Educational Evaluation and Policy Analysis
March 2016, Vol. 38, No. 1, pp. 171–196
DOI: 10.3102/0162373715603504

How Can Placement Policy Improve Math Remediation Outcomes? Evidence From Experimentation in Community Colleges

Federick Ngo
Tatiana Melguizo
University of Southern California

Changing placement policy may help to improve developmental education student outcomes in community colleges, but there is little understanding of the impacts of these reforms. We take advantage of heterogeneous placement policy in a large urban community college district in California to compare the effects of math remediation under different policy contexts. District colleges either switched from using math diagnostics to using computer-adaptive tests, or raised placement cutoffs. We use quasi-experimental methods to identify the impact of remediation under each policy and the change in impact following placement policy experimentation. We find that switching to a computer-adaptive test exacerbated the penalty of remediation for marginal students and resulted in more placement errors. Modestly raising placement cutoffs had no significant effects.

Keywords: *community colleges, remedial/developmental education, placement tests, regression discontinuity*

One of the first steps an incoming community college student takes during matriculation is sitting for a placement test in reading, writing, or math. This enables the college to identify the student's preparedness for college-level academics and direct the student toward appropriately leveled coursework. Nationally, about 60% of students are referred to remedial or developmental coursework after this screening process (National Center for Public Policy and Higher Education & Southern Regional Education Board [NCPPHE & SREB], 2010), but this figure is more than 80% in states such as California, which serves about one fifth of all community college students in the country and is the setting for this study (Foundation for California Community Colleges, 2015). A prevailing concern is that the majority of students placed in developmental courses following the assessment and placement process do not progress through course sequences to complete college-level courses or earn postsecondary credentials (Bailey, Jeong, & Cho, 2010; Fong, Melguizo, & Prather, 2015). This has prompted increased attention to policies and practices that can improve remediation outcomes.

Changing assessment and placement policy is increasingly being seen as a lever to achieve this goal. Estimates from recent research suggest that placement tests, which are commonly used across the country to assess incoming students, may be mis-assigning nearly a quarter of students in math, with most of these being underplacement errors into remedial math courses that are below student skill levels (Scott-Clayton, Crosta, & Belfield, 2014). It follows that if more accurate placement instruments and measures are used, and used more accurately, then students will be more likely to complete the courses in which they are placed and persist toward their academic goals.

Aside from these prediction-based estimates, there are just a handful of studies on how actual placement policies influence and affect developmental math student outcomes. This includes research investigating how placement test results affect enrollment decisions (Martorell, McFarlin, & Xue, 2015; Scott-Clayton & Rodriguez, 2015), whether cutoffs are set correctly (Melguizo, Bos, Ngo, Mills, & Prather, in press), and the usefulness of multiple measures for making placement decisions (Ngo & Kwon, 2015). This type of evidence is important because establishing placement policies is a complex endeavor, with colleges needing to select placement instruments, set cutoffs, and decide whether to incorporate additional measures. The reality of assessment policy in community colleges is that placement measures are not routinely validated, and faculty and administrators often do not feel as though they have adequate tools and support to select and use tests and set cutoffs appropriately (Melguizo, Kosiewicz, Prather, & Bos, 2014). There is scant evidence to inform these practitioner decisions, resulting in continual experimentation with assessment policy that may or may not be beneficial to students. Some examples include the use of holistic advising in developmental education in Texas (Texas Success Initiative, Senate Bill [SB] 162) and the development of a common assessment system in California (Common Assessment Initiative, SB 1456).

In this study, we take advantage of heterogeneous assessment policy in a large urban community college district (LUCCD) in California to compare the effects of math remediation under different placement policy contexts. Because each college uses a placement test and a set of cutoffs to assign students to math courses, we use a regression discontinuity (RD) design to identify the impact of placement into lower-level math remediation in each college, before and after a placement policy change. Three district colleges switched between using a math diagnostic (Mathematics Diagnostic Testing Project, MDTP) and using a computer-adaptive test (ACCUPLACER), and three colleges raised placement cutoffs. Of these six, three serve as focus colleges for this study. We argue that within a narrow bandwidth of students around each cutoff, the RD estimate provides an indicator of the impact of placement decisions for students just below the cutoff. Comparing the differences in these RD estimates before and after each policy change provides insights into how this impact changes, and how placement policy can improve math remediation outcomes.

Using an RD design within a difference-in-difference framework to account for student achievement trends in the district, we find that there is an increase in the negative effect of remediation on students' early college outcomes after two colleges in the district switched from diagnostics to computer-adaptive tests. We focus on the cutoff between pre-algebra (PA) and elementary algebra (EA), where the majority of incoming students are assigned, and find that students just under the cutoff and placed in PA were less likely to enroll in college, less likely to attempt and complete gatekeeper math courses within a year, and completed fewer credits on average a year after taking the computer-adaptive placement test. In contrast, we observed no significant changes in the effect of remediation when colleges modestly raised test score cutoffs, and we demonstrate that these findings are robust to sensitivity analyses.

Because the switch to computer-adaptive tests may have also changed the level of the cutoff with respect to the student ability distribution, we extend the analysis and provide placement accuracy analytics that help us to better understand the mechanism by which placement policies can influence developmental math outcomes. The results indicate that the observed negative effects of remediation in colleges that switched from diagnostics to computer-adaptive tests coincided with an increase in the proportion of severe placement errors. This suggests that diagnostics may improve the accuracy of placement decisions relative to commonly used computer-adaptive tests, and potentially may be a more useful remedial screening tool in developmental education.

The article proceeds as follows: We first discuss the role of assessment and placement policies in developmental education and relevant research on the impacts of placement policies. We then describe the heterogeneous placement policy contexts of the LUCCD and outline our methodological approach to understanding the effectiveness of remediation in these contexts. We then present the findings and the sensitivity checks we conducted in conjunction with our primary analyses. Finally, we discuss the results of the study and opportunities for improving placement policies in community colleges.

Placement Policies in Developmental Math

Given that community colleges are open-access institutions that do not typically select students on the basis of academic achievement, colleges need some means of identifying student skill and directing students toward courses that can serve their academic needs (Cohen & Brawer, 2003). This typically occurs during the assessment and placement process at matriculation, and in a survey of placement testing practices in community colleges across the country, Fields and Parsad (2012) found that virtually all public 2-year colleges use a mathematics test to screen students for math remediation.

Researchers have documented the ways in which establishing assessment and placement policy is a complex task involving multiple decisions (Hughes & Scott-Clayton, 2011; Melguizo et al., 2014). One of the first judgments colleges or other coordinating bodies make is selecting which placement instruments to use. Although multiple options exist, Fields and Parsad (2012) reported that most public 2-year colleges across the country have chosen to use either the ACCUPLACER (42%) or COMPASS (60%), although some use results from the SAT (32%) or ACT (17%) and others use alternative instruments (14%).[1] ACCUPLACER and COMPASS, developed by the College Board and ACT Inc., respectively, are commercially available computer-adaptive tests that identify student skill in arithmetic, algebra, and college-level math using an algorithm that responds to student performance (Mattern & Packman, 2009).

Once an instrument is chosen, colleges or districts must then determine where to set placement cutoffs. Despite the fact that such tests as the ACCUPLACER or COMPASS provide suggestions for where cutoffs should be set, there is considerable variability in cutoffs for placement into developmental versus college-level math in community colleges across the country (Fields & Parsad, 2012). Furthermore, there is a lack of technical support for evaluating and adjusting cutoffs, although Melguizo et al. (in press) suggests a method using RD design, and Scott-Clayton et al. (2014) suggest a method using probit regressions and extrapolations. In terms of outcomes, Martorell and McFarlin (2011) examined whether the effects of remediation on college outcomes in Texas differed after the placement cutoff was raised by about a third of a standard deviation. They found slightly more negative effects after cutoffs were raised, suggesting that lower-ability students may have benefitted more from placement in remedial courses.

In addition, colleges can decide whether to incorporate additional measures such as high school grade point average (GPA) or prior math achievement, which have been demonstrated to increase student access to higher-level courses without compromising success rates, or other background measures into the placement decision (Marwick, 2004; Ngo & Kwon, 2015). In most states, and as recommended by the College Board and ACT, Inc., it is expected that institutions validate the measures used based on student achievement data and adjust policies as necessary (Fulton, 2012; Mattern & Packman, 2009). The reality, however, is that these tasks are easier said than done; community college faculty report feeling unsupported in selecting and validating placement measures and in setting placement cutoffs, meaning that implemented policies are often the result of continual experimentation (Melguizo et al., 2014).

The Impact and Influence of Placement Decisions

It is important to examine this experimentation because placement policy plays a tremendous role in shaping the pathways, opportunities, and outcomes of students in community colleges. Initial placement decisions can place students on a track of remedial coursework that can be as many as five semesters long and restrict access to college-level courses. Long remedial sequences can be extremely costly to students in terms of time and money (Melguizo, Hagedorn, & Cypers, 2008), and the more remedial courses that are required, the less likely that students will complete the sequence and earn college credit (Bailey et al., 2010). If placement polices are not achieving the goal of accurate placement into appropriate coursework, then this may affect the likelihood that students will achieve their longer-term academic goals.

Placement testing is also one of the first interactions incoming students have with postsecondary institutions. The signal that tests scores or their associated labels (e.g., "remedial") send may influence students' beliefs about their readiness for college and their ability to succeed, and be a mechanism by which community colleges discourage students from persisting and instead trigger a "cooling out" process (Clark, 1960). There is some evidence that these labels affect the college investment decisions of secondary students (Papay, Murnane, & Willett, 2011), but both qualitative and quantitative studies investigating the "discouragement hypothesis" among community college students have found that placement test results do not appear to discourage students from enrolling (Deil-Amen & Tevis, 2010; Martorell et al., 2015; Scott-Clayton & Rodriguez, 2015). However, Scott-Clayton and Rodriguez (2015) did find one group of students—those who were potentially mis-assigned in developmental English courses—who may have been discouraged by their placement exam scores and were less likely to enroll.

Placement Policy Reforms

Although placement tests do not appear to affect enrollment behavior, there has still been growing scrutiny of placement policies across the country (Burdman, 2012). There is increasing uncertainty about whether the tests and cutoff scores in use across the country are appropriate, and this has prompted reform in choice and use of tests, as well as investigation of whether utilizing additional or alternative measures such as information from high school transcripts could improve placement accuracy (Hughes & Scott-Clayton, 2011).[2] Recent research has also called the usefulness of popular placement tests into question. Studying both a large statewide community college system and a large community college district, Scott-Clayton et al. (2014) found that nearly a quarter of students may be mis-assigned to their math courses by placement tests such as the COMPASS or ACCUPLACER, with most of these being under-placement errors (i.e., students being placed into courses that are below their skill level). A meta-analysis conducted by the College Board on placement accuracy of the ACCUPLACER also found that just about two thirds of students were correctly placed into lower-level math courses (Mattern & Packman, 2009). Such findings have recently led ACT, Inc. to phase out the COMPASS placement test (Fain, 2015).

Despite the number of studies that have concluded that commonly used placement tests are only weakly if at all correlated with student outcomes (e.g., Armstrong, 2000; Jenkins, Jaggars, & Roksa, 2009; Medhanie, Dupuis, LeBeau, Harwell, & Post, 2012), there is scant research on how placement tests affect remediation outcomes and how alternatives can improve on current practices. To our knowledge, only one study has examined heterogeneity in the effects of remediation by test type. Using an RD design, Scott-Clayton and Rodriguez (2015) found that the effects of placement in remediation were more negative after the community college system in their study switched from using an in-house single-score test to a new test, the COMPASS, to make placement decisions in math. Following the districtwide switch, students at the margin of the cutoff were even less likely to take and pass college-level math than cohorts under the previous testing regime.

Diagnostics as Alternatives

What is also missing from the literature is research on how alternative placement instruments such as math diagnostics can improve placement policy. Math diagnostics may be a useful alternative in the developmental math setting because they can provide detailed information about developmental math students' content knowledge (Stigler, Givvin, & Thompson, 2010). In contrast to the computer-adaptive format, which adjusts in response to student performance and can therefore be over after just a handful of questions, diagnostics are designed to gather information on student proficiency on a range of topics and provide skill-specific information that practitioners can use to inform placement decisions and tailor classroom instruction. For example, the MDTP EA subtest, developed by a California State University (CSU) and University of California (UC) partnership, provides a report of student understanding of algebraic expressions, functions, graphing, and other math topics.

Although diagnostics such as the MDTP have been shown to improve placement accuracy in secondary school settings (Betts, Hahn, & Zau, 2011; Huang, Snipes, & Finkelstein, 2014), they are surprisingly rarely used in developmental math settings in community colleges (Burdman, 2012). In fact, we found just one recent study by Rodríguez (2014) that provides preliminary descriptive evidence on the use of diagnostics for community college placement decisions in Virginia. Following the shift to using the Virginia Placement Test (VPT), a diagnostic tool, college-level math placements increased 22% but pass rates in those courses declined 7%. Although this study provides some evidence on the effects of placement policies and the use of diagnostics, the findings may reflect the whole gamut of developmental education redesign in Virginia, which in addition to diagnostic use includes the introduction of modularized courses. Because reforms in developmental math are often bundles of assessment and instructional interventions, there is still relatively little understanding of the role of placement policies, and specifically of using alternative assessment instruments to improve placement decisions and community college student outcomes.

Placement Policy Experimentation in the LUCCD

Our study builds on this work by examining changes in assessment and placement policies and subsequent effects on student outcomes. We capitalize on the heterogeneity in placement policies in the California Community College (CCC) system, a highly decentralized system of higher education. The CCC system includes 112 colleges, roughly one tenth of all community colleges in the country, and serves more than 2 million students each year, about one fifth of all community college students in the country (Foundation for CCCs, 2015). We focus specifically on one LUCCD in California that enrolls a diverse student population: 56% identify as Latina/o, 14% as African American, 13% as Asian, and 16% as White; 51% are below the poverty line, and nearly 20% are from homes where parents received only elementary education.

The LUCCD includes nine colleges and, much like other CCCs, each has considerable autonomy over choice and use of assessment, including choice of test (e.g., ACCUPLACER, COMPASS, or MDTP), choice of cutoff, and choice of additional measures that can be factored into the placement decision (Perry, Bahr, Rosin, & Woodward, 2010). The CCC system provides some guidelines for assessment practices, such as requirements to conduct validations, but the devolved autonomy largely enables colleges to continually experiment with placement policies without having to work through district- and state-level bureaucratic structures.

Six out of the nine LUCCD colleges implemented a placement policy change within the 2005–2012 window of this study. Three colleges raised placement cutoffs, and three switched their placement tests entirely—two of these latter three colleges switched from using a diagnostic test (i.e., MDTP) to using a commercially available computer-adaptive test (i.e., ACCUPLACER), and the third made a switch in the opposite direction.[3] The other colleges made no placement policy changes during the study window, although some made an instructional change in the delivery of developmental math.[4] There were other district policies that may have influenced student outcomes during the study window, such as funding to support basic skills and budget changes related to the 2008 recession that may have affected enrollment trends. We discuss how we account for these in our empirical strategy. Doing so enables us to focus on the effects of placement policy changes in those colleges that either switched tests or raised placement cutoffs. Characteristics of the colleges in the district are described in Table 1.

Data

We utilize administrative records from the LUCCD that link students' demographic data, assessment records, course enrollment history, and academic outcomes. Students taking the assessments are those who have the intention of enrolling in courses that lead to certificates or associate's degrees, or award transfer credit for the CSU or UC systems. We excluded from the analysis students who were concurrently enrolled in high school, those who had already received an associate's or bachelor's degree at the time of testing, and those who were above 65 years old.

We further focus our analysis on students around the PA/EA cutoff because nearly two thirds of all students entering LUCCD are placed into these two levels, and because EA is considered to be a gatekeeper course (e.g., for degree completion and as a prerequisite). It is also reasonable to assume that most students have the goal of completing EA for the reasons described above. The final PA/EA sample includes 68,300 students who were placed into either PA or EA in the LUCCD colleges between 2005 and 2012. The sample distribution across district colleges is also shown in Table 1.

Empirical Strategy

RD Design

Given that LUCCD students are assigned to courses by a system of placement cutoffs, we can use an RD design to identify the effect of placement in a lower-level course relative to a higher-level course. The RD design relies on the assumption that (a) an exogenously determined rule assigns students to different conditions, and (b) observations within a narrow bandwidth above and below the cutoff are statistically equivalent, and thus, assignment to the treatment condition (PA) or control condition (EA) is akin to randomized assignment (Murnane & Willett, 2010). Differences in outcomes between the two groups can be attributed to the policy rule, here, the placement policy.

To enable comparison of outcomes across assessment cohorts, we examine the effects of placement in PA versus EA on enrollment, persistence, and credit completion within a year of the assessment. These outcomes are important success milestones for students in developmental education (Melguizo, 2011). The RD design enables us to identify the difference in the probabilities of achieving these persistence and completion outcomes for students placed in PA, relative to statistically similar students placed directly in EA. A baseline RD model for each college is as follows:

$$Y_{it} = \alpha_0 + g\left(S_{it}\right) + \beta_1 T_{it} + \gamma' X_{it} + \varphi_t + \varepsilon_{it}. \qquad (1)$$

Here, i indexes students and t indexes assessment cohorts. Y is the outcome variable, and S is the normalized and cutoff-centered total placement test score for each student in each campus and assessment cohort, serving as the running variable in the RD design and taking on some functional form g.[5] Because the MDTP scores are integer values and the ACCUPLACER scores in some colleges are noninteger values, it was necessary to normalize the scores to enable comparison across colleges.[6] The variable T is a dichotomous indicator of treatment status, equal to 1 if the student was assigned to PA and 0 if assigned to EA. $\boldsymbol{X}$ is a set of student-level covariates including age, sex, racial/ethnic group, language, and citizenship status, which we include to increase the precision of the RD estimates (Murnane & Willett, 2010). The model includes cohort fixed effects φ_t, and standard errors are clustered by cohort.

The coefficient β_1 gives the effect of placement in the treatment condition (PA) for each college, averaged over each assessment cohort for each college and policy period. Yet as discussed previously, our research question concerns how placement policies themselves can affect student outcomes in developmental math. With these panel data, we can estimate the *change* in the RD effect after placement policy experimentation in each college, a first difference estimate. To do so, we add a dummy variable P that equals 0 if the observation is in the prepolicy period and equals 1 in the postpolicy period in each college. This is interacted with the test score and treatment indicator from Equation 1. The model inclusive of the policy change is as follows:

$$\begin{aligned} Y_{it} = {} & \alpha_0 + g\left(S_{it} \times P_t\right) + \beta_1 T_{it} + \beta_2 P_t + \beta_3 T_{it} \times P_t \\ & + \gamma' X_{it} + \varphi_t + \varepsilon_{it}. \end{aligned} \qquad (2)$$

We estimate Equation 2 for each college to obtain a difference-in-RD effect, captured here by the coefficient β_3. This first difference coefficient can be interpreted as the change in the RD effect after the policy change in each college. A positive estimate would suggest that remediation became more beneficial (less negative) to students, whereas a negative estimate would suggest that its benefit decreased (more negative). A null estimate here would indicate that the effect of remediation remained constant through the policy change.

Three Focus Colleges

The estimated first difference coefficient β_3 from Equation 2 may be biased if there were

TABLE 1

Changes in Placement Policy at PA/EA Cutoff, LUCCD Colleges 2005–2012

College	Policy change	Date	Placement test	% students placed in PA	% students placed in EA	Sample in PA	Sample in EA	Total sample size
A	Switched from MDTP to ACCUPLACER	Spring 2011	MDTP/ACCUPLACER	37.9	19	12,125	1,988	14,113
B	Switched from MDTP to ACCUPLACER	Fall 2009	MDTP/ACCUPLACER	19.8	31	6,206	4,742	10,948
C	Raised cutoff by 7 points	Fall 2009	ACCUPLACER	23.9	18.7	2,379	2,379	4,758
D	No changes[a]	—	COMPASS	48.3	40.5	6,793	4,403	11,196
E	No changes	—	COMPASS	33.8	31.6	4,310	2,667	6,977
F	Instructional change in EA (Extended EA)	Fall 2008	ACCUPLACER	27.7	19.2	5,019	3,641	8,660
G	Switched from ACCUPLACER to MDTP[b]	Fall 2009	ACCUPLACER/MDTP	54.3	23.1	3,763	1,399	5,162
H	Removed extended algebra; Raised PA/EA cutoff by 41 points	Fall 2011	ACCUPLACER	13.3	16.1	2,623	1,377	4,000
I	Raised PA/Extended EA cutoff by 5 points	Spring 2011	ACCUPLACER			1,545	941	2,486
ALL	—	—		28.2	23.7	44,763	23,537	68,300

Note. PA = pre-algebra; EA = elementary algebra; LUCCD = large urban community college district; MDTP = Mathematics Diagnostic Testing Project.

[a]College D instated a self-placement policy in summer and fall, 2008. Because no tests were administered, it is not possible to include students enrolling in those semesters in the analysis. The placement cutoffs for PA/EA prior to summer 2008 and after fall 2008 were identical.

[b]College G could not be included because of the unavailability of accurate diagnostic test data.

other changes in the district that were also related to student achievement outcomes (e.g., enrollment trends, demographic shifts, budget changes related to the recession, district policies, etc.). To account for these secular trends, we embed the RD design within a difference-in-difference framework, using Colleges D and E as control colleges. Colleges D and E did not implement any placement policy changes at the PA/EA cutoff during the 2005 to 2012 study window[7] and therefore provide a baseline trend from which we can identify a difference-in-difference-in-RD (DDRD) estimate for each college that did enact a placement policy change. The model is as follows:

$$Y_{ijt} = \alpha_0 + g\left(S_{ijt}\right) + \left(\beta_1 T_{ijt} + \rho' C_j + \beta_2 T_{ijt} C_j\right) + \left(g\left(S_{ijt}\right) + \beta_3 T_{ijt} + \xi' C_j + \beta_4 T_{ijt} C_j\right) P_j + \gamma' X_{ijt} + \mu_j + \varphi_t + \varepsilon_{ijt}. \tag{3}$$

This is similar to Equation 2 above, except that *j* indexes colleges, and there is now a dummy indicator C_j, which is equal to 1 if the student is in the treatment college and 0 if the student is in one of the control colleges (D or E). The three-way interaction between *C, T*, and *P* is the DDRD estimate, identifying the change in the RD effect after each college implemented a placement policy change. Standard errors are two-way clustered by campus and cohort in the pooled model (Cameron, Gelbach, & Miller, 2011).

Switch in Placement Test. We chose three focus colleges for this analysis. Of the three colleges (A, B, and G) that switched between using diagnostics and computer-adaptive tests, Colleges A and B had diagnostic test data available and made no observable instructional changes. We unfortunately were unable to include College G due to the unavailability of MDTP subscore data, which is necessary for determining the running variable in the RD design. In addition, it is important to note that College A, which is included, uses the MDTP diagnostic information in a unique way. The MDTP provides skill-specific information regarding student math proficiencies, and College A has used this feature prior to 2011 to establish a system of multiple cutoffs to determine math course placement. For example, students who earned a high enough composite score to be placed into EA may still have been placed into PA if they did not earn a high enough score on one of the subtopics of the MDTP (e.g., fractions, exponents). To account for this system of multiple cutoffs, we use a binding-score RD approach in which the minimum of the set of four scores serves as the running variable in the RD design (Reardon & Robinson, 2012). This binding-score approach essentially transforms the multiple dimensions of the assignment process to a one-dimensional process along the lines of a traditional RD design.

Raising Cutoffs. Of the three colleges that raised placement cutoffs (C, H, & I), Colleges H and I were not included as focus colleges because they assigned students to extended algebra, a two-semester EA sequence.[8] College C was included as it raised cutoffs by 7 points and maintained a traditional one-semester EA course.

We present in Table 2 sample means of background characteristics and outcomes of the PA and EA samples before and after the policy changes in each of the three focus colleges.[9] Overall, it is evident that the student populations are quite different across colleges, reflecting LUCCD's diversity. Students assigned to PA are slightly older and are more likely to be female, Latina/o, and African American.

Outcome Measures

Table 2 also provides sample mean outcomes. The first set focuses on early persistence in developmental math because the change in placement test policy provides a unique opportunity to examine whether placement test results discourage students from enrolling. Although this has been studied in recent literature, with results indicating no signs of placement test discouragement (Martorell et al., 2015; Scott-Clayton & Rodriguez, 2015), our study tests this hypothesis in a context where diagnostics and computer-adaptive tests are used. In addition, if students were not placed correctly and therefore not served well by the time spent in remediation, then this may also affect subsequent persistence decisions. We therefore examine whether or not students attempted EA within a year of the placement exam. The second set of outcomes captures completion within a year of assessment—whether students completed EA, the gatekeeper

TABLE 2
Characteristics and Mean Outcomes of Samples in PA and EA, 2005–2012

	College A				College B				College C			
	Prepolicy		Postpolicy		Prepolicy		Postpolicy		Prepolicy		Postpolicy	
	PA	EA	PA	EA	PA	EA	PA	EA	PA	EA	PA	EA
Characteristics												
Age	23.40	25.96	22.77	22.31	26.11	26.03	24.59	21.40	27.42	24.47	27.64	25.06
Female	0.54	0.48	0.56	0.40	0.63	0.55	0.52	0.52	0.67	0.65	0.65	0.58
Asian/PI	0.11	0.26	0.00	0.00	0.10	0.13	0.11	0.11	0.01	0.01	0.03	0.02
African American	0.02	0.01	0.02	0.02	0.11	0.07	0.11	0.09	0.79	0.76	0.72	0.68
Latina/o	0.81	0.62	0.84	0.79	0.45	0.29	0.47	0.47	0.14	0.17	0.18	0.25
White	0.02	0.03	0.01	0.02	0.23	0.39	0.26	0.26	0.01	0.01	0.01	0.01
other	0.05	0.08	0.07	0.07	0.10	0.11	0.06	0.06	0.05	0.04	0.06	0.04
Nonnative English	0.32	0.40	0.22	0.19	0.27	0.22	0.20	0.15	0.07	0.09	0.07	0.08
Permanent resident	0.09	0.13	0.05	0.06	0.13	0.10	0.11	0.08	0.05	0.05	0.05	0.06
Other visa	0.09	0.10	0.06	0.08	0.08	0.07	0.06	0.07	0.02	0.02	0.04	0.05
Outcomes												
Enrolled	0.81	0.88	0.83	0.87	0.88	0.89	0.74	0.83	0.88	0.91	0.78	0.84
Enrolled in math	0.53	0.57	0.44	0.65	0.70	0.71	0.48	0.66	0.62	0.64	0.39	0.62
Compliance	0.94	0.92	0.89	0.97	0.94	0.91	0.93	0.92	0.91	0.91	0.82	0.94
Attempted EA in 1 year	0.23	0.48	0.17	0.60	0.24	0.59	0.20	0.60	0.15	0.53	0.15	0.55
Pass EA in 1 year (A/B)	0.15	0.30	0.10	0.35	0.13	0.41	0.13	0.35	0.07	0.20	0.07	0.31
Total units 1 year	10.21	11.05	10.08	11.53	10.91	13.61	9.57	11.49	8.67	9.07	7.70	10.71
n	9,524	946	2,601	1,042	2,722	1,916	3,484	2,826	1,406	1,861	973	518

Note. Students in compliance are those who enrolled in the math course to which they were assigned. PA = pre-algebra; EA = elementary algebra; PI = Pacific Islander.

course, and total units completed within a year's time. These completion outcomes can help us to understand the impacts of placement decisions on academic achievement.

Because attrition from developmental math sequences affect the observability of subsequent outcomes, the proposed RD analysis may produce biased estimates if we did not input zeroes for unobserved outcomes (e.g., outcomes of students who did not enroll, dropped out, or withdrew). For example, if we included only students who were placed in PA and have observable outcomes in EA, that is, those who chose to continue on and enroll, we may be positively biasing the RD estimates. This is because this approach does not account for the treatment effect on persistence (i.e., enrollment or dropout), and we only observe students who chose to continue on in the math sequence. We could adopt an instrumental variables strategy that accounts for compliance with the treatment assignment, but as Scott-Clayton and Rodriguez (2015) explain, the relevant treatment is not the taking of remedial courses but rather assignment to remediation, which may affect initial enrollment decisions. The total effect of assignment to remediation should include both the effect on persistence and the effect of developmental education coursework. Because we are interested in this total effect of placement decisions, we believe it is appropriate to input zeroes for unobserved outcomes and to estimate the Intent to Treat (ITT) effect. We also reason that in the LUCCD context, students taking the assessments likely have the intent of completing EA, and it is also the intent of the college, in offering these developmental math courses, to prepare students to pass higher-level courses by first assigning them to PA. Inputting zeroes enables us to identify the total effect of assigning students to remediation.

Validity of the RD Design

Manipulation. Before presenting the results, we discuss customary checks of the assumptions underlying the validity of the RD design. The

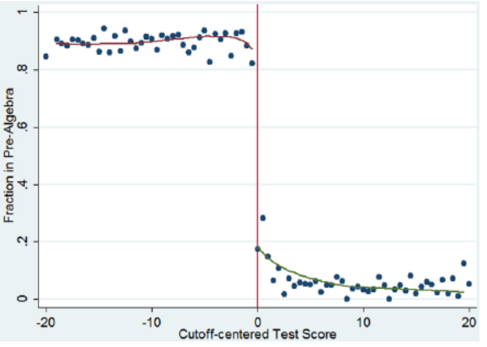

FIGURE 1. *Compliance with pre-algebra assignment for math enrollers.*

first is that there is an exogenously defined policy that determines treatment status (Murnane & Willett, 2010). That is, the policy rule or cutoff assigns participants to either a treatment or control condition by way of a running variable such as a placement test score, and this status cannot be manipulated. If students were able to cheat to attain a score just above the cutoff, then this would pose a threat to the internal validity of the treatment effect estimate. According to our analysis of assessment regulations, students typically receive a printout following the placement test and are at no point made aware of test score cutoffs. They are also typically not allowed to retest before 1 year's time. In addition, the ACCUPLACER and COMPASS are computer-adaptive tests, which would be difficult to manipulate. To inspect manipulation, we plotted histograms and densities of test scores in each college before and after each policy change. There was no indication of any discontinuities at the PA/EA cutoff in any college (see Figure A1 in the online appendix, available at http://epa.sagepub.com/supplemental).

Compliance. Second, there would be a threat to validity if PA students complied with their placement assignments at a different rate compared with students placed directly in EA. We observe that compliance with placement assignment is high, but not perfect. Roughly 92% of students complied with the initial placement based on assessment results (see Table 2 and Figure 1). We also observed that of the 8% who did not comply, 3% enrolled in a lower-level course and 5% enrolled in a higher-level course. Although students are typically blocked by the enrollment management system from registering in a higher-level course than the one in which they were placed, students are allowed to enroll in a lower-level course and by law are allowed to challenge their placement decision. This generally entails a process of the student presenting evidence of his or her math preparedness and obtaining approval from math faculty. Given the small percentage of noncompliers and the fact that similar numbers of students enrolled in higher- and lower-level courses than the one to which they were assigned, we include these students and proceed with what is referred to as "sharp" RD to estimate the ITT effect. This approach, in contrast to "fuzzy" RD, enables us to examine the effect of placement decisions on enrollment behavior. We provide results of fuzzy RD estimation for each policy period using two-stage least squares and assignment to PA as an instrument for compliance in Online Appendix Table A4. As expected with high rates of compliance, the estimates are of similar magnitude as the sharp RD results presented in the "Findings" section below.

Covariate Balance. Finally, a third key assumption in RD design is that no other variables vary discontinuously at the cutoff. The descriptive statistics presented in Table 2 reveal differences in the composition of the students in the PA and EA samples in each college. However, this is not necessarily a threat to the validity of the RD estimates if the covariates are not discontinuous at the cutoff. To check this, we ran a set of parallel RD regressions using Equation 1 with each covariate serving as the dependent variable in the regression (Lee & Lemieux, 2010). The results, which are available in Online Appendix Table A2, suggest that there is evidence of discontinuities in some covariates, such as age and race. However, this may be due to the number of covariates and random chance, so we performed a joint test of these discontinuities in a seemingly unrelated regression setup (Lee & Lemieux, 2010). Most of these are not significant. Although we observe a significant discontinuity in College B in the .33 *SD* bandwidth, this disappears in the narrower bandwidth. This covariate test allows us to be confident that any significant differences in outcomes around the cutoff can be attributed to assignment to the treatment and control groups. With these checks of the internal validity of the RD estimates complete, we proceed with the presentation of the findings.

Findings: The Effects of Placement Policy Experimentation

We use the LUCCD data to focus on 2 types of placement policy experimentation—switching from a diagnostic placement tool to a computer-adaptive tool in Colleges A and B, and raising the cutoff by 7 points in College C—and examine the impact of placement in PA relative to EA near the margin of the placement cutoff in each policy context. As an initial step, we first obtained the baseline RD estimate for each college and policy period using the optimal bandwidth suggested by Imbens and Kalyanaraman (2012).

The results, available in Online Appendix Table A3, indicate that the RD coefficients generally became more negative after each policy change with respect to attempting and completing EA, but not for enrolling or credit completion. This simple before–after comparison gives an idea of the level change in the RD estimate, but it does not account for changes in student characteristics over time.

We therefore present here the results from estimation of Equation 2, a pooled model that enables us to compare the treatment effect before and after the policy change in each college and control for student-level characteristics over time. The *Treat (PA)* coefficient shown in Table 3 provides an estimate of the baseline pre-policy impact of being placed in PA versus EA in each focus college. We present a set of fixed bandwidths ranging from .25 *SD* to 1 *SD* to enable comparison across policy periods and to be able to more easily interpret the estimates over time.

We do not observe any significant negative effects on initial enrollment for students within narrower bandwidths of the PA/EA cutoff in any of the colleges, mirroring other studies that find no evidence of placement exam discouragement (Martorell et al., 2015; Scott-Clayton & Rodriguez, 2015). We do observe significant negative effects of placement in PA on both attempting and passing EA within a year of the assessment, but PA students at the margin of the cutoff appeared to complete about the same total number of units in 1 year's time as those students placed in EA.

The coefficient of interest for the first difference estimate is the *Treat* × *Post* indicator, which captures the change in the RD effect after placement policy experimentation in each college. The estimates in Table 3 indicate that there are mostly significant negative effects on enrollment in the two colleges (A and B) that switched from diagnostics to computer-adaptive tests. The difference in enrollment for students at the margin of the cutoff was about 10 to 13 percentage points lower following the test switch. Rates of attempting EA also decreased at the margin in these 2 colleges by about 25 percentage points. The effects on completing EA in Colleges A and B are negative, but not all significant at the 5% significance level. For example, the treatment effect of placement in PA decreased by 18.4 percentage points ($p < .01$) within the .50 bandwidth and 10.1 percentage points in the .33 bandwidth ($p < .10$). There was also a statistically significant decrease in College B, ranging from 9.1 percentage points in the .33 bandwidth ($p < .10$) and 14.4 percentage points in the .25 bandwidth ($p < .01$). There were no significant changes in the RD effect on total units completed. Interestingly, the RD effect does not appear to change in College C after it raised placement cutoffs by 7 points, and this is consistent for all four outcomes. It may be that raising the cutoff by this amount did not significantly change the composition of academic ability around the cutoff. Although not presented here, we did find a significant decrease in the RD effects after College H raised its PA/EA cutoff by 41 points, suggesting that modestly raising the cutoff may be innocuous, but substantially raising it may be detrimental.[10]

We illustrate these policy changes in Figures 2 to 5. Each figure shows the estimated discontinuity coefficient with 95% confidence intervals for each outcome and assessment cohort. Although there is variation in the RD estimates over time, the trends and levels appear to be different in the pre- and postpolicy periods in Colleges A and B, but not in College C.

Difference-in-Difference-in-RD

Although these first difference estimates provide an idea of the level change in the RD

TABLE 3

Change in Regression Discontinuity Estimates Before and After a Policy Change in Each College (First Difference)

	Enrolling				Attempting EA				Passing EA				Total units 1 year			
	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*
College A																
Treat (PA)	−0.039*	−0.024	−0.015	−0.015	−0.20***	−0.20***	−0.22***	−0.23***	−0.090**	−0.084*	−0.121*	−0.122*	−0.312	−0.363	−1.617	−1.632
	(0.015)	(0.014)	(0.029)	(0.029)	(0.028)	(0.039)	(0.056)	(0.056)	(0.026)	(0.034)	(0.050)	(0.050)	(0.736)	(1.116)	(1.515)	(1.518)
Treat × Post	0.04	−0.02	−0.102*	−0.137*	−0.121	−0.29***	−0.287**	−0.254*	−0.014	−0.184**	−0.101	−0.037	1.424	−2.92	−2.902	−2.764
	(0.048)	(0.020)	(0.044)	(0.065)	(0.060)	(0.056)	(0.078)	(0.102)	(0.045)	(0.059)	(0.060)	(0.075)	(0.892)	(1.572)	(2.163)	(1.718)
Constant	0.970***	1.011***	1.040***	1.038***	0.509***	0.490***	0.565***	0.563***	0.340***	0.317***	0.383***	0.385***	12.43***	12.42***	14.45***	14.56***
	(0.020)	(0.026)	(0.015)	(0.016)	(0.040)	(0.042)	(0.050)	(0.049)	(0.035)	(0.046)	(0.058)	(0.060)	(0.926)	(1.020)	(1.242)	(1.269)
R^2	.040	.046	.037	.038	.087	.091	.114	.116	.046	.037	.039	.040	.037	.037	.035	.037
n	11,387	5,422	2,781	2,697	11,387	5,422	2,781	2,697	11,387	5,422	2,781	2,697	11,387	5,422	2,781	2,697
College B																
Treat (PA)	0.014	0.025	0.05	0.068*	−0.23***	−0.24***	−0.22***	−0.17***	−0.10***	−0.095*	−0.080*	−0.062	0.334	0.817	1.718*	2.035
	(0.015)	(0.017)	(0.025)	(0.030)	(0.035)	(0.039)	(0.039)	(0.042)	(0.025)	(0.034)	(0.032)	(0.033)	(0.577)	(0.892)	(0.765)	(0.997)
Treat × Post	−0.076*	−0.097**	−0.115*	−0.122	−0.110*	−0.107*	−0.19***	−0.240**	0.003	0.012	−0.091	−0.144**	0.574	−0.631	−2.856	−3.575
	(0.029)	(0.034)	(0.046)	(0.064)	(0.039)	(0.050)	(0.047)	(0.071)	(0.043)	(0.078)	(0.051)	(0.050)	(0.841)	(1.573)	(1.376)	(2.342)
Constant	0.933***	0.924***	0.908***	0.887***	0.592***	0.603***	0.565***	0.542***	0.243***	0.164***	0.109***	0.117***	10.97***	7.76***	7.51***	7.59***
	(0.014)	(0.020)	(0.026)	(0.034)	(0.015)	(0.024)	(0.041)	(0.038)	(0.017)	(0.022)	(0.027)	(0.023)	(0.633)	(0.616)	(0.753)	(0.977)
R^2	.031	.031	.032	.041	.131	.111	.100	.111	.086	.067	.060	.074	.042	.041	.053	.061
n	7,165	3,871	2,404	1,827	7,165	3,871	2,404	1,827	7,165	3,871	2,404	1,827	7,165	3,871	2,404	1,827
College C																
Treat (PA)	0.007	−0.011	−0.016	−0.014	−0.33***	−0.33***	−0.32***	−0.31***	−0.08***	−0.09***	−0.084**	−0.094**	0.257	−0.32	−0.646	−0.489
	(0.019)	(0.014)	(0.016)	(0.020)	(0.027)	(0.017)	(0.019)	(0.029)	(0.014)	(0.012)	(0.023)	(0.023)	(0.562)	(0.367)	(0.408)	(0.447)
Treat × Post	−0.038	−0.05	−0.027	−0.043	−0.038	−0.008	−0.06	−0.044	−0.092*	−0.04	−0.057	−0.03	−2.742*	−1.518	−1.561	−1.688
	(0.058)	(0.072)	(0.081)	(0.069)	(0.069)	(0.068)	(0.088)	(0.122)	(0.042)	(0.044)	(0.056)	(0.075)	(1.094)	(0.714)	(0.879)	(1.437)
Constant	0.942***	0.964***	0.957***	0.962***	0.587***	0.595***	0.601***	0.595***	0.180***	0.181***	0.178***	0.169***	10.94***	11.16***	10.97***	11.03***
	(0.034)	(0.032)	(0.035)	(0.035)	(0.027)	(0.024)	(0.028)	(0.032)	(0.034)	(0.029)	(0.041)	(0.033)	(0.889)	(0.846)	(0.822)	(0.759)
R^2	.034	.031	.034	.026	.176	.170	.158	.158	.076	.065	.058	.054	.038	.038	.037	.032
n	4,715	3,524	2,839	2,468	4,715	3,524	2,839	2,468	4,715	3,524	2,839	2,468	4,715	3,524	2,839	2,468

Note. Full model includes test score (*z*), age, race, sex, language, resident status, campus, and cohort. EA = elementary algebra; PA = pre-algebra.
*$p < .05$. **$p < .01$. ***$p < .001$.

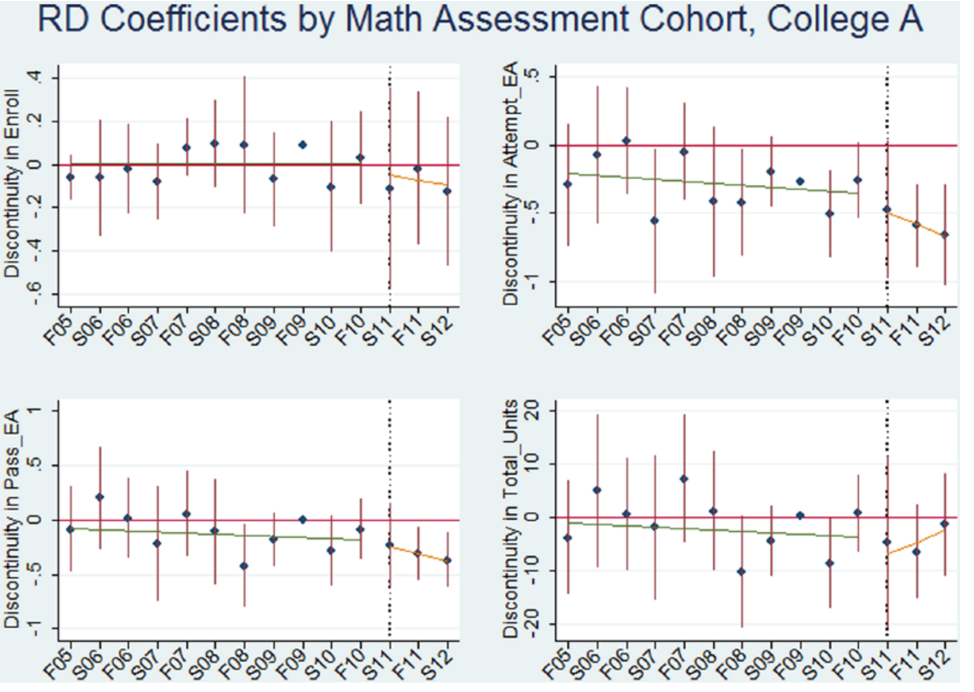

FIGURE 2. *RD coefficients by math assessment cohort, College A.*
Note. RD = regression discontinuity; EA = elementary algebra.

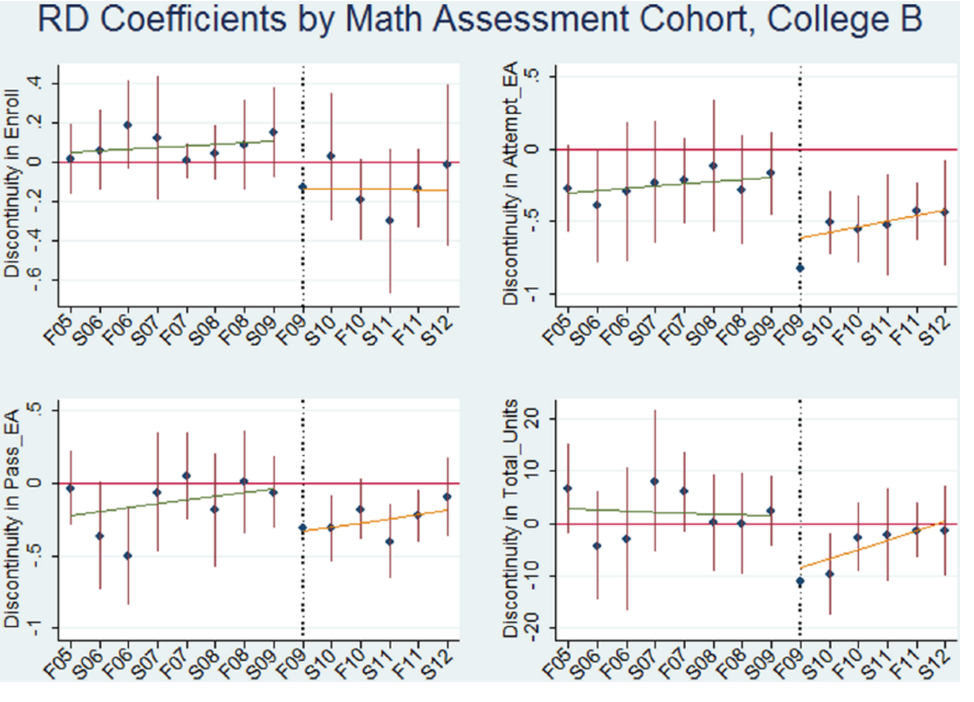

FIGURE 3. *RD coefficients by math assessment cohort, College B.*
Note. RD = regression discontinuity; EA = elementary algebra.

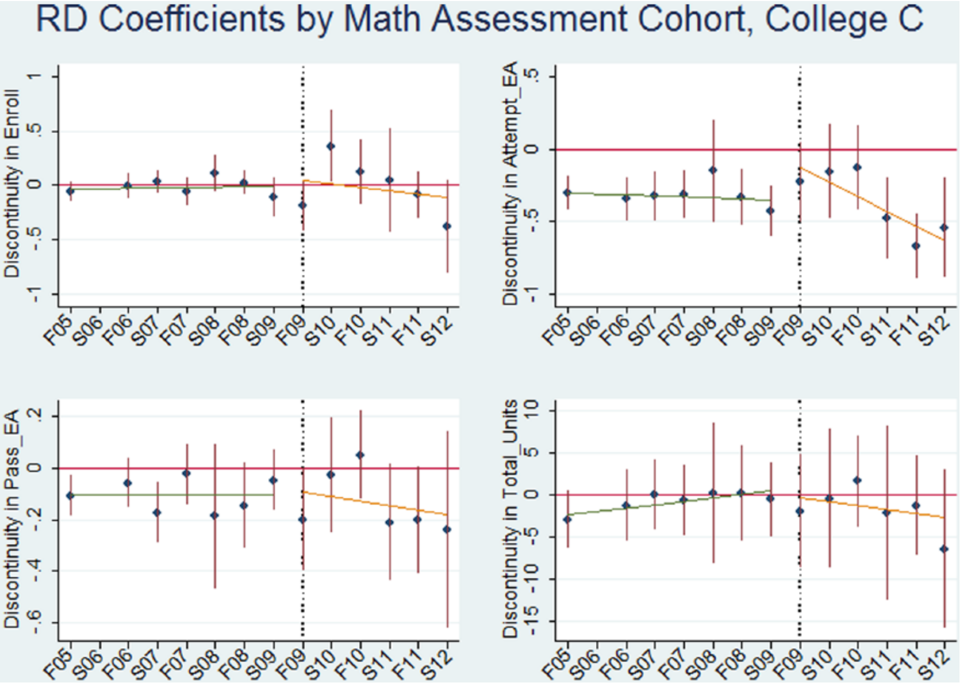

FIGURE 4. *RD coefficients by math assessment cohort, College C.*
Note. RD = regression discontinuity; EA = elementary algebra.

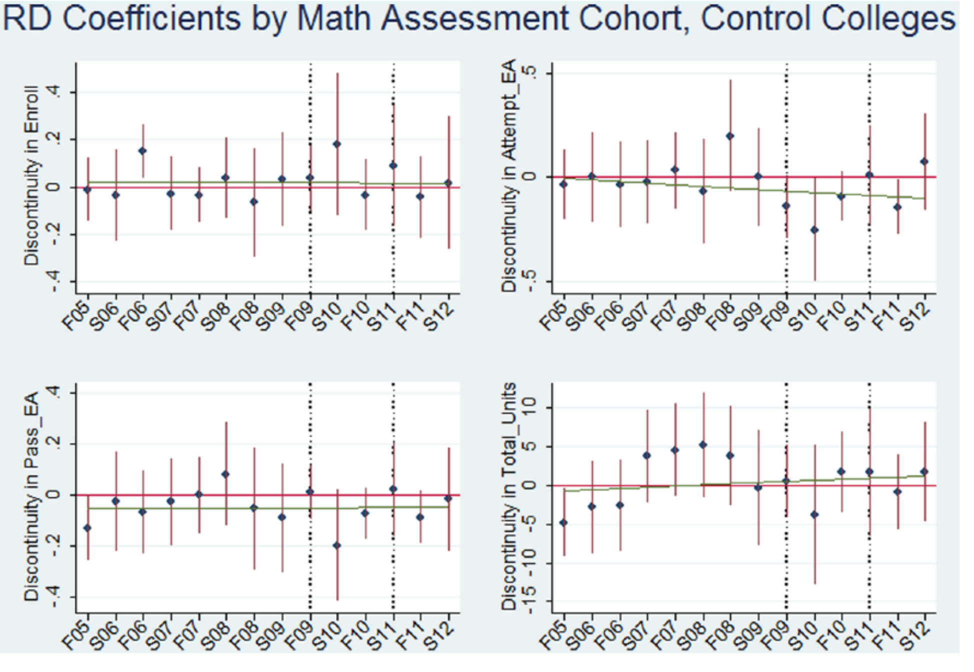

FIGURE 5. *RD coefficients by math assessment cohort, control colleges.*
Note. RD = regression discontinuity; EA = elementary algebra.

effect following each placement policy change, it is plausible that concurrent trends may be correlated with the outcomes of interest and pose a threat to the validity of these results. These include factors such as enrollment trends and demographic shifts, changes in resource allocation, or other policies that could influence success outcomes across the district.

An ideal comparison to isolate the effect of a placement policy change such as College B's switch from the MDTP to ACCUPLACER in 2009 would be to compare College B with an identical counterfactual College B that did not make the switch. Then, the differences in outcomes could be attributed directly to the policy change. In the absence of this ideal counterfactual control college in the district, we use Colleges D and E, both of which did not enact any placement policy changes between 2005 and 2012.[11] Changes between the pre-2009 and post-2009 trends in Colleges D and E provide a second difference that can be used to identify a DDRD effect in College B. We therefore use Equation 3 for each focus college and estimate the DDRD coefficient, which is the change in the RD estimate accounting for secular trends in the district that my influence student outcomes.

The validity of the DDRD estimate in this context hinges on the assumption that the trend in the control colleges does not also significantly change after the policy change in the treatment college (Murnane & Willett, 2010). We see in Figure 5 that the trends across colleges are close to parallel prior to the policy change, and that there does not appear to be any drastic change in the RD estimates following fall, 2009, when Colleges B and C made a change, or spring, 2011, when College A made a change. These are also captured in estimation by the $T \times P$ coefficients presented in Tables 4 and 5. We observe that these are not significant at most bandwidth sizes and not significant at narrow bandwidths, thus providing convincing evidence that the underlying trend in the control colleges is constant through the policy change in the treatment colleges. That there are no changes in the control colleges provides important evidence assuring the internal validity of the difference-in-difference estimate.

The main effect of interest is the *Treat* × *Post* × *College A/B/C* variable, which provides an estimate of the DDRD effect in each of the three focus colleges. We present these for the set of persistence outcomes in Table 4 and the set of completion outcomes in Table 5.

Enrollment. We find that relative to diagnostics, there does appear to be a discouragement effect for the results of computer-adaptive tests. Students assigned to PA under the computer-adaptive test policy in Colleges A and B were about 12 to 15 percentage points less likely to enroll than their counterparts in the same colleges who were assigned under the MDTP. This reduction in cohort size was further compounded as the group placed in PA faced a second persistence decision. We see that an additional 10% to 15% of students did not enroll in EA at the margin of the cutoff, for a total of about a 25% difference in EA attempt rates between students placed in PA versus EA.

Completion. There are similar findings for the completion outcomes, shown in Table 5. In Colleges A and B, who made the switch from MDTP to ACCUPLACER, we see no change in the larger bandwidths, but statistically significant and consistently negative effects on passing EA following the placement policy change within narrow bandwidths. This ranges from a 9 to 17 percentage point reduction in the RD effect in College A, and a 10 to 19 percentage point reduction in College B in the narrower RD bandwidths. Similar to the first difference model, the estimate for the .25 *SD* bandwidth in College A is negative but not statistically significant. We suspect that this may be due to the small sample of students in the postpolicy period in the .25 *SD* bandwidth, because the sample sizes in the .33 *SD* and .25 *SD* bandwidths are the same in the prepolicy period in College A. These lower completion rates also seem to have translated into a reduction in credit completion for students placed in PA in these 2 colleges, with students in PA completing about 2.4 to 3.5 fewer credits than students placed in EA.

Overall, these estimates suggest that students around the placement cutoff after the switch in placement tests in Colleges A and B experienced a larger penalty to placement in PA relative to EA. In other words, students placed using results from computer-adaptive tests were more negatively

TABLE 4

Difference-in-Difference-in-Regression Discontinuity Estimates, Enrollment Outcomes

	Enrolling				Attempting EA within 1 year			
	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*
College A: Switch								
Treat × Post × College A	0.048	−0.025	−0.126*	−0.130***	−0.113***	−0.273***	−0.270***	−0.235**
	(0.042)	(0.045)	(0.059)	(0.035)	(0.029)	(0.059)	(0.068)	(0.077)
Treat × Post	−0.02	0.005	0.023	−0.012	−0.018	−0.014	−0.008	−0.007
	(0.018)	(0.029)	(0.040)	(0.017)	(0.038)	(0.069)	(0.076)	(0.091)
Treat × College A	−0.071***	−0.066*	−0.047	−0.051	−0.145**	−0.161***	−0.188***	−0.194***
	(0.020)	(0.029)	(0.035)	(0.044)	(0.051)	(0.048)	(0.056)	(0.058)
Post × College A	0.050*	0.074***	0.153***	0.127***	0.087**	0.124**	0.118	0.123
	(0.024)	(0.017)	(0.037)	(0.038)	(0.032)	(0.045)	(0.064)	(0.070)
Constant	0.922***	0.918***	0.907***	0.910***	0.228***	0.209***	0.270***	0.280***
	(0.008)	(0.024)	(0.043)	(0.045)	(0.033)	(0.038)	(0.053)	(0.057)
R^2	.033	.037	.034	.037	.118	.117	.14	.142
N	25,014	12,379	7,265	5,969	25,014	12,379	7,265	5,969
College B: Switch								
Treat × Post × College B	−0.072***	−0.100***	−0.120***	−0.151***	−0.051	−0.028	−0.142*	−0.231***
	(0.014)	(0.011)	(0.021)	(0.017)	(0.052)	(0.048)	(0.060)	(0.052)
Treat × Post	−0.002	0.007	0.005	0.016	−0.063	−0.075	−0.044	−0.015
	(0.014)	(0.015)	(0.027)	(0.028)	(0.047)	(0.040)	(0.054)	(0.046)
Treat × College B	−0.021	−0.014	0.01	0.044	−0.207**	−0.234***	−0.192**	−0.125*
	(0.011)	(0.021)	(0.040)	(0.041)	(0.067)	(0.052)	(0.064)	(0.053)
Post × College B	0.021	0.041	0.041	0.044	0.071*	0.044	0.098	0.143**
	(0.022)	(0.021)	(0.023)	(0.045)	(0.031)	(0.034)	(0.053)	(0.049)
Constant	0.905***	0.888***	0.849***	0.856***	0.118*	0.104**	0.159**	0.164***
	(0.011)	(0.016)	(0.019)	(0.021)	(0.052)	(0.039)	(0.050)	(0.044)

(continued)

TABLE 4 (CONTINUED)

	Enrolling				Attempting EA within 1 year			
	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*
R^2	.033	.033	.034	.039	.225	.219	.199	.202
N	20,792	10,828	6,888	5,099	20,792	10,828	6,888	5,099
College C: Raise								
Treat × Post × College C	−0.032*	−0.054***	−0.035	−0.069	0.017	0.066	−0.02	−0.045
	(0.014)	(0.016)	(0.031)	(0.042)	(0.054)	(0.055)	(0.061)	(0.044)
Treat × Post	−0.002	0.008	0.013	0.024	−0.062	−0.075	−0.042	−0.007
	(0.016)	(0.013)	(0.021)	(0.023)	(0.047)	(0.041)	(0.054)	(0.037)
Treat × College C	−0.023	−0.047*	−0.049	−0.03	−0.313***	−0.321***	−0.291***	−0.268***
	(0.021)	(0.022)	(0.035)	(0.039)	(0.066)	(0.053)	(0.062)	(0.051)
Post × College C	0.024	0.042	0.024	0.062	0.033	−0.006	0.044	0.079***
	(0.026)	(0.023)	(0.025)	(0.039)	(0.031)	(0.034)	(0.035)	(0.021)
Constant	0.910***	0.887***	0.837***	0.849***	0.122**	0.110***	0.165***	0.182***
	(0.020)	(0.023)	(0.020)	(0.026)	(0.038)	(0.028)	(0.039)	(0.025)
R^2	.034	.034	.034	.034	.174	.155	.152	.151
N	18,681	10,690	7,470	5,861	18,681	10,690	7,470	5,861

Note. Full model includes age, race, gender, language, residence status, cohort, and college. EA = elementary algebra.
$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

TABLE 5

Difference-in-Difference-in-Regression Discontinuity Estimates, Completion Outcomes

	Passing EA within 1 year (B or better)				Total units completed within 1 year			
	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*
College A: Switch								
Treat × Post × College A	−0.001	−0.173***	−0.082**	−0.025	1.067	−3.026***	−2.394***	−0.743
	(0.013)	(0.020)	(0.026)	(0.033)	(0.658)	(0.599)	(0.598)	(0.690)
Treat × Post	−0.021	−0.01	−0.008	−0.003	−0.158	−0.104	−0.64	−2.013
	(0.013)	(0.029)	(0.033)	(0.046)	(0.472)	(0.676)	(0.816)	(1.040)
Treat × College A	−0.056***	−0.062***	−0.102***	−0.108***	−0.277	−0.871	−2.560***	−2.316**
	(0.016)	(0.013)	(0.021)	(0.022)	(0.635)	(0.601)	(0.769)	(0.832)
Post × College A	−0.009	0.032	−0.013	−0.054	−0.073	2.004***	1.783**	0.226
	(0.015)	(0.018)	(0.029)	(0.035)	(0.648)	(0.516)	(0.546)	(0.867)
Constant	0.138***	0.125***	0.152***	0.175***	10.44***	10.60***	11.09***	11.88***
	(0.014)	(0.012)	(0.032)	(0.030)	(0.362)	(0.265)	(0.309)	(0.394)
R^2	.07	.069	.084	.084	.037	.032	.033	.032
n	25,014	12,379	7,265	5,969	25,014	12,379	7,265	5,969
College B: Switch								
Treat × Post × College B	0.036	0.045	−0.099**	−0.185***	0.206	0.14	−2.514**	−3.531**
	(0.021)	(0.025)	(0.031)	(0.030)	(0.928)	(0.789)	(0.768)	(1.178)
Treat × Post	−0.029	−0.036**	−0.002	0.015	0.389	−0.92	−0.943*	−1.213*
	(0.020)	(0.012)	(0.030)	(0.036)	(0.792)	(0.472)	(0.462)	(0.573)
Treat × College B	−0.076**	−0.075***	−0.035	−0.011	0.507	0.331	1.121	2.063*
	(0.026)	(0.015)	(0.027)	(0.024)	(0.650)	(0.786)	(1.016)	(0.939)
Post × College B	0.033***	0.001	0.046	0.083*	−0.278	−0.312	0.314	0.156
	(0.008)	(0.008)	(0.034)	(0.038)	(0.469)	(0.651)	(0.749)	(0.992)
Constant	0.052*	0.012	0.014	0.024	10.63***	10.14***	10.25***	10.93***
	(0.021)	(0.036)	(0.042)	(0.050)	(0.535)	(1.483)	(1.815)	(2.047)

(continued)

TABLE 5 (CONTINUED)

	Passing EA within 1 year (B or better)				Total units completed within 1 year			
	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*	1 *SD*	.50 *SD*	.33 *SD*	.25 *SD*
R^2	.143	.12	.101	.101	.047	.039	.044	.044
n	20,792	10,828	6,888	5,099	20,792	10,828	6,888	5,099
College C: Raise								
Treat × Post × College C	−0.065**	−0.006	−0.058	−0.06	−2.989***	−0.572	−0.647	−0.555
	(0.023)	(0.033)	(0.041)	(0.045)	(0.769)	(0.744)	(0.442)	(0.466)
Treat × Post	−0.027	−0.035*	0.001	0.023	0.529	−0.875	−0.865*	−1.214*
	(0.019)	(0.016)	(0.023)	(0.019)	(0.791)	(0.481)	(0.365)	(0.499)
Treat × College C	−0.057*	−0.081***	−0.047	−0.048**	0.648	−0.872*	−1.480**	−0.932
	(0.027)	(0.015)	(0.024)	(0.015)	(0.350)	(0.378)	(0.522)	(0.711)
Post × College C	0.109***	0.057**	0.069*	0.077**	1.892**	−0.103	−0.487	−0.434
	(0.014)	(0.022)	(0.027)	(0.030)	(0.717)	(0.742)	(0.728)	(0.587)
Constant	0.082***	0.055***	0.053***	0.070***	11.10***	11.29***	11.14***	11.98***
	(0.013)	(0.011)	(0.013)	(0.011)	(0.489)	(0.672)	(1.234)	(1.339)
R^2	.075	.054	.051	.05	.043	.046	.05	.048
n	18,681	10,690	7,470	5,861	18,342	10,481	7,323	5,740

Note. Full model includes age, race, gender, language, residence status, cohort, and college. EA = elementary algebra.
*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 6

Sensitivity Analyses for BW = .33 SD, DDRD Estimates

	(1) Full	(2) No covariates	(3) Quadratic	(4) Cubic	(5) Smaller analytical window[a]	(6) Placebo policy change (1 sem. before)	(7) Conditioning on enrollment
Enrolling							
T × P × College A	−0.126*	−0.115*	−0.129*	−0.123*	−0.128	−0.073	—
	(0.059)	(0.056)	(0.056)	(0.059)	(0.085)	(0.059)	
T × P × College B	−0.120***	−0.112***	−0.135***	−0.108***	−0.100*	−0.089	—
	(0.021)	(0.016)	(0.021)	(0.026)	(0.045)	(0.044)	
T × P × College C	−0.037	−0.028	−0.039	−0.031	0.019	−0.089	—
	(0.033)	(0.028)	(0.035)	(0.033)	(0.099)	(0.062)	
Attempting EA within 1 year							
T × P × College A	−0.270***	−0.245***	−0.267***	−0.254**	−0.217	−0.183	−0.271
	(0.068)	(0.069)	(0.072)	(0.082)	(0.098)	(0.090)	(0.277)
T × P × College B	−0.142*	−0.140*	−0.161*	−0.154	−0.123	−0.049	−0.133
	(0.060)	(0.062)	(0.071)	(0.089)	(0.088)	(0.054)	(0.097)
T × P × College C	−0.015	−0.014	−0.013	0.007	0.050	−0.046	−0.166
	(0.064)	(0.066)	(0.063)	(0.066)	(0.108)	(0.074)	(0.109)
Passing EA within 1 year							
T × P × College A	−0.082**	−0.072*	−0.083*	−0.058	−0.083	−0.041	−0.125
	(0.026)	(0.030)	(0.033)	(0.039)	(0.075)	(0.064)	(0.081)
T × P × College B	−0.099**	−0.102**	−0.122**	−0.118*	−0.124	−0.045	−0.092
	(0.031)	(0.033)	(0.044)	(0.055)	(0.062)	(0.062)	(0.049)
T × P × College C	−0.058	−0.057	−0.043	−0.027	−0.033	−0.005	−0.185
	(0.041)	(0.047)	(0.047)	(0.047)	(0.091)	(0.066)	(0.129)
Total units in 1 year							
T × P × College A	−2.394***	−1.337*	−2.322***	−2.098*	−2.496	0.161	−1.284
	(0.598)	(0.635)	(0.571)	(0.875)	(2.943)	(2.256)	(0.815)
T × P × College B	−2.514**	−2.466***	−3.110***	−2.410***	−0.846	−1.286	−0.839
	(0.768)	(0.634)	(0.662)	(0.699)	(1.486)	(1.986)	(0.730)
T × P × College C	−0.647	−0.425	−0.532	−0.198	0.039	0.330	−0.117
	(0.442)	(0.425)	(0.508)	(0.425)	(1.452)	(1.208)	(0.535)

Note. Full model is .33 *SD*, with covariates, campus and cohort fixed effects, and linear interaction term of treatment and assignment variables. DDRD = difference-in-difference-in-RD; EA = elementary algebra; RD = regression discontinuity.

[a]The analytical window is 2 years before and after the policy change in Colleges B and C, and 4 semesters before and after the policy change in College A because we did not have data for the full 2 years after spring, 2011.

*$p < .05$. **$p < .01$. ***$p < .001$.

affected by the placement decision than prior cohorts placed by MDTP.[12]

Although the overall completion rates dropped in College C, we see that there were no significant changes in any outcomes in the narrower bandwidths around the PA/EA cutoff after College C raised the cutoff. This suggests that although higher-scoring students in College C were placed in PA instead of EA after cutoffs were raised, this did not significantly exacerbate the penalty of remediation for those students. We suspect that this may be related to the noisiness of the test and the possibility that raising the cutoff did not improve placement accuracy. We investigate this hypothesis further in the next section after testing the robustness of the results.

Robustness, Sensitivity, and Extensions

Each robustness check presented in Table 6 focuses on the .33 *SD* bandwidth, and the main results are presented in column 1. In column 2, we present the model estimated without covariates and see that the coefficients are of roughly the same magnitude and significance. In accordance with standard practice for RD analysis, we examined sensitivity to functional form by testing the inclusion of test score polynomials up to the third degree. We found that including these polynomial interactions with the treatment variable did not produce markedly different results from the linear form. This is demonstrated in columns 3 and 4.

We also performed checks to test the validity of the DDRD model to determine whether the observed treatment effects are merely artifacts of time. First, we narrowed the analytical window to just 2 years before and after the policy change.[13] The results presented in column 5 indicate that these estimates are similar to those of the full model. There would also be a threat to validity if the change in the treatment effect we observed was not unique to the timing of the policy change. Although we found that the postpolicy control college coefficients were not significantly different from zero (see Tables 4 and 5), we also performed a falsification exercise with a placebo policy change. We ran the DDRD model one semester before the actual policy change in each campus and provide the results in column 6. Again, we do not find any significant placebo policy cutoff coefficients. This provides further assurance that the difference estimates we obtained are specific to the policy change in each college.

Finally, in column 7, we present estimates of the average treatment on the treated (ATT) effect, for which no zeroes were inputted for unobserved outcomes. This estimate therefore reflects differences in outcomes of students placed in PA relative to EA conditional on enrollment. These results can help us understand whether what is driving the negative findings is nonenrollment or the quality of developmental education courses, because a possible explanation for the more negative impact of placement in PA is that educational quality decreased. Conditioning on enrollment suggests that students who chose to enroll and persist performed no differently from their counterparts placed in EA. This analysis underscores the impact of the placement results on enrollment behavior.

Test or Cutoff Effects?

Although the significant DDRD effects were only observed in the two colleges that switched from diagnostics to computer-adaptive placement instruments, we nevertheless caution against interpreting the estimates of the change as solely attributable to the use of the computer-adaptive ACCUPLACER and the MDTP diagnostic. We examined the effects of a placement *policy* change, which encompasses more than just a change in test type. The MDTP, for example, was typically offered in a paper and pencil format during the study window, and students, generally under the advisement of the testing coordinator, were able to choose which MDTP subtest to start with. In contrast, the ACCUPLACER is a computer-adaptive test with predetermined starting points. It is also possible that the relative cutoffs after the policy changes in Colleges A and B were higher with respect to incoming student ability, and that the estimates we obtained in the diagnostic colleges may be more attributable to a high cutoff than to the use of the computer-adaptive test itself.

Reducing Placement Errors

To investigate this further, we present supplementary analyses focusing on placement accuracy, which can help us to better understand the role of placement tests in explaining developmental math outcomes. A possible logic for the observed decline in the RD estimates is that students were more accurately placed when diagnostics were used and less accurately placed when computer-adaptive instruments were used. Drawing on a method to examine placement accuracy described in Scott-Clayton et al. (2014), we first obtain predicted probabilities of passing and failing EA by analyzing only students who were placed directly in and who enrolled in EA. The probit model includes test scores, multiple measure indicators based on student's academic background, which are used in the placement process (*MM*), and a set of demographic controls, and is given by the following two equations:

$$\Pr[Fail\ EA = 1]_{ijt} = \alpha_0 + \beta_1 (TestScore)_{ijt} + \beta_2 MM_{ijt} + \gamma' X_{ijt} + \varepsilon_{ijt},$$

$$\Pr[Pass\ EA\ w/B\ or\ better = 1]_{ijt} = \alpha_0 + \beta_1 (TestScore)_{ijt} + \beta_2 MM_{ijt} + \gamma' X_{ijt} + \varepsilon_{ijt}.$$

The estimated parameters are used to extrapolate predicted probabilities of passing and failing EA for the full sample (i.e., including those students placed in PA). Students who were either predicted to fail EA but were placed there, or get a B or better in EA but are not placed there, are considered severely misplaced. The average proportion of

TABLE 7
SER of Placement, Before and After a Policy Change

Campus	Policy change	Prepolicy change		Postpolicy change		Difference	
		All	Narrow	All	Narrow	All	Narrow
	Switched placement test						
A	MDTP to ACCU in spring, 2011	0.062	0.076	0.108	0.443	+0.046	+0.367
B	MDTP to ACCU in fall, 2009	0.027	0.045	0.111	0.390	+0.084	+0.345
	Raised PA/EA cutoff						
C	By 7 points (ACCU) in fall, 2009	0.274	0.326	0.118	0.313	−0.156	−0.013

Note. SER is calculated as the sum of severe misplacements. This is the average of the proportion of students who are predicted to pass EA with a B or better but are placed in PA and the proportion of students who are predicted to fail/withdraw from EA but are placed there. The narrow bandwidth is 5 points above and below the cutoff, and the narrow cutoff in College A is 1 point above and below the cutoff. SER = severe error rate; MDTP = Mathematics Diagnostic Testing Project; PA = pre-algebra; EA = elementary algebra.

these two groups provides the severe error rate (SER; Scott-Clayton et al., 2014).

The findings from this analysis, presented in Table 7, indicate that the rate of severe placement errors increased by about 5 percentage points in College A and 9% in College B. After the placement instrument change, about 11% of students in PA and EA were severely misplaced. That is, these students were likely to have passed EA with a B or better, but were placed in PA, or were likely to fail EA but placed there anyways. The rates are considerably higher within a narrower bandwidth of the cutoff, jumping to 44% in College A and 39% in College B, both increases of about 35%. This suggests that the content of the computer-adaptive placement test may be misaligned with the algebra courses (Hodara, Jaggars, & Karp, 2012). In College C, which raised the PA/EA cutoff, the estimated SERs actually decreased in the full sample of students and stayed roughly the same within a narrow bandwidth. This helps to explain why we did not observe any differences in the RD estimate in College C. It appears that raising the cutoff by 7 points did not improve placement accuracy.

These errors rates seem to complement the RD results. An increase in the SER of placement around the cutoff may explain the larger negative impact of remediation. If students were less accurately placed by the computer-adaptive test than by the diagnostic, this may have exacerbated the effects of placement in PA relative to EA. That College C did not experience a decline in placement accuracy following an increase in the placement cutoff suggests that the increase in Colleges A and B may be related more to the use of the computer-adaptive test than to the change in cutoff relative to student ability. Further research would be needed to investigate this hypothesis in more depth.

Discussion and Policy Implications

Because community colleges are a pathway toward college certificates and degrees for a large fraction of the population, many of whom are first-time college-goers, it is crucial to increase access and ensure success through more accurate assessment and placement practices. Indeed, research shows that the cost associated with misplacement, in terms of the opportunity cost to the student and the cost of providing remediation, is estimated to be much larger than the cost of the test itself (Rodríguez, Bowden, Belfield, & Scott-Clayton, 2014). If colleges pay more attention to placement test policy such that accuracy is improved and misplacement errors are minimized, then colleges and students can both stand to benefit. Statewide reforms with this goal are already underway: The Texas Success Initiative, for example, has the goal of incorporating diagnostic assessment and holistic advising as part of its statewide developmental education plan; Virginia, North Carolina, and Florida likewise have revised assessment practices to improve placement decisions (Asera, 2011; Burdman, 2012; Rodríguez, 2014). California is moving toward centralization by piloting a common assessment system, although it remains yet to be seen what types of tools and measures will be utilized (Burdman, 2015).

The findings of the study demonstrate that although there may not be a perfect placement policy or tools devoid of error, there are those that can make marginal improvements over current practices and minimize the damage of placement errors. In the absence of large-scale evaluations of these policies and tools, we capitalized on the local-level heterogeneous policy context in LUCCD to examine placement policy experimentation in the form of test choice and cutoff changes. The findings highlight a possible advantage to using diagnostics to make placement decisions in developmental math. All else constant, the two community colleges that switched from diagnostics to computer-adaptive tests experienced a larger negative impact of remediation, with fewer students at the margin of the cutoff enrolling and moving on to EA after being placed in PA. Although we cannot completely disentangle the effects of the test switch from a change in the level of the cutoff with respect to student ability, our supplementary analyses show that there were higher proportions of severe placement errors following the switch from diagnostics to computer-adaptive tests. This provides some insights into a potential explanation for the observed decline in the effectiveness of remediation—that diagnostics may be better able to identify student skill than computer-adaptive tests. Diagnostics such as the MDTP can provide information on student proficiency for a range of subtopics such as fractions, exponents, and reasoning; this skill-specific information can be incorporated into placement policies to improve math placement decisions or used to tailor instruction in math courses.

Of course, making placement policy improvements may be easier said than done. The results of the study also underscore the reality that assessment and placement policy decisions are often made without much guidance or technical support, and faculty feel unsupported in choosing placement tools, and setting and evaluating cutoffs (Melguizo et al., 2014). Colleges may raise cutoffs simply because they observe low passing rates in higher-level courses, or to manage enrollment levels in college-level courses (Melguizo et al., 2014; Hodara et al., 2012). There are also competing pressures that may make certain policy choices more viable than others. For example, computer-adaptive tests, although potentially less accurate, may be more cost-effective and easier to administer in large batches, and they also can take less time than a diagnostic test such as the MDTP. Clearly, these tensions between equity and efficiency necessitate that colleges make trade-offs that may come at the cost of student success (Jaggars & Hodara, 2013).

Irrespective of the placement instruments chosen, one way to support practitioners in evaluating current placement policies is through RD analyses similar to the ones presented here. This stems from Robinson's (2011) notion of using RD estimates to identify whether students experienced a "smooth transition" between academic contexts (e.g., from English language learner [ELL] courses to English-dominant courses). Drawing on this idea for developmental math, if cutoffs are "correct," then we should observe no significant discontinuity in later outcomes. If students experienced a positive benefit to placement in a lower-level course, evidenced by a positive discontinuity, this would suggest that the cutoff could be raised so that more students can obtain the benefit of remediation. However, if the RD estimate is negative, as we observe here during each college's prepolicy period, then this implies a negative impact of remediation and that the cutoff could be lowered to minimize the penalty (Melguizo et al., 2013). Further research should continue to investigate this hypothesis and interpretation with respect to other contexts of placement policy experimentation, such as when placement cutoffs are lowered rather than raised.

Conclusion

Although these findings are drawn from a sample of large urban community colleges in California and have limited external generalizability, growing concern about the usefulness of commercially available placement tests obliges research on alternative placement policies and tools that can guide practitioner efforts. Overall, our analysis shows that switching from using a math diagnostic for placement to using a computer-adaptive test led to a larger negative effect of assignment to remediation on early college outcomes, including initial math enrollment and completion of gatekeeper math courses. It appears that these differences may have stemmed from an increase in placement errors. We did not observe significant changes in the colleges where placement cutoffs were raised, although the findings indicate that the appropriate experimentation would have been to lower the test score cutoffs. It seems evident that community colleges should experiment with using diagnostic information to more systematically inform placement decisions or lower cutoffs. This may improve placement accuracy and the early college outcomes of incoming community college students placed in developmental math.

Acknowledgments

We wish to thank the officials at the community college system for providing access to the data, and Holly Kosiewicz, Kristen E. Fong, and W. Edward Chi at the University of Southern California, and William Doyle at Vanderbilt University for comments on prior drafts of the article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A100381 to the University of Southern California. Additional support was received from an internal grant from the Advancing Scholarship in the Humanities and Social Sciences (ASHSS) Initiative of the University of Southern California, Office of the Provost.

Notes

1. Some colleges reported using more than one instrument, so these do not add up to 100%.

2. See Burdman (2012) for an overview of reforms across the country.

3. Unfortunately, we are unable to include College G due to the unavailability of diagnostic data.

4. College F made a change in instructional delivery by adding a cutoff that assigned students to either pre-algebra (PA) or two semesters of elementary algebra (EA). Fong, Melguizo, and Prather (2015) provide an analysis of this practice.

5. We normalized the test score by campus and cohort to enable comparison across colleges. We also ran models with raw test scores and obtained similar results.

6. Note that the total score used to derive the course placement consists of both the raw test score and additional points factored in from multiple measures. These three colleges, like the other large urban community college district (LUCCD) colleges, awarded between −2 and 5 additional points based on student responses to an educational background questionnaire administered at the time of the placement exam. Depending on the institution, students could earn additional points for the highest level of math taken, the importance of math, their high school grade point average (GPA), whether they have a diploma or general education development (GED), and their college plans. Importantly, the areas for which additional points were awarded and the amount of possible points did not change in any of the three colleges.

7. College D implemented a self-placement policy during the 2008 summer and fall semesters. Because no assessments were given, there are no assessment records for these two terms, and they are not included in the sample.

8. See Fong et al. (2015) for an examination of the impact of assigning students to extended algebra.

9. Means for a narrower bandwidth can be found in Online Appendix Table A1.

10. In College H, students were placed into PA versus extended EA, a two-semester EA sequence. These results are available from the authors on request.

11. We also compared the difference-in-difference-in-RD (DDRD) results when just College E was used as a control college and found results of similar direction and magnitude as those presented in Tables 4 and 5.

12. We provide figures in Online Appendix Figures A2 to A4 that illustrate this visually for each college.

13. We restricted the check to 1 year before and after in College A because 2 years after were not possible with the data.

References

Armstrong, W. B. (2000). The association among student success in courses, placement test scores, student background data, and instructor grading practices. *Community College Journal of Research and Practice*, *24*, 681–695.

Asera, R. (2011). *Innovation at scale: How Virginia community colleges are collaborating to improve developmental education and increase student success*. Boston, MA: Jobs for the Future.

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*, 255–270.

Betts, J. R., Hahn, Y., & Zau, A. C. (2011). *Does diagnostic math testing improve student learning?* San Francisco: Public Policy Institute of California. Retrieved from http://www.ppic.org/content/pubs/report/R_1011JBR.pdf

Burdman, P. (2012). *Where to begin? The evolving role of placement exams for students starting college*. Boston, MA: Jobs for the Future.

Burdman, P. (2015). *Degrees of freedom: Probing math placement policies at California colleges and universities* (No. 3). Stanford: Policy Analysis for California Education.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, *29*(2), 238–249.

Clark, B. (1960). *The open door college*. New York, NY: McGraw-Hill.

Cohen, A. M., & Brawer, F. B. (2003). *The American community college*. San Francisco, CA: John Wiley.

Deil-Amen, R., & Tevis, T. L. (2010). Circumscribed agency: The relevance of standardized college entrance exams for low SES high school students. *The Review of Higher Education*, *33*, 141–175.

Fain, P. (2015, June 18). Finding a new compass. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/news/2015/06/18/act-drops-popular-compass-placement-test-acknowledging-its-predictive-limits

Fields, R., & Parsad, B. (2012). *Tests and cut scores used for student placement in postsecondary education: Fall 2011*. Washington, DC: National Assessment Governing Board.

Fong, K., Melguizo, T., & Prather, G. (2015). Increasing success rates in developmental math: The complementary role of individual and institutional characteristics. *Research in Higher Education*. Advance online publication. doi:10.1007/s11162-015-9368-9

Foundation for California Community Colleges. (2015). *California community colleges*. Retrieved from https://www.foundationccc.org/Portals/0/Documents/NewsRoom/FactSheets/ccc-facts-figures.pdf

Fulton, M. (2012). *Using state policies to ensure effective assessment and placement in remedial education*. Denver, CO: Education Commission of the States.

Hodara, M., Jaggars, S. S., & Karp, M. M. (2012). *Improving developmental education assessment and placement: Lessons from community colleges across the country* (CCRC Working Paper No. 51). New York, NY: Community College Research Center, Columbia University.

Huang, C.-W., Snipes, J., & Finkelstein, N. (2014). *Using assessment data to guide math course placement of California middle school students* (REL 2014–040). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from http://ies.ed.gov/ncee/edlabs

Hughes, K. L., & Scott-Clayton, J. (2011). Assessing developmental assessment in community colleges. *Community College Review*, *39*, 327–351.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, *79*, 933–959.

Jaggars, S. S., & Hodara, M. (2013). The opposing forces that shape developmental education. *Community College Journal of Research and Practice*, *37*, 575–579.

Jenkins, D., Jaggars, S. S., & Roksa, J. (2009). *Promoting gatekeeper course success among community college students needing remediation: Findings and recommendations from a Virginia Study*. New York, NY: Community College Research Center. Retrieved from http://eric.ed.gov/?id=ED507824

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*, 281–355.

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

Martorell, P., McFarlin, I., Jr., & Xue, Y. (2015). Does failing a placement exam discourage underprepared students from enrolling college? *Education Finance and Policy*, *10*, 46–80.

Marwick, J. D. (2004). Charting a path to success: The association between institutional placement policies and the academic success of Latino students. *Community College Journal of Research and Practice*, *28*, 263–280.

Mattern, K. D., & Packman, S. (2009). *Predictive validity of ACCUPLACER scores for course placement: A meta-analysis* (Research Report No. 2009-2). New York, NY: College Board.

Medhanie, A. G., Dupuis, D. N., LeBeau, B., Harwell, M. R., & Post, T. R. (2012). The role of the ACCUPLACER mathematics placement test on a student's first college mathematics course. *Educational and Psychological Measurement*, *72*, 332–351.

Melguizo, T. (2011). A review of the theories developed to describe the process of college persistence and attainment. In J. C. Smart & M. B. Paulsen (Eds.) *Higher education: Handbook of theory and research* (pp. 395–424). Dordrecht, The Netherlands: Springer.

Melguizo, T., Bos, H., Ngo, F., Mills, N., & Prather, G. (in press). Using a regression discontinuity design to estimate the impact of placement decisions in developmental math. *Research in Higher Education*.

Melguizo, T., Hagedorn, L. S., & Cypers, S. (2008). Remedial/developmental education and the cost of community college transfer: A Los Angeles County sample. *The Review of Higher Education*, *31*, 401–431.

Melguizo, T., Kosiewicz, H., Prather, G., & Bos, H. (2014). How are community college students assessed and placed in developmental math? Grounding our understanding in reality. *Journal of Higher Education*, *85*, 691–722.

Murnane, R. J., & Willett, J. B. (2010). *Methods matter: Improving causal inference in educational and social science research*. Oxford, UK: Oxford University Press.

National Center for Public Policy and Higher Education & Southern Regional Education Board. (2010). *Beyond the rhetoric: Improving college readiness through coherent state policy*. Atlanta, GA: National Center for Public Policy and Higher Education. Retrieved from http://publications.sreb.org/2010/Beyond%20the%20Rhetoric.pdf

Ngo, F., & Kwon, W. (2015). Using multiple measures to make math placement decisions: Implications for access and success in community colleges. *Research in Higher Education*, *56*, 442–470.

Papay, J. P., Murnane, R. J., & Willett, J. B. (2011). *How performance information affects human-capital investment decisions: The impact of test-score labels on educational outcomes* (No. w17120). Cambridge, MA: National Bureau of Economic Research.

Perry, M., Bahr, P. M., Rosin, M., & Woodward, K. M. (2010). *Course-taking patterns, policies, and practices in developmental education in the California Community Colleges*. Mountain View, CA: EdSource.

Reardon, S. F., & Robinson, J. P. (2012). Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness*, *5*, 83–104.

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis*, *33*, 267–292.

Rodríguez, O. (2014). *Increasing access to college-level math: Early outcomes using the Virginia Placement Test*. Retrieved from http://ccrc.tc.columbia.edu/publications/increasing-access-to-college-level-math.html

Rodríguez, O., Bowden, B., Belfield, C., & Scott-Clayton, J. (2014, September). *Remedial placement testing in community colleges: What resources are required, and what does it cost?* (CCRC Working Paper No. 73). New York, NY: Community College Research Center.

Scott-Clayton, J., Crosta, P., & Belfield, C. (2014). Improving the targeting of treatment: Evidence from college remediation. *Educational Evaluation and Policy Analysis*, *36*, 371–393.

Scott-Clayton, J., & Rodriguez, O. (2015). Development, discouragement, or diversion? New evidence on the effects of college remediation. *Education Finance and Policy*, *10*, 4–45.

Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator*, *1*(3), 4–16.

Authors

FEDERICK NGO is a PhD candidate in urban education policy at the Rossier School of Education, University of Southern California. His research interests include community colleges, college access and persistence, and math education.

TATIANA MELGUIZO is associate professor at the Rossier School of Education, University of Southern California. Her research interests include economics of education, community colleges, and developmental education.

Manuscript received March 11, 2015
Revision received June 23, 2015
Accepted August 6, 2015