Research in Higher Education
https://doi.org/10.1007/s11162-021-09630-2

Heterogeneity in the Effects of College Course Placement

Lindsay Daugherty[1] · **Russell Gerber**[2] · **Francisco Martorell**[3] · **Trey Miller**[4] · **Emily Weisburst**[5]

Received: 21 October 2019 / Accepted: 15 February 2021

Abstract

Traditionally colleges have relied on standalone non-credit-bearing developmental education (DE) to support students academically and ensure readiness for college-level courses. As emerging evidence has raised concerns about the effectiveness of DE courses, colleges and states have been experimenting with approaches that place students into credit-bearing coursework more quickly. To better understand which types of students might be most likely to benefit from being placed into college-level math coursework, this study examines heterogeneity in the causal effects of placement into college-level courses using a regression discontinuity design and administrative data from the state of Texas. We focus on student characteristics that are related to academic preparation or might signal a student's likelihood of success or need for additional support and might therefore be factors considered for placement into college-level courses under "holistic advising" or "multiple measures" initiatives. We find heterogeneity in outcomes for many of the measures we examined. Students who declared an academic major designation, had bachelor's degree aspirations, tested below college readiness on multiple subjects, were designated as Limited English Proficiency (LEP), and/or were economically disadvantaged status were more likely to benefit from placement into college-level math. Part-time enrollment or being over the age of 21 were associated with reduced benefits from placement into college-level math. We do not find any heterogeneity in outcomes for our high school achievement measure, three or more years of math taken in high school.

Keywords Postsecondary education · College readiness · Developmental education · College advising · Equity

✉ Lindsay Daugherty
ldaugher@rand.org

[1] RAND Corporation, 1776 Main St, Santa Monica, CA 90401, USA

[2] Texas Higher Education Coordinating Board, 1200 E Anderson Ln, Austin, TX 78752, USA

[3] School of Education, University of California, Davis, 1 Shields Ave, Davis, CA 95616, USA

[4] School of Economic, Political and Policy Sciences, University of Texas at Dallas, 800 West Campbell Road, Richardson, TX 75080, USA

[5] Luskin School of Public Affairs, University of California, Los Angeles, 337 Charles E Young Dr E, Los Angeles, CA 90095, USA

Published online: 26 March 2021

Introduction

Over the past decade, higher education policymakers and college administrators have recognized the need to improve developmental education (DE). DE courses are designed to help students become "college-ready," but do not provide credit towards an academic degree and are typically required before a student can enroll in credit-bearing courses. While more than 60 percent of community college students enroll in DE (Bailey et al. 2010), researchers have found little evidence these courses improve academic outcomes (Boatman & Long, 2018; Calcagno & Long, 2008; Martorell & McFarlin, 2011; Scott-Clayton & Rodriguez, 2012) and have raised concerns that many students "get stuck" in lower-level DE courses (Jenkins & Bailey, 2017).

To address these concerns, states and colleges are changing policies and practices around DE delivery and placement to move students into college-level courses more quickly or avoid DE courses altogether. Some colleges are looking to "holistic advising" and "multiple measures" approaches to make placement decisions rather than relying solely on placement exam scores. These approaches consider a wide range of measures including indicators of high school achievement and course-taking, college enrollment characteristics (e.g. full time vs part time), life circumstances, and other indicators of readiness (Cullinan et al. 2018). However, it remains unclear which student characteristics meaningfully distinguish between students who benefit from placement into college-level courses and who might be harmed. Furthermore, policymakers need to know which group identified by a potential measure is likely to benefit from placement into college-level courses.

This study examines heterogeneity in the impact of placement into college-level math. We focus on student characteristics that might be considered for placement decisions under holistic advising and multiple measures placement approaches. We used a sample of first-time in college (FTIC) enrollees entering all Texas public community colleges between 2013 and 2015. We employed a regression discontinuity (RD) design to estimate the impact of placement into college-level math coursework on passing a college-level math course and persisting in community college. By estimating effects separately for student subgroups and testing for differential impacts across subgroups, we examined which measures are associated with more positive outcomes from placement into college-level math. We examined a wide range of measures, including a measure of high school achievement, characteristics of enrollment (e.g., part-time/full-time enrollment, major and degree type), and other student-level characteristics that might be considered in making placement decisions (e.g., English proficiency status, economic barriers). We made adjustments for multiple comparisons to guard against finding spurious effects.

We find that many of the factors we considered were associated with differential effects of placement into college-level coursework. Students pursuing an academic major and students pursuing a bachelor's degree were more likely to benefit from college-level course placement, while part-time students were less likely to benefit. Interestingly, we found that several student subgroups that might be viewed as facing additional barriers to success in college-level courses—students with Limited English Proficiency (LEP) status, students designated as economically disadvantaged, and students testing below college readiness in more than one subject—were actually more successful when placed into college-level math coursework. And finally, students who were older than 21 years of age were less likely to be successful when placed directly into college-level math coursework. We did not find differential effects for our high school achievement measure, taking three or more years of high school math courses. We interpret these findings as being instructive for the development of placement policies and advising practices in settings where practitioners have to take numerous factors into account when making placement decisions and advising students about what courses are most appropriate for them.

Background

Policy Background

Texas and other states have passed major policies requiring reform to the delivery of DE and placement into DE coursework, with a particular focus on reforms that place students more immediately into college-level courses. A national survey finds that as of 2016, more than half of all community colleges across the country were experimenting with accelerated models of developmental education (Rutschow et al. 2019). Some state policies encourage placement directly into the credit-bearing college course without any mandated additional support, such as Florida's elimination of mandated DE (Hu et al. 2016). Other states like California, Tennessee, and Texas have passed policies that call for colleges to pair college-level coursework with additional academic support, a strategy referred to as "corequisites" (Cuellar Mejia et al. 2016; Daugherty et al. 2018; Ran & Lin, 2019). Recent studies suggest that the placement of students directly into college courses can, on average, improve rates of success in early math, reading and writing coursework (Cho et al. 2012; Logue et al. 2019; Miller et al. 2020; Park et al. 2016; Ran & Lin, 2019). What is less clear is whether these benefits apply broadly or whether some groups of students benefit while others do not.

To determine whether students will be placed into college-level coursework or DE, most states and colleges across the country continue to rely heavily on standardized placement exams (Rutschow et al. 2019). However, over the past decade, colleges have increasingly moved to adopt "multiple measures" (or in Texas, "holistic advising") reforms, meaning that they supplement or replace placement exam scores with other measures to determine the optimal placement for students. As of 2016, 57 percent of community colleges reported using multiple measures to place students in math (Rutschow et al. 2019), and the reforms have continued to be scaled across the country since that time. Policies that use more information to make placement decisions can reduce the rate of misplacement into developmental education by providing additional information on academic readiness and other factors that may predict success in college-level coursework (Scott-Clayton et al. 2014). This is especially true given recent research suggesting that estimates of college readiness might vary by student characteristics (Klasik & Strayhorn, 2018). Moreover, policies that aim to generally expand placement into college-level courses (through accelerated models of instruction) that have demonstrated positive benefits (Logue et al. 2019; Miller et al. 2020; Park et al. 2016; Ran & Lin, 2019).

A key question for policies that increase discretion in placement decisions relative to strict test-based placement is what factors might be appropriate for determining placement. Academic measures like high school grades and courses are most commonly incorporated into multiple measures systems and thus critical to examine. Most of the research literature on alternative measures for placement focuses on high school grade-point average (GPA), with evidence consistently demonstrating that GPA can help to improve placement (Bahr et al. 2017; Hodara & Cox, 2016; Ngo et al. 2013; Scott-Clayton, 2012; Scott-Clayton et al.

2014). Ngo and Kwan (2015) find that the highest math course taken and grade earned can also be a valuable predictor.

However, recent multiple measures implementation research (e.g., Barnett and Reddy, 2017; Cullinan et al. 2018) suggests that colleges might drawing on a broader set of non-academic measures to determine placement under more flexible holistic advising policies. For example, a national survey found that 13 percent of colleges were considering measures of motivation and commitment for placement (Rutschow et al., 2019). Texas policy guidance on holistic advising recommended transportation challenges and other financial barriers as possible factors that might be considered in determining placement under the assumption that students facing these challenges might require additional support. Our technical assistance and implementation work in Texas community colleges (e.g., Daugherty et al. 2018; Gehlhaus et al. 2018) suggested that factors such as enrollment intensity, major, age, and English Language support needs were also being considered by some advisors as predictors of success in college-level coursework and being used to guide placement recommendations. Yet there is little evidence around how these non-academic student characteristics are related to success in college-level coursework, and the use of some of these factors for placement purposes raise concerns about equity. More research is needed to understand whether there is heterogeneity around student outcomes for these characteristics.

It is worth noting that other studies have also examined heterogeneity in the impacts of developmental education (or acceleration) by student characteristics like race/ethnicity and gender, though we assume that these features are not being used for the purposes of placement and thus exclude them from our analysis (e.g., Boatman & Long, 2018; Hu et al. 2016).

DE Policy in Texas

In 2013 the state began to require that all public colleges in Texas use a common assessment for placement, the Texas Success Initiative Assessment (TSIA). Under the TSIA, the state established a single set of statewide cutoff scores in math, reading, and writing that all public Texas colleges were required to use for placement. Students scoring below the common college-ready cut score were required by state policy to enroll in DE, while students scoring above the cut score could take college-level math courses without having to enroll in DE. The cutoff scores used during our study period form the basis for our regression discontinuity research design, as explained below.

State policy in Texas offered some students waivers and exemptions from placement testing, allowing them to be placed directly into college coursework without taking the placement exam. For example, students who enrolled in short-term technical programs were exempted from TSI requirements. Waivers from placement testing are awarded to students for demonstrating college readiness through other assessments (e.g., SAT, high school exit exam) or prior course-taking at another institution, and waivers are also provided to military and veteran students. Students who did not receive waivers or exemptions and took the TSIA prior to enrollment are the students who were being advised into DE or college-level coursework, the population relevant to our study.

The college-level math course that most college students took during our study period was a college algebra course that could be offered for three or four credit hours. Some colleges were beginning to experiment with "math pathways" reforms, which allowed students to substitute another entry-level math course like statistics or contemporary math for the college algebra course depending on the policies of the college and a student's major. Common learning objectives and allowable course hours were set for courses at the state level, while colleges had the flexibility to determine other course features like curricula, grading policies, and class sizes. The most common DE course taken by students during the study period was intermediate algebra, though some colleges were starting to develop alternative math pathways at the DE level as well.[1]

Following the trends in other states, Texas has made a number of changes to its DE policies in a series of reforms between 2011 and 2015. These reforms are commonly referred to as "Texas Success Initiative" (TSI) reforms. The guidance required colleges to begin using holistic advising for placement into accelerated coursework in 2015, though colleges were provided with discretion regarding which measures would be used, how they would be combined, and the process for incorporating measures into advising. The state's guidance on holistic advising highlighted high school grade point average and class rank, prior academic coursework and/or workplace experiences, noncognitive factors (e.g., motivation, self-efficacy), and family-life issues (e.g., job, childcare, transportation, finances) as factors that should be considered. While the state's holistic advising policy had not yet been implemented during our study period, some of the variables we use in the impact heterogeneity analysis are motivated by these policy changes.

Data

Sample and Data Sources

Our study examined FTIC students at public Texas community colleges. There are 50 community college systems in Texas that encompass 77 community colleges and over a hundred separate campuses. Each year, roughly one million students enroll in at least one course at these public community colleges, of which two hundred thousand are FTIC enrollees. We focused on students first enrolling in the fall or spring semesters between fall of 2013 and the fall of 2015.

We used administrative data from the Texas Higher Education Coordinating Board (THECB) to identify students enrolled in public Texas postsecondary institutions. These files include extensive demographic information, information on enrollment characteristics, and course enrollment and grades. These data were used for our outcomes and most of the placement measure for which we tested for heterogeneity across outcomes. We describe the outcomes and these variables in more detail in the following sections.

Our sample was limited to students who took the TSIA placement exam, which is the relevant population given that these are the individuals who were being considered for placement into DE by their colleges.

We supplemented the state's postsecondary administrative data with information on placement exam scores from the College Board. Since students could retake the placement tests, the College Board data includes multiple score records. We used the first observed math assessment score from College Board files in our analysis and describe the analytic reasons for this decision in greater detail below.

[1] Toward the end of the study period, some colleges in Texas also began to pilot corequisites, where students who tested below college-ready were able to directly enter the college course and receive concurrent DE support rather than taking a standalone course.

To provide colleges with additional information that might be used for the purposes of placement, THECB developed a questionnaire that was appended to the TSIA and requested information on a range of different factors, including educational history—high school diploma or equivalent, years since enrolled in high school and courses, number of high school math and English courses taken—parental educational attainment, race/ethnicity, and indicators of English language learner status. We drew our measure on math high school course-taking from this survey. Response options ranged from 1 course to more than 4, and we created a binary indicator of whether a student had completed at least 3 math courses to represent a threshold that might be used for placement purposes.

Outcomes

We examined two outcomes of interest (see Table 1 for sample means), passing a first college-level class in math, and continued enrollment and/or completion in subsequent semesters (defined in the paper as "persistence"). Passing a first college-level course, or "gateway" course is required to gain entry into other courses and is a necessary degree requirement for all associate degree-seeking students. However, in practice many students do not take such a course because they drop out before completing this program requirement (Bailey et al. 2010). Passing a college-level math course is therefore driven both by being able to enroll in the course and then going on to perform well in the course. Studies indicate that momentum in passing gateway courses is important to later college success (Belfield et al. 2019; Jenkins & Bailey, 2017). These studies also highlight persistence through the initial few semesters of college to be an important predictor of completion.

We examined outcomes through the first two fall/spring semesters of enrollment to allow time for those students initially placed into DE (those under the college-ready cut score) to have completed DE courses in the first semester, then enrolled in and passed their first college-level courses in math. We choose two semesters to ensure a sufficient sample size while also allowing sufficient time to observe outcomes for students on both sides of the cut score.

We coded each of our outcomes as a binary variable, so coefficients can be interpreted as the change in the probability of a student achieving the milestone. Once a student had passed a first college-level course, we considered them as having achieved this milestone for every subsequent semester. For persistence, our outcome accounts only for the enrollment in the semester of interest (the third semester after enrollment), for example, a student who enrolls in fall 2013, does not enroll in spring 2014, but does enroll in fall 2014 is marked as achieving the persistence enrollment milestone. Students who transferred to another Texas college or completed a degree or certificate were also counted as having persisted.[2]

Moderators

We consider level of enrollment (i.e., part-time/full-time), academic or technical major type, bachelor's degree or other aspirations, tested below college ready in subjects other

[2] We observe transfers to other Texas colleges (including private colleges and four-year universities), but we do not observe transfers to colleges outside of Texas, so these transfers are not included in our persistence measure.

Table 1 Summary statistics

	Total		Below Math Cut Score		Above Math Cut Score	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Measures Examined for Heterogeneity						
3+ Years Math in High School	0.659		0.648		0.729	
Academic Major	0.689		0.675		0.780	
B.A. Intent	0.201		0.178		0.351	
Part-time Enrollment	0.611		0.617		0.573	
Tested Below College Ready in 2+ Subjects	0.768		0.813		0.476	
Limited English Proficiency	0.105		0.096		0.161	
Economic Disadvantage	0.504		0.524		0.374	
Age >21	0.399		0.415		0.293	
Other Baseline Characteristics						
Asian	0.032		0.021		0.105	
Black	0.212		0.227		0.110	
Hispanic	0.407		0.423		0.301	
White	0.287		0.273		0.377	
Female	0.561		0.577		0.456	
International	0.022		0.015		0.066	
Certificate Intent	0.063		0.068		0.035	
Technical Major	0.241		0.252		0.168	
Technical Prep Major	0.070		0.072		0.052	
Parent has Some College	0.729		0.718		0.798	
3+ Years English in High School[1]	0.762		0.755		0.807	
Test Scores						
Centered Math TSIA Score	-16.689	15.393	− 20.600	12.106	9.123	8.145
Centered Reading TSIA Score	− 3.489	14.555	− 4.999	14.064	6.475	13.791
Centered Writing TSIA Score	0.446	15.016	− 1.199	14.499	11.303	13.797
Outcomes						
Pass FCL Math (1 year)	0.186		0.128		0.566	
Persistence (1 year)	0.509		0.483		0.682	
N	61208		53155		8053	

[1]Variables Years English in High School > 3, Years Math in High School > 3, and Parent has Some College are obtained via a survey given during TSIA administration and was obtained via the College Board

than math, LEP status, economically disadvantaged status, and age over 21. We created binary indicators for each of our measures of interest. Table 1 shows the sample average levels of each of these variables overall, and for students above and below the statewide cutoff score for college readiness in math.

Table 2 lists the rationale for using each of the moderators along with an explanation for how these variables were created. Broadly speaking, our analysis is motivated by recent research suggesting that measures of college readiness vary by student characteristics (Klasik & Strayhorn, 2018), which suggest that it is important to understand how

Table 2 Measures examined for evidence of heterogeneity

Measures	Description and Justifications for Examining
3 or more math courses taken in high school	Binary measure indicating whether a student took 3 or more math courses in high school. High school achievement measures were highlighted in TSI guidance as being one of the measures Texas colleges should consider for placement. High school achievement measures were the most commonly used across the country (Rutschow et al., 2019), and our on-the-ground work suggested that high school factors were commonly considered for multiple measures placement
Academic major, bachelor degree aspirations	Binary measures for (1) whether a student declared an "academic" major (i.e., from programs intended to transfer credits to a 4-year institution as opposed to "technical majors" which provide occupation-specific skills); binary measure for whether or not a student intends to earn a B.A. (both measured at in the student's first semester). Course of study was reported as commonly being used for placement across the county (Rutschow et al., 2019), and our on-the-ground work with Texas community colleges suggested that these factors were commonly considered in advising
Part-time enrollment	Binary measure for whether a student enrolled in fewer than 12 units in her or his first semester. On-the-ground work with Texas community colleges suggested that enrollment levels were sometimes considered for placement decisions, with part-time enrollees less likely to be offered accelerated options due to concerns that students didn't have the time to devote to college-level coursework
Testing below college ready in multiple areas	Binary measure for scoring below college-level on at least two subject placement tests. On-the-ground work with Texas community colleges suggested that students testing below college ready in multiple areas were sometimes advised out of college-level coursework because of concerns about stronger needs for preparation
Limited English Proficiency status	Binary measure indicating a student received LEP services enroll in LEP courses or who were determined, based on a local placement test, to be Limited English Proficient during his or her first semester. On-the-ground work with Texas colleges indicated that some colleges offered separate DE pathways for students needing English language support and LEP status was sometimes reported as a risk factor for success in college-level coursework
Economic disadvantaged designation	Binary measure indicating a student classified by THECB as economically disadvantaged (triggered by any of the following: annual income at or below the federal poverty line; receipt of public assistance; receipt of a Pell Grant or comparable state program of need-based financial assistance; participation in federal job training). TSI policy recommended that transportation issues and other financial issues should be considered for placement, and economically disadvantaged status can act as a proxy for these factors

Table 2 (continued)

Measures	Description and Justifications for Examining
Age over 21	Binary indicator for whether a student was 21 years of age or more at the start of their first semester of enrollment. TSI policy recommended that work experience and life responsibilities should be considered, and older students might be more likely to have both of these things. Our on-the-ground work with Texas colleges suggested that older, returning students were sometimes placed differently due to concerns about breaks in education and exposure to math coursework

placement decisions could have differential effects by student academic background and other characteristics. Some of the measures we consider are directly related to academic preparation for college-level coursework such as placement test scores and high school course-taking. However, we also consider a broader set of measures that were being used on the ground by advisers to guide placement in some Texas community colleges.

It is important to note that the measures we used may not have been the exact measures used in holistic advising and may instead be proxies. For example, we did not have information on the transportation challenges and other financial barriers that Texas policy guidance recommended might be considered for placement decisions and instead rely on the state's indicator for economic disadvantage as a proxy to identify students most likely to be facing these challenges. To identify students who might be most likely to face life responsibilities (e.g., work, family), we used age over 21 as a proxy. We were also limited by data availability. For example, we were unable to examine some THECB-recommend factors like high school GPA and non-cognitive measures because we did not have access to these data.

Methodology

Research Design and Empirical Specification

Our research design centers on the comparability of students who score very close to the DE placement threshold. The assumption underlying the analysis is that these students are similar aside from their likelihood of being placed into DE classes. We use the variation in being placed in DE at the cut score to isolate the causal impact of enrolling directly in a college-level math course, which can only occur if a student is not placed in DE. To learn about whether some students might benefit from enrolling directly in a college-level math course more than others, we obtain separate estimates of this impact for different student subgroups, as well as estimating the differences in impacts across students.

A key practical issue for implementing this approach is that students can retake the placement exam to score high enough to place out of the DE sequence (there is no additional reward for scoring high beyond passing the cut score). We used the first test score as the "running variable" in the RD analysis to avoid bias related to students strategically retaking the exam. Students whose highest score is just below the placement threshold are those who either did not retake the test or stopped taking it before they could pass it. Conversely, students whose highest score is just above the placement threshold are those that

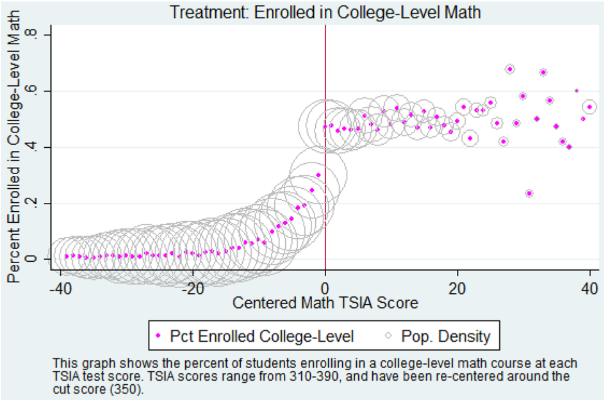


Fig. 1 Math Placement Score and Treatment Variable (Enrolled in College-Level Math)

either passed on a first attempt or kept on retaking the exam until they passed it. Thus, students whose highest test scores are just above or below the placement threshold are likely to systematically differ in motivation and persistence to be placed out of DE. In contrast, scoring barely above or below the threshold on the first attempt should not be related to differences in motivation (or any other characteristic).[3]

Using the first score means that there is a "fuzzy" relationship between whether a student is placed out of DE and the test score running variable we use in this analysis, or students are not perfectly assigned to courses based on their test score. Another source of fuzziness stems from students sometimes being assigned to the college-level course even if they do not score above the college-readiness cutoff.[4] Nonetheless, the likelihood of being in a DE course decreases sharply at the cut score, as can be seen in the plot of placement test scores in Fig. 1. We use fuzzy regression discontinuity (FRD) methods (Hahn, Todd and van der Klaauw, 2001) to exploit this variation and estimate the impact of placement into college-level math versus DE math courses. The FRD approach entails using an indicator for scoring above the threshold as an instrumental variable (IV) for college-level course taking. Specifically, we estimate linear IV models of the following form:

$$\mathrm{T_i} = \delta + \theta\, \mathrm{g}(\mathrm{S_i}) + \eta\, \mathrm{R_i} + \psi\, \mathrm{X_i} + \mathrm{u_i} \tag{1}$$

$$\mathrm{Y_i} = \alpha + \varphi\, \mathrm{T_i} + \gamma\, \mathrm{f}(\mathrm{S_i}) + \beta\, \mathrm{X_i} + \mathrm{e_i} \tag{2}$$

where T_i is a dichotomous variable indicating student i enrolled directly in a college-level math course in the first semester (or was placed out of DE courses), Y_i is student persistence or whether a student passed an first college-level course, $g(S_i)$ and $f(S_i)$ are flexible functions of the TSIA score (S_i), R_i is an indicator for scoring above the cutoff score, X_i is a set of baseline student covariates, and u_i and e_i are residuals. Covariates in the model

[3] This approach has been used by other papers on this topic including Martorell and McFarlin (2011); Calcagno and Long (2008); Boatman and Long (2018).

[4] For instance, this can be because of an exemption or waiver obtained after the student initially tested, in which case the student appears below the cut score in our data but was able to enroll in a college course, or the placement of a student into a corequisite (immediate enrollment in a college course with concurrent DE support).

included the measures we are testing for heterogeneity—three or more years of high school math, academic program designation and plans to seek a degree, part-time enrollment, lack of college readiness in more than one subject, economically disadvantaged status, LEP status, and whether a student was older than 21—as well as other relevant student covariates that are less likely to be used directly for math placement such as race/ethnicity, gender, parental education, placement exam scores in reading and writing, and years of English coursework. Equation (1) denotes the "first stage" relationship between TSIA performance and placement into a DE math course rather than a college-level math course. The parameter η measures the discontinuity in the probability of placement directly in a college-level math course at the cutoff score seen in Fig. 1. Equation (2) is the "structural equation," and relates taking a college-level math course to the outcome Y_i, with the parameter φ capturing the effect of DE course taking on Y_i.

To estimate differential effects for student subgroups, we use a modified version of the model above and estimate the following system of equations:

$$T_i = \delta_1 + \theta_1 g(S_i) + \eta_1 R_i + \phi_1 R_i * W_i + \psi_1 X_i + u_{1i} \tag{1a}$$

$$T_i * W_i = \delta_2 + \theta_2 g(S_i) + \eta_2 R_i + \phi_2 R_i * W_i + \psi_2 X_i + u_{2i} \tag{1b}$$

$$Y_i = \alpha + \varphi T_i + \varphi_{int} T_i * W_i + \gamma f(S_i) + \beta X_i + e_i \tag{2a}$$

where W_i is an indicator for a particular student characteristic (note W_i is a subset of the covariate vector X_i so the "main effect" of W_i is not explicitly written in these equations). To see how this model works, consider the case when W_i is an indicator for being a part-time student. In this example, the effect of enrolling directly in a college-level course for part-time students is given by $\varphi + \varphi_{int}$, and the effect for full-time students is given by φ. Thus, the parameter φ_{int} gives the differential effect of enrolling in a college math course for part-time and full-time students, and more generally for students with $W_i = 1$ and $W_i = 0$.

Several other comments about this model bear mention. First, since we want to estimate two different treatment effects (one for students with $W_i = 1$ and one for students with $W_i = 0$), we need a second instrumental variable, which is created by interacting R_i and W_i; that is why there are two first-stage equations. Second, the "main effect" of the characteristic W_i is explicitly emphasized in the structural Eq. (2). This notation distinguishes it from its usual place in covariate vector X_i. Third, we choose to estimate the models with interactions rather than splitting the sample by W_i to help with statistical power. Fourth, we estimate the model via two-stage least squares regression; this generates the same point estimates as would using the predicted values of T_i and T_i*W_i from the first stage equations in Eq. 2.

To estimate these models, we need to choose a method for estimating the relationship between Y and S (and also between T and S) away from the cutoff, as well as choosing a bandwidth around the cutoff. Our preferred specification does not impose a bandwidth limit and instead uses a flexible polynomial to approximate the functions f(S) and g(S). Specifically, we parameterize these functions with a cubic in S where the parameters of the polynomial are allowed to differ on either side of the cutoff. We use this approach to help increase statistical precision that would be lost if we were to use narrower bandwidths. While this introduces the possibility that observations far from the cutoff might influence the results, the graphical evidence in Fig. 1 suggests this is unlikely to be a significant concern. A fully interacted cubic polynomial should provide enough flexibility to prevent observations far from the cutoff from unduly influencing the estimates.

Because we are running many comparison tests on the same data set, we are likely to find some statistically significant results by random chance. To ensure that our results are robust to the large number of statistical tests, we perform a Benjamini–Hochberg correction for multiple comparisons.[5] We only show significance for cases in which the p-value estimate is sufficiently small to pass the more stringent standards imposed by the Benjamini–Hochberg procedure.

Threats to Internal and External Validity

There are two key assumptions required for the fuzzy RD model described above to identify the effect of enrolling directly in a college-level math course. First, there must be a discontinuity in the likelihood of the treatment (i.e., that η is not equal to zero). Figure 1 clearly demonstrates that to be the case, and below we show estimates of η that corroborate the visual impression. Second, scoring above or below the cutoff must be "as good as random" for students scoring close to the cutoff. In practice, this condition means students scoring just above and below the cutoff do not differ systematically in ways that are related to the outcomes. In the context of this study, the mathematics test is multiple choice and machine scored, making it very unlikely for there to be systematic sorting of students around the cutoff for any particular administration of the TSIA exam. As discussed above, we address potential biases related to test retakes by using the initial TSIA score in the College Board data.

The assumption of no systematic differences between students on either side of the cut score has two testable implications. One is that the density of the running variable ought to be smooth through the cutoff (McCrary, 2008). The second is that there should be no discontinuities in baseline covariates at the cutoff. Intuitively, this test amounts to examining whether the students on either side of the cutoff "look alike" in terms of baseline characteristics. While we cannot prove the key identification assumption of comparability across the cutoff, the empirical evidence on the smoothness of the test score density and of baseline covariates provides reassurance.

In terms of external validity, our approach provides evidence on a local average treatment effect (LATE) of enrolling directly in a college-level math course. In this case, the LATE refers to a student subgroup characterized by two conditions. First, our estimates provide evidence relevant only to students near the cutoff score. The effects for students scoring farther away from the cutoff might differ from what we estimate. This is a standard limitation of RD approaches. Second, our estimates pertain to the "compliers" – that is, students who are induced to enroll in a college-level math course because their initial TSIA cutoff score is below the cut point. In particular, our estimates may not be applicable to students who retake the TSIA exam and score above the cutoff, or to students who go into college-level courses despite scoring below the cutoff. Despite these caveats, the expansive sample of administrative data on students across the state of Texas, as well as the rigorous quasi-experimental research design make this analysis informative for other contexts.

[5] The Benjamini–Hochberg procedure provides modified standards for statistical significance that become more stringent as the number of statistical tests within the same domain rises. See Benjamini and Hochberg (1995).

Table 3 First stage, regression discontinuity results

First Stage Fuzzy RD	Outcome: Placed into College Level Math		
	(I)	(II)	(III)
Above Cut Indicator	0.341*** (74.73)	0.239*** (39.27)	0.167*** (21.58)
Polynomial Degree	Linear	Quadratic	Cubic
Interact Above Cut Score with Polynomial	Yes	Yes	Yes
N	61,208	61,208	61,208

Columns show results of different estimations

T-Statistics in parentheses

Polynomial Degree indicates the order of a polynomial in the TSIA score that is used to control for the relationship between test score and probability of enrollment in developmental education classes away from the cut score

Interactions indicates the presence of an interaction term included with the polynomial, so that polynomial may vary in coefficients above and below the cut score

Demographic Covariate Control Set: Age, Asian, Black, Economic Disadvantage, Female, Hispanic, International, LEP, White, Part Time, Above cut score in reading and writing, Type of Major Declared, and Degree Intent

Sample: First Time in College Sample that took the TSIA in a given subject, Fall 2013-Fall 2015

Fixed Effects: Semester

*p < 0.1, **p < 0.05, ***p < 0.01

Results

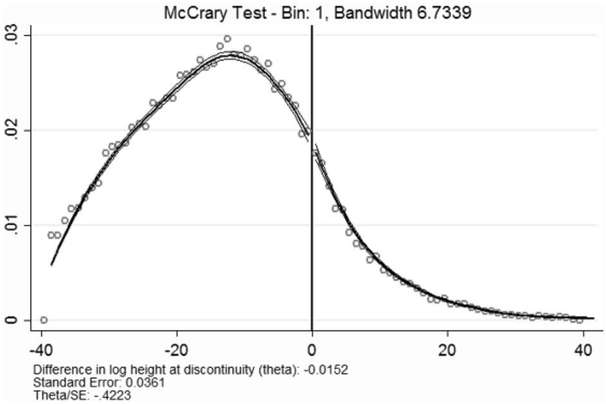
First Stage Results

As seen in Fig. 1, there is a sharp increase in the fraction of students enrolling in a college-level math course in the first semester at the cut score. Table 3 shows first stage point estimates with different parametric specifications for the TSIA math subject score. The estimated discontinuities range from 33 percentage points for the linear model to 17 percentage points for the cubic model. All estimates are precisely estimated, with the F-statistic (which is the square of the t-statistic on the "above cut indicator") range from 466 to 5580. These results indicate that "weak instruments" are not a concern in this analysis and that there is considerable statistical power to estimate treatment effects of enrolling in college-level math course in the first semester using the fuzzy RD design. The sensitivity of the estimates to polynomial choice can be explained by the curvature in the relationship between placement in a college-level course and TSIA DE placement test score and the very steep slope near the cutoff.[6] To address potential bias due to misspecification, we use the fully interacted cubic polynomial specification in subsequent analyses.

The first stage results in Fig. 1 and Table 3 clearly show that placement exam scores affect placement outcomes, but the "compliance" with the placement policy might differ across subgroups. For the pooled sample, the first-stage F-statistic (the square of the t-statistic on the excluded instrument) is always greater than 400, indicating a strong first-stage relationship. Appendix Table 1 shows first stage estimates by subgroup. There is a strong

[6] This steepness reflects test retaking; many students who initially score near the TSIA cutoff are able to meet the college-readiness standard.

Fig. 2 McCrary Test Results



first stage for all subgroups, although the magnitude does vary somewhat across subgroups. One implication of this pattern is that some of the heterogeneity we document may be due to our approach picking up different local effects across subgroups because the complier set might be different in one group relative to another.

Testing Identification Assumptions

Figure 2 shows the estimated density of the running variable used in the analysis. This is based on the non-parametric density estimator proposed in McCrary (2008) for testing smoothness of the running variable's density. The results shown in Fig. 2 use the default bandwidth choice (approximately 6.7), which results in no problematic discontinuity estimate in the density at the cutoff, as would be expected when capturing students' true first exams in our data. The estimated discontinuity is small and statistically insignificant, consistent with a smooth distribution at the cutoff point, with slightly more data on the left of the cutoff than on the right.[7]

While this evidence suggests there are no serious concerns with sorting around the cutoff, what ultimately matters is whether students on either side of the cutoff are similar in terms of other factors related to outcomes. To examine this, we analyze potential discontinuities in baseline covariates at the cutoff. The graphs in Appendix 1 show that the covariates used in the analysis are balanced around the cut score. Table 4 displays estimated discontinuities in these covariates using our preferred parametric specification. These estimates are small and statistically insignificant. These results suggest that students above and below the cut score are similar in all dimensions measured by our data.

[7] Given the discrete running variable, we use a bin size of 1 for this analysis.

Table 4 Covariance balance tests

Covariate balance tests—first stage regression, cubic polynomial with cut score interactions	Cut Score Indicator Coeff	Standard Error	P-value
Measures examined for heterogeneity			
3 + Years Math in High School	− 0.005	0.016	0.746
Academic Major	0.012	0.015	0.434
B.A. Intent	0.005	0.013	0.708
Part-time Enrollment	− 0.010	0.016	0.513
Tested Below College Ready in 2 + Subjects	− 0.021	0.013	0.106
Limited English Proficiency	0.013	0.010	0.191
Economic disadvantage	0.020	0.016	0.226
Age > 21	0.012	0.015	0.435
Other baseline characteristics			
Asian	0.009	0.006	0.118
Black	0.001	0.013	0.962
Hispanic	− 0.012	0.016	0.450
White	0.001	0.015	0.968
Female	− 0.004	0.017	0.800
International	0.006	0.005	0.193
Certificate Intent	− 0.003	0.008	0.728
Technical Major	− 0.020	0.014	0.169
Technical Prep Major	0.007	0.008	0.377
Parent has Some College	− 0.004	0.015	0.808
3 + Years English in High School[a]	− 0.001	0.014	0.970
Test scores			
Centered Reading TSIA Score	0.320	0.430	0.457
Centered Writing TSIA Score	0.471	0.375	0.209
N = 61,208			

Running Variable is TSIA score

Sample: First Time in College Sample that took the TSIA in a given subject, Fall 2013-Fall 2015

Fixed Effects: Semester

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

[a]Predicted success is a binary variable created in the following way: We run a probit using the full covariate set and generate predicted probability of successful persistence one year from entry. The variable is equal to 1 if a student scores at or above the median of the predicted values, zero if not

Main Results

Baseline results (see Appendix 2) suggest that the college readiness standard imposed by Texas policy was well-constructed for math, as there was no differential effect overall on passing a first college-level math course or persisting in college for students placed directly in a college-level math course near the cut score. The estimated effect on passing a FCL math course is small in magnitude and not statistically significant (coefficient = -0.008, p-value = 0.897) and also for persistence (coefficient = 0.010, p-value = 0.904). If the

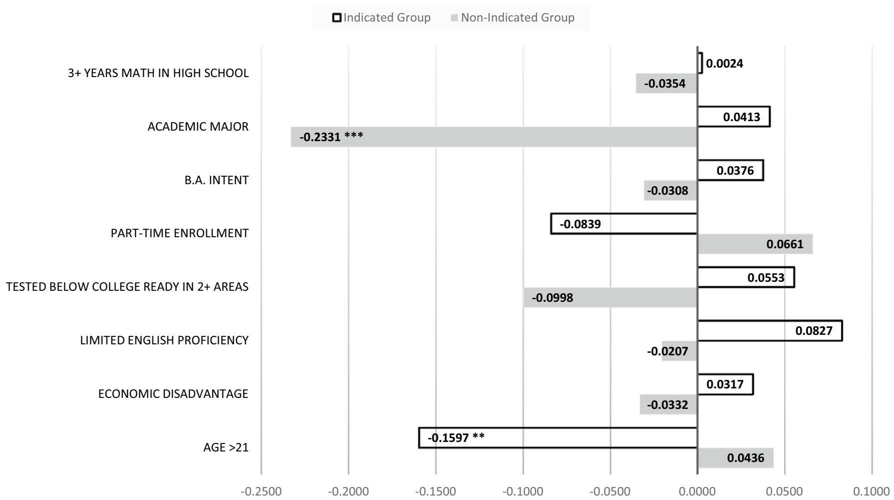


Fig. 3 Total Treatment Effects by Group, Outcome of Pass First College-level Math Course. Bars indicate estimates of the subgroup-specific treatment effects of enrolling directly in a college-level math course. Clear bars indicate the estimated impact of being placed directly in a college-level math course for students in the group with the characteristic indicated in the left-hand column of the figure, and shaded bars depict the estimated treatment effect for the non-indicated group. *$p<0.1$, **$p<0.05$, ***$p<0.01$. We only show significance for cases in which the p-value estimate is sufficiently small to pass the more stringent standards imposed by the Benjamini–Hochberg procedure

marginal student had a positive (negative) effect of being placed directly in a college-level course, then, assuming a smooth gradient in actual expected readiness along with the test score, we would expect that moving the cut score lower (higher) would be beneficial – that students near the cut score are more able (unable) to handle the material. However, we find evidence of substantial heterogeneity in impacts across student characteristics that might be used for multiple measures placement.

Figures 3 and 4 present our estimates of the subgroup-specific treatment effects of being placed directly in a college-level math course. For each row, the clear bar shows the estimated impact of being placed directly in a college-level math course for students in the group with the characteristic indicated in the left-hand column of the figure, and the shaded bar depicts the estimated treatment effect for the non-indicated group.[8] For example, in Fig. 3, the impact of being placed directly in a college-level course on the likelihood of passing a first college-level math class within a year for students older than 21 at the time of college entry was − 15.97 percentage points. Meanwhile, the same effect for students who were 21 years or younger was 4.36 percentage points. The stars contained in the bar labels indicate statistical significance for a subgroup's estimated treatment effect.[9] Figures 5 and 6 show the differences in treatment effects of being placed directly in a college-level math

8 The full baseline covariate set is included in each regression, but is not reported. Each regression adds one covariate interaction term individually to the analysis, while retaining the baseline covariate set.

9 The total effects shown in Figs. 3 and 4 are the result of adding the baseline effect φ and the interaction effect ϕ for the indicated group (as well as making the appropriate variance addition for the standard errors of the resulting estimate), and maintain the baseline effect and standard error for the non-indicated group in each regression. In Appendix 2, we include a table that shows the regression results in their direct form, including the baseline treatment effect, the additional interaction treatment effect, and the total effect, each with standard errors.

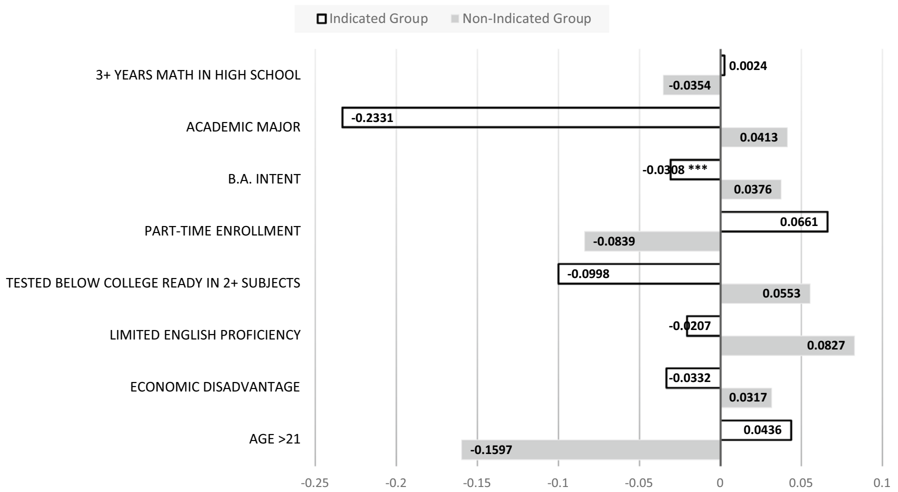


Fig. 4 Total Treatment Effects by Group, Outcome of Persistence. Bars indicate estimates of the subgroup-specific treatment effects of enrolling directly in a college-level math course. Clear bars indicate the estimated impact of being placed directly in a college-level math course for students in the group with the characteristic indicated in the left-hand column of the figure, and shaded bars depict the estimated treatment effect for the non-indicated group. *p < 0.1, **p < 0.05, ***p < 0.01. We only show significance for cases in which the p-value estimate is sufficiently small to pass the more stringent standards imposed by the Benjamini–Hochberg procedure

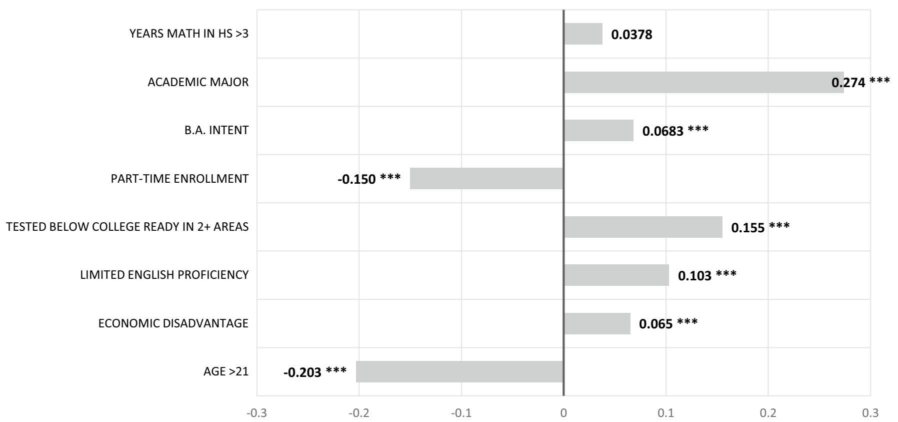


Fig. 5 Difference in Treatment Effects by Group, Outcome of Pass First College-level Math Course. Bars indicate differences in treatment effects of enrolling directly in a college-level math course between the indicated and non-indicated group. *p < 0.1, **p < 0.05, ***p < 0.01. We only show significance for cases in which the p-value estimate is sufficiently small to pass the more stringent standards imposed by the Benjamini–Hochberg procedure

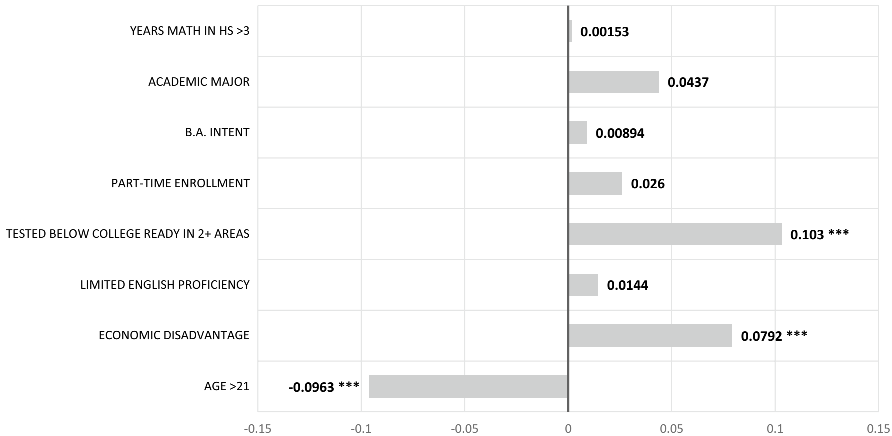


Fig. 6 Difference in Treatment Effects by Group, Outcome of Persistence. Bars indicate differences in treatment effects of enrolling directly in a college-level math course between the indicated and non-indicated group. $*p<0.1$, $**p<0.05$, $***p<0.01$. We only show significance for cases in which the p-value estimate is sufficiently small to pass the more stringent standards imposed by the Benjamini–Hochberg procedure

course for students with and without a particular characteristic. Again, the stars in the bar labels indicate the statistical significance of the difference in treatment effects.

While other studies have found high school achievement measures to be predictive of course success (e.g., Bahr et al. 2017; Ngo and Kwan, 2015; Scott-Clayton, Crosta, & Belfield, 2014), we found that our measure of three or more years of math in high school was not associated with greater likelihood of success when students were placed into college-level math coursework. It may be that this particular measure may not be as predictive as other high school achievement measures, as we might expect measures that incorporate more information on course performance (e.g., high school GPA, grade in highest math course) to provide additional value beyond knowing the number of courses taken.

We found that students who had declared an academic major and students who were pursuing a bachelor's degree were more likely to succeed when placed directly into college coursework than their peers pursuing technical programs and sub-baccalaureate credentials. There are a number of potential reasons these measures might be related to success when placed directly into college-level coursework. The completion of college-level math coursework might be a higher priority for academic, degree-seeking students, and these courses may be more likely limit to limit student access to other courses required for transfer and/or graduation. For example, completion of college algebra is often a prerequisite for many science, technology, engineering and mathematics courses. In addition, students pursuing different majors and types of credentials may differ in terms of levels of math preparation. The results also suggested that students who enrolled part-time were less likely to succeed than full-time enrolling students when placed into college-level coursework. Possible explanations for these relationships might include fewer course options fitting into part-time schedules, less time to devote to the college-level math coursework, and more external life circumstances that might hinder success.

It is worth noting that when looking at our academic major, bachelor's degree and part-time enrollment measures, we find heterogeneity with regard to our course passing outcome but not to our persistence measure. One possible explanation for this is that course placement (our treatment) is critical to a student's chances of enrolling in and succeeding in that specific course while many things contribute to student persistence, so we would expect to see a smaller degree of heterogeneity in persistence rates across all measures. Another explanation is that technical students and part-time students may be groups of students most likely to postpone math enrollment while they retested or dealt with other course requirements. These students were not necessarily failing or dropping out of college (i.e., "less successful"), but simply delaying math course-taking. However, to the degree that early math course-taking is important to performance in other coursework and long-term success, delaying college math may have been problematic.

Interestingly, results for several other measures— including requirements to take DE in other subject areas, LEP status, and economically disadvantaged status—ran counter to how we understand them to be used in the field for placement. Colleges might be hesitant to place students who are deemed "not college ready" in multiple subjects or LEP students into accelerated approaches because of concerns that the need for academic support might be greater. But in fact, the findings suggest that students testing below college readiness in multiple subjects and LEP students benefited more when placed into college-level coursework. The findings that LEP students benefit to an even greater degree from acceleration through DE mirror those found on another study (Hodara, 2015).

One possible explanation for these counterintuitive results is that math TSIA scores may have underestimated true math ability for students who struggled with reading the English-language content on a timed assessment. It also may be that low assessment scores in multiple subjects are driven by other assessment issues—such as a lack of preparation for the assessment or test anxiety issues—that limit the ability of the assessment to accurately measure a student's math ability. Testing below college ready also impacts the other courses a student is placed into. Students who tested below college readiness in multiple subjects may have a schedule full of DE courses, and the lack of momentum (i.e., progress in meeting early college milestones) may be demotivating more so than if a student is only required to take one DE course. Alternatively, students who test college-ready in reading and writing are more likely to be enrolled in college-level coursework in other subjects, and some students on the margins of readiness may face more challenges balancing tough course loads.

While TSIA guidance suggested that colleges consider transportation and financial barriers for placement, measures that are likely to be driven by economic disadvantage, our findings did not provide evidence supporting the use of these measures to limit access to college-level coursework. Economically disadvantaged students benefited more than their peers when placed directly into college-level math despite whatever additional barriers they may have faced. Again, these findings might be driven by assessment issues that disproportionately disadvantaged students with economic challenges and led to scores that underestimate math ability.

Finally, the evidence suggests that older students were worse off when placed into college-level coursework, and less likely to succeed relative to younger students. While we cannot determine exactly what drove these patterns, possible factors might have included the substantial time since students were in enrolled high school and engaged in math coursework or differences in the life circumstances that older students were dealing with that acted as barriers to academic success.

Conclusion

Multiple measures placement and holistic advising approaches have become a common practice in many colleges across the country. State policies often allow colleges substantial autonomy to determine which measures and approaches they will use, and allow for wide variation in placement practices across institutions, and sometimes across advisers within institutions. And while there is strong theory and emerging experimental evidence to support the use of multiple measures, there is limited evidence available on how best to implement these approaches. While evidence suggests that high school GPA is one valuable measure for colleges to incorporate, many colleges do not have access to transcript data for the purposes of placement. Colleges may instead be incorporating measures into placement for which there is little rigorous evidence suggesting that the factors predict success in the expected ways, and with little consideration of the implications of using these measures for equity.

We examined heterogeneity in the impacts of college-level course assignment for a set of factors being considered for the purposes of placement by some Texas colleges. Our findings suggest that major and degree program may be valuable to consider in making placement decisions, consistent with the guidance colleges in Texas and across the country are receiving through the Guided Pathways movement. And while our evidence suggests that part-time enrollees might be somewhat less successful in college-level coursework, it will be important for colleges to think about the equity implications of disproportionately offering accelerated options to full-time enrollees.

On the other hand, our analysis identified several measures where the relationships run counter to common assumptions and suggest that colleges should be cautious in the use of these measures for placement. For example, our findings suggest that students who test below college readiness in multiple subjects might be better off when placed into college-level math, as these students might have less strenuous course requirements in other subject areas or might be in need of some opportunity to gain momentum in at least one subject. LEP students and economically disadvantaged students also saw greater success when placed into college-level math. In each of these cases of counterintuitive findings, a lack of predictive value of the placement exam due to testing issues is a likely factor. The use of these three measures by colleges to place students into DE courses (and withhold accelerated options) may lead to negative impacts on student success rates as well as having negative implications for equity.

These results require careful use and interpretation. First, it is important to note that the findings are generalizable to students who fall within test score ranges that are near college ready. In many cases colleges are focusing their use of multiple measures on students in these higher ranges so this is a key population of interest, but patterns may differ for students as multiple measures practices are scaled to students testing further below the college-ready cut score. In addition, our study focuses on students placed into college-level coursework without any additional support, so our findings cannot necessarily be generalized to placement into corequisite models that pair college-level placement with additional academic support. More generally, our findings may not be generalizable to other states and institutions with different placement exams, and course placement options. In particular, some of the non-academic measures we consider may only be proxies for the factors used by advisers in practice; for example, a more targeted discussion with a student may elicit more precise measures of responsibilities, needs, and barriers than our measures of age 21 or older and economic disadvantage.

It is also worth noting that many of our findings on variation were driven by one group of students seeing no impacts from enrolling in the college-level course and another group of students seeing positive impacts. In this case the best policy may be to accelerate both groups of students. There is a growing body of evidence suggesting that acceleration into college-level coursework with concurrent DE support (i.e., corequisites) might be more effective than standalone DE courses for most students, and the movement of many states to eliminate standalone DE suggests a quickly evolving context for multiple measures placement. It is unclear whether the heterogeneity we found in the traditional placement context (standalone DE versus standalone college-level math) would translate to this new context where corequisites are the primary DE option, so further research would be needed to identify promising multiple measures placement factors in this context.

Appendix 1

See Fig. 7.

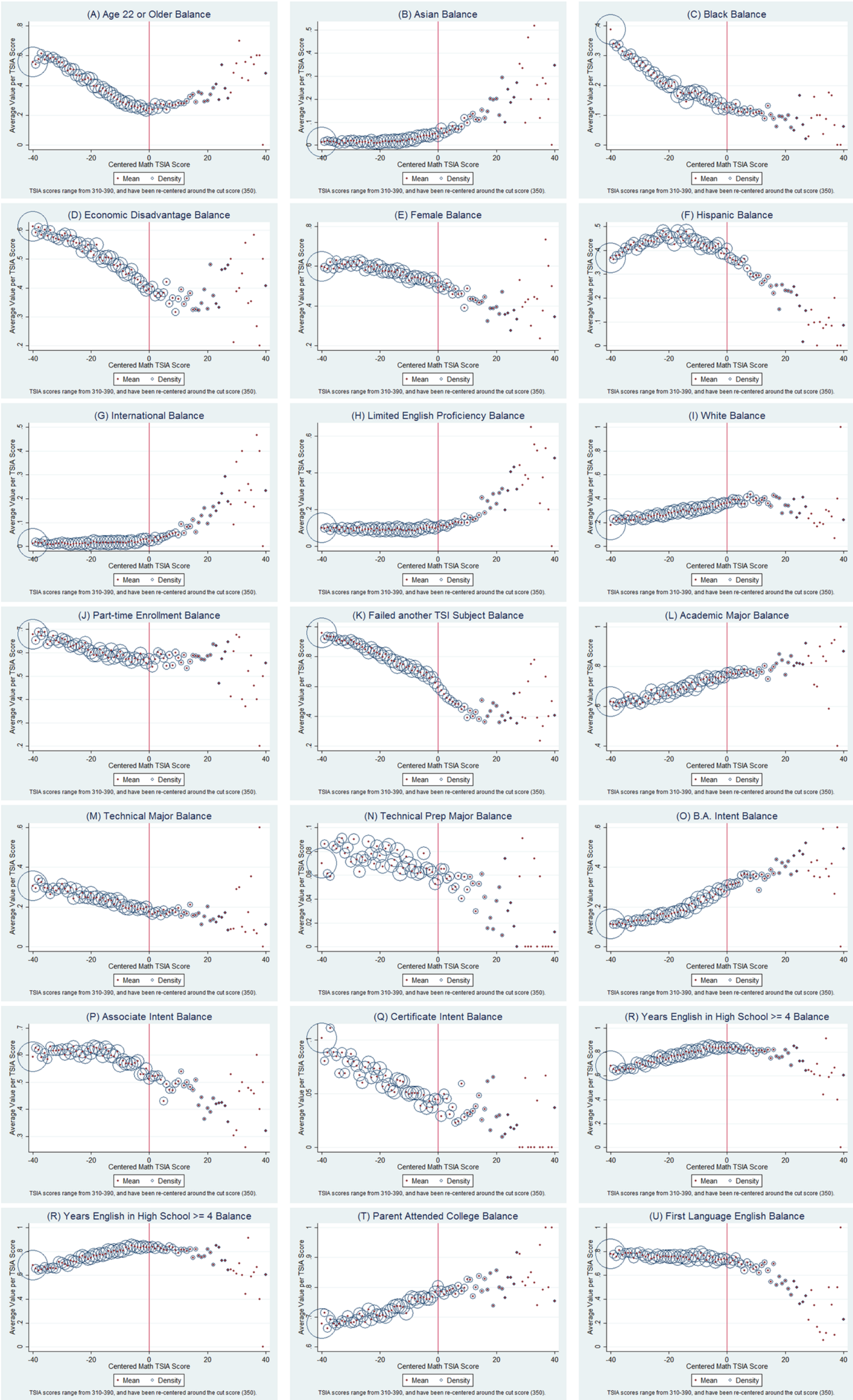


Fig. 7 Covariate balance graphs

Appendix 2

See Tables 5, 6.

Table 5 First-stage estimates for key measures

First Stage Fuzzy RD	Estimated Size of Discontinuity at Cut Score		
	(I)	(II)	(III)
All Student Average	0.341***	0.239***	0.167***
(N = 61,208)	(74.73)	(39.27)	(21.58)
Among Students with:			
Years Math in HS > 3	0.220***	0.136***	0.0846***
(N = 25,418)	(35.26)	(16.31)	(7.89)
Academic Major	0.238***	0.156***	0.109***
(N = 25,452)	(38.28)	(18.69)	(10.08)
B.A. Intent	0.211***	0.114***	0.0562***
(N = 7519)	(17.87)	(7.10)	(2.69)
Part-time Enrollment	0.176***	0.121***	0.0850***
(N = 19,549)	(28.15)	(14.43)	(7.88)
Tested below college ready in 2 + areas	0.253***	0.154***	0.0929***
(N = 27,903)	(42.57)	(19.67)	(9.39)
Limited English Proficiency	0.315***	0.256***	0.197***
(N = 4,272)	(18.60)	(10.91)	(6.40)
Economic Disadvantage	0.250***	0.170***	0.113***
(N = 18,534)	(34.64)	(17.61)	(9.09)
Age > 21	0.229***	0.162***	0.133***
(N = 14,048)	(28.92)	(14.85)	(9.37)
Polynomial Degree	Linear	Quadratic	Cubic
Interact Above Cut Score with Polynomial	Yes	Yes	Yes

Columns show results of different estimations, T-statistics in parentheses

Polynomial Degree indicates the order of a polynomial in the TSIA score that is used to control for the relationship between test score and probability of enrollment in developmental education classes away from the cut score

Interactions indicates the presence of an interaction term included with the polynomial, so that polynomial may vary in coefficients above and below the cut score

Demographic Covariate Control Set: Age, Asian, Black, Economic Disadvantage, Female, Hispanic, International, LEP, White, Part Time, Above cut score in reading and writing, Type of Major Declared, and Degree Intent

Sample: First Time in College Sample that took the TSIA in a given subject, Fall 2013-Fall 2015

Fixed Effects: Semester

* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$

Table 6 Separate Main and Interaction Effects

	Placed into College Level		College Level*Group		Total Effect for Group	
	(Coeff.)	(P-Value)	(Coeff.)	(P-Value)	(Coeff.)	(P-Value)
Outcome: Pass FCL Math						
Base	– 0.00769	(0.8966)				
3 + Years Math in High School	– 0.0354	(0.5847)	0.0378*	(0.0738)	0.0024	(.0580)
Academic Major	– 0.233***	(0.0027)	0.274***	(0.0000)	0.0413	(.0566)
B.A. Intent	– 0.0308	(0.6221)	0.0683***	(0.0011)	0.0376	(.0559)
Part-time Enrollment	0.0661	(0.1488)	– 0.150***	(0.0000)	– 0.0839	(.0743)
Tested below college ready in 2 + areas	– 0.0998	(0.1222)	0.155***	(0.0000)	0.0553	(.0568)
Limited English Proficiency	– 0.0207	(0.7270)	0.103***	(0.0001)	0.0827	(.0636)
Economic Disadvantage	– 0.0332	(0.5906)	0.0650***	(0.0005)	0.0317	(.0567)
Age > 21	0.0436	(0.4386)	– 0.203***	(0.0000)	– .1597**	(.0708)
Outcome: Persistence						
Base	0.0101	(0.9043)				
3 + Years Math in High School	0.00891	(0.9224)	0.00153	(0.9590)	0.0104	(.0820)
Academic Major	– 0.0254	(0.8163)	0.0437	(0.3198)	0.0184	(.0794)
B.A. Intent	0.00773	(0.9301)	0.00894	(0.7618)	0.0167	(.0789)
Part-time Enrollment	– 0.00272	(0.9664)	0.0260	(0.5861)	0.0233	(.1051)
Tested below college ready in 2 + areas	– 0.0510	(0.5763)	0.103***	(0.0006)	0.0518	(.0803)
Limited English Proficiency	0.00825	(0.9215)	0.0144	(0.6993)	0.0227	(.0901)
Economic Disadvantage	– 0.0211	(0.8095)	0.0792***	(0.0029)	0.0581	(.0803)
Age > 21	0.0344	(0.6654)	– 0.0963***	(0.0063)	– 0.0620	(.0998)
N = 61,208						

Fuzzy RD, Cubic Score Specification: Running Variable is TSIA score, Instrument is at or above the state cutoff

Sample: First Time in College Sample that took the TSIA in a given subject, Fall 2013-Fall 2015

Demographic Covariate Control Set: Age, Asian, Black, Economic Disadvantage, Female, Hispanic, International, LEP, White, Part Time, Above cut score other subjects, Type of Major Declared, and Degree Intent

Fixed Effects: Semester

Outcome horizon is two long semesters (discounting summer terms) from initial enrollment semester

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Acknowledgements The research reported here was supported, in whole or in part, by the Institute of Education Sciences, U.S. Department of Education, through grant R305H150069 to the RAND Corporation. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education. The research benefitted tremendously from helpful input and strong guidance and support from Texas Higher Education Coordinating Board staff. In particular we thank Julie Eklund, Suzanne Morales-Vale, Melissa Humphries, and David Gardner.

References

Bahr, P. R., Hetts, J., Hayward, C., Willett, T., Lamoree, D., Newell, M. A., Sorey, K., & Baker, R. B. (2017). *Improving placement accuracy in california's community colleges using multiple measures of high school achievement*. The Research and Planning Group for California Community Colleges.

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.

Barnett, E., & Reddy, V. (2017). College Placement Strategies: Evolving Considerations and Practices 1. In *Preparing Students for College and Careers* (pp. 82–93). Routledge.

Barnett, E. A., Bergman, P., Kopko, E., Reddy, V., Belfield, C., & Roy, S. (2018). *Multiple Measures Placement Using Data Analytics An Implementation and Early Impacts Report*. Center for the Analysis of Postsecondary Readiness (CAPR): New York NY.

Belfield, C. R., Jenkins, D., & Fink, J. (2019). *Early momentum metrics leading indicators for community college improvement (CCRC Research Brief)*. Columbia University, Teachers College, Community College Research Center.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57*(1), 289–300.

Boatman, A & Long, B.T. (2013) The role of remediation and developmental courses in access and persistence. A. Jones, L. Perna (Eds) The state of college access and completion improving college success for students from underrepresented groups. Routledge Books, New York

Boatman, A., & Long, B. T. (2018). Does remediation work for all students? How the effects of postsecondary remedial and developmental courses vary by level of academic preparation. *Educational Evaluation and Policy Analysis, 40*(1), 29–58.

Calcagno, J. C., & Long, B. T. (2008). The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance (No. w14194). New York: National Center for Postsecondary Research.

Cho, S. W., Kopko, E., Jenkins, D., & Jaggars, S. S. (2012). New evidence of success of community college remedial English students: Tracking the outcomes of students in the Accelerated Learning Program (ALP) (CCRC Working Paper No. 53). New York, NY: Columbia University, Teachers College, Community College Research Center.

Cuellar Mejia, C., Rodriguez, O., & Johnson, H. (2016). *Preparing students for success in California's community colleges*. Public Policy Institute of California.

Cullinan, D., Barnett, E.A., Ratledge, A., Welbeck, R., Belfield, C. & Lopez, A. (2018). Toward better college course placement: a guide to launching a multiple measures assessment system. New York NY: Center for the Analysis of Postsecondary Readiness (CAPR).

Daugherty, L., Gomez, C. J., Carew, D. G., Mendoza-Graf, A., & Miller, T. (2018). *Designing and implementing corequisite models of developmental education: findings from texas community colleges*. RAND Corporation.

Gehlhaus, D., Daugherty, L., Karam, R., Miller, T. & Mendoza-Graf, A. (2018). Practitioner Perspectives on Implementing Developmental Education Reforms: A Convening of Six Community Colleges in Texas. Working Paper, Santa Monica, CA: RAND Corporation.

Hodara, M. (2015). The effects of english as a second language courses on language minority community college students. *Educational Evaluation and Policy Analysis, 37*(2), 243–270.

Hodara, M., & Cox, M. (2016). *Developmental education and college readiness at the University of Alaska (No. 2016-123)*. Portland OR: Regional Educational Laboratory Northwest.

Hu, S., Park, T. J., Woods, C. S., Tandberg, D. A., Richard, K., & Hankerson, D. (2016). Investigating developmental and college-level course enrollment and passing before and after Florida's developmental education reform (REL 2017–203). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

Jenkins, D. & Bailey, T. (2017). Early Momentum Metrics: Why They Matter for College Improvement (CCRC Research Brief No. 65). New York, NY: Columbia University, Teachers College, Community College Research Center.

Klasik, D., & Strayhorn, T. L. (2018). The complexity of college readiness: Differences by race and college selectivity. *Educational Researcher, 47*(6), 334–342.

Logue, A. W., Douglas, D., & Watanabe-Rose, M. (2019). Corequisite mathematics remediation: results over time and in different contexts. *Educational Evaluation and Policy Analysis*. https://doi.org/10.3102/0162373719848777

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.

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

Miller, T., Daugherty, L., Martorell, P., & Gerber, R. (2020). *Assessing the Effect of Corequisite English Instruction Using a Randomized Controlled Trial*. American Institutes for Research.

Ngo, F., & Kwon, W. W. (2015). Using multiple measures to make math placement decisions: Implications for access and success in community colleges. *Research in Higher Education, 56*(5), 442–470.

Ngo, F., Kwon, W., Melguizo, T., Prather, G., & Bos, J. M. (2013). Course placement in developmental mathematics: Do multiple measures work. Los Angeles, CA: The University of Southern California Working Paper, as of 12/1/2019: https://pullias.usc.edu/wp-content/uploads/2013/10/Multiple_Measures_Brief.pdf.

Park, T., Woods, C., Tandberg, D., Richard, K., Cig, O., Hu, S., & Bertrand Jones, T. (2016). Examining student success following developmental education redesign in Florida. Teachers College Record.

Ran, F. X., & Lin, Y. (2019). *The effects of corequisite remediation: evidence from a statewide reform in Tennessee*. Columbia University, Teachers College, Community College Research Center.

Scott-Clayton, J. and Rodriguez, O. (2012). Development, discouragement, or diversion? new evidence on the effects of college remediation. NBER Working Paper No. 18328.

Scott-Clayton, J., Crosta, P. M., & Belfield, C. R. (2014). Improving the targeting of treatment: evidence from college remediation. *Educational Evaluation and Policy Analysis, 36*(3), 371–393.

Rutschow, E. Z., Cormier, M. S., Dukes, D., & Zamora, D. E. C. (2019). *The Changing Landscape of Developmental Education Practices: Findings from a National Survey and Interviews with Postsecondary Institutions*. Center for the Analysis of Postsecondary Readiness.