *Postsecondary education evidence review protocol.* Retrieved from http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=242