National Poverty Center Working Paper Series

#11-14

April 2011
Revised October 2011
Revised January 2013

Does Failing a Placement Exam Discourage Underprepared Students from Going to College?

Paco Martorell, RAND
Isaac McFarlin Jr., University of Michigan
Yu Xue, University of Texas-Dallas

Does Failing a Placement Exam Discourage Underprepared Students from Going to College?

Paco Martorell
RAND
paco@rand.org

Isaac McFarlin Jr.
University of Michigan
imcfar@umich.edu

Yu Xue
University of Texas at Dallas
yxx036000@utdallas.edu

December 2012

Abstract

About one-third of college students are required to take remedial courses to redress the lack of preparation for college-level coursework. Assignment to remediation is generally made on the basis of performance on a placement exam. In some states, students are required to take a placement test prior to enrolling in college-level courses. In this type of setting, assignment to remediation may have the unintended effect of dissuading students from actually going to college. This is because remedial courses typically do not count towards a college degree, so remediation might increase the cost of college by increasing the time required to complete a degree. Furthermore, being identified as needing remediation could directly affect enrollment rates via stigma effects or by providing students with information about their suitability for college. This paper examines this issue empirically using administrative data from Texas. Using regression discontinuity methods, we find that students whose placement exam scores would require them to be in remediation are no less likely to enroll in college than are those scoring just above the remediation placement cutoff.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B07581 to the University of Texas at Dallas. The opinions expressed are those of the authors and do not represent views of the Institute, the U.S. Department of Education, or other organizations. We wish to thank Sheldon Danziger, Mark Long, and participants at the Western Economic Association conference for helpful comments. All errors are our own.

1. Introduction

Nearly one-third of entering students at public colleges and universities participate in remedial education programs (NCES, 2003). At the college level, remedial education consists of high-school-level coursework (that does not contribute to a college degree) and services such as tutoring that are designed to provide underprepared students with the academic skills needed to succeed in college. Expenditures on college remediation are considerable, with estimates suggesting that public colleges spend at least $1 billion per year on remediation (Breneman and Haarlow, 1998).

Assignment to remediation is typically determined by performance on a placement exam, with students scoring below some specified threshold required to enter remediation (Collins, 2008). Although these assessments are generally not used for admissions purposes, concerns have been expressed about whether they might have unintended consequences. Failing a placement exam imposes the additional requirement of having to participate in remedial education on degree-seeking students. As such, college placement tests may dissuade some students with weaker academic skills from pursuing postsecondary schooling. Some college administrators cite the stringency in placement exam standards as a contributing factor to lower enrollments, particularly for the less-advantaged and racial minorities (Stewart, 1995).

In this paper, we examine whether the additional burden of being required to enter remediation discourages students who do poorly on a placement exam in a way that affects their college enrollment behavior. Essentially, we estimate the effect of informing students that they are academically deficient and that they therefore must take non-credit bearing remedial coursework to pursue a degree. This "treatment" could have a discouragement effect for at least three reasons. First, the placement exam may provide students with new information about their academic skills. Many students overestimate their ability to succeed in college (Venezia et al., 2003). If students are informed that they are not immediately ready to undertake a college-level curriculum, they may decide to pursue other endeavors. Another reason is that being assigned to remediation increases college costs since remedial courses do not count towards a degree, yet students must still pay tuition for remediation (Collins, 2008). Moreover, having to take remedial courses could lengthen the time needed to complete a degree, which would increase the opportunity cost of college. Finally, there may be a "stigma" effect whereby students decide not to enroll to avoid being stigmatized as academically deficient (Kingan and Alfred, 1993).

Understanding whether such a discouragement effect exists has important policy implications that depend on societal priorities. If maximizing the number of students attending college is a policy aim, as President Barack Obama suggested in a recent speech, then uncovering discouragement effects might suggest that remediation assignment and implementation should be redesigned to minimize the negative effects on college-going.[1] An alternative view, however, is that policies aimed at maximizing post-secondary enrollment are not efficient since some students lack the ability and motivation to succeed in college, as evidenced by low completion rates of underprepared students (Rosenbaum, 2002). In this light, discouragement effects brought about from failing placement tests might be viewed positively.

Although we are unaware of previous empirical work examining the link between assignment to remediation and college enrollment behavior, some evidence suggests that students do take steps to avoid remediation. Bostian (2008) finds that students in need of remediation at North Carolina two-year colleges are more likely to transfer to non-selective four-year schools that do not have a remediation requirement. Likewise, there is anecdotal evidence of students choosing community colleges in a given metropolitan area on the basis of having lower passing scores for remedial placements (Moltz, 2009a).[2]

We analyze administrative data on four cohorts of Texas high school graduates. Texas is a useful state for conducting this study for three reasons. First, up until 2003, Texas used a statewide exam and a common cutoff score for remedial placements at most public colleges and universities in the state. Thus, unlike other states where there is no uniform standard or where some institutions do not offer

[1] President Barack Obama, August 5, 2010, University of Texas at Austin.

[2] Lindo et al. (2010) and Fletcher and Tokmouline (2010) examine the effects of being placed on academic probation while in college, which can be thought of as a type of informational shock among students who have already enrolled in college. Our study, on the other hand, seeks to understand whether being labeled unprepared for college affects enrollment behavior. Howell et al. (2010) find that California's Early Assessment Program, a voluntary school-based program designed to increase the quality of information on academic preparedness among high school juniors, reduces the need for remediation at one large California public university. In contrast, we assess whether the "informational shock" from being labeled academically underprepared decreases the chances that students actually go on to higher education

remediation, poorly prepared students would find it challenging to avoid remediation by choosing different institutions. Second, during the study period, students were required to take the placement exam before enrolling in college-level courses (THECB, 1998). This implies that there is some scope for the outcome of the placement exam to potentially affect college enrollment behavior. Finally, detailed college and high school administrative records are available for Texas students which make it feasible to track high school graduates who take the placement test and observe their college enrollment behavior.

To account for the fact that students who perform poorly on placement tests may be less likely to go to college for other reasons, we use a regression discontinuity design based on the test score cutoffs used to assign students to remediation. In particular, we use data on a sample of students who take the state's primary placement exam prior to entering college and compare the college-going outcomes of students who score just above and just below the passing cutoff for assignment to remediation. Under relatively weak assumptions supported by the data, this comparison identifies the effect of failing the placement exam on the enrollment behavior of students at the margin for being placed into remediation.

We begin by documenting that failing the state's college placement exam results in a sharp increase in the probability of participation in remediation (as intended by the policy). We then show that students who would be assigned to remediation because of low placement exam scores are significantly less likely to enroll in college than are students who would not be assigned to remediation. However, the regression discontinuity estimates suggest that, for students near the remediation assignment threshold, this association does not reflect a causal relationship. The point estimates are statistically insignificant and our preferred estimates suggest we can rule out effects half as large as the raw difference in college enrollment rates between exam passers and failers. The results are robust to a variety of specifications and also hold for several student subgroups. One qualification to these findings is that they do not address whether the presence of a mandatory remediation policy dissuades any students from enrolling in college. Instead, our results only speak to the impact of placement test performance on enrollment behavior.

This paper is organized as follows. Section 2 provides background information on remedial education in general and the key institutional details for Texas during our study period. Section 3 describes the data. Section 4 presents the empirical framework. The results are discussed in Section 5, and Section 6 concludes the paper and discusses possibilities for future research.

2. Background

2.1 Remedial Education and Mandatory Placement Exams

Virtually all public community colleges and most four-year colleges provide some form of remedial instruction (NCES, 2003). Although relatively informal methods, such as referral by an academic advisor, are sometimes used to place students into remediation, assignment is typically determined on the basis of performance on an entry assessment known as a college placement exam. Nonetheless, considerable variation exists in the tests that are used to make remedial placements. Abraham (1986) found that hundreds of combinations of placement tests across states are used for assigning first-year students into remediation. Some states, such as Ohio, grant institutions considerable latitude in their choice of assessments and standards for making placements (Bettinger and Long, 2009). However, it appears that the national trend is toward greater standardization (Hughes and Scott-Clayton, 2010). Twenty-seven states currently require public institutions to administer placement tests, and nineteen states use a common test and passing standards (Collins, 2008).

There is also variation in the timing for when students must take these tests. Many students take entry assessments shortly before starting classes, while some do so during their first term in college. However, there is a nascent movement in some states toward testing students much earlier in high school to obviate the need for remediation in college (Long and Riley, 2007; Howell et al., 2010). Still, another policy adopted by states is to embed college-ready diagnostics within perennial assessments such as a state's high school exam. This strategy, adopted by several states, including Texas, allows highachieving students to place out of remediation unknowingly, well before applying for admission (Dounay, 2006).

Another feature of remedial education is that schools often provide different tiers of remedial classes because of variation in the level of preparation of remedial students. In particular, students who score extremely poorly on the placement exam are assigned to the lowest tier while those who score better, but still in the range where remediation is required, are placed into more advanced tiers. For instance in Texas, most public colleges have multiple levels of remedial courses for a given discipline (THECB, 2005).

2.2 How Might Mandatory Placement Exams Affect Enrollment Behavior?

Among students considering postsecondary schooling, performance on placement tests largely determines their subsequent assignment to remediation. Several recent studies have examined the impact of college remediation. Bettinger and Long (2009) find that remedial coursework in Ohio improves student outcomes. However, discontinuity-based estimates from Florida (Calcagno and Long, 2009) and Texas (Martorell and McFarlin, 2011) offer little evidence that remediation confers benefits. This paper examines a different question: namely, whether performing poorly on a placement exam and being subsequently labeled "academically underprepared" affects enrollment-related decisions among students assessed before they enter college.

There are multiple reasons why performing poorly on a placement test and being subsequently assigned to remediation could affect enrollment. First, students must pay for remedial coursework even though it does not provide college credits. For this reason, being assigned to remediation could increase total college costs by increasing the total number of courses that students must complete, which has direct tuition costs as well as opportunity costs in the form of foregone wages (Alliance for Excellent Education, 2006). Thus, one might expect that a fully-informed student who behaves rationally and takes into account the full cost of a college education may be less likely to enroll in college upon learning they must take remedial courses.[3] On the other hand, effective remedial programs could speed up student progress in subsequent courses, thereby offsetting the time spent in remedial classes, and hence the costs of being assigned to remediation.

Second, students who fail a mandatory placement exam may avoid college because of the "social stigma" associated with remediation. Although inherently difficult to measure, it has been argued that students could develop a negative self-image upon reviewing inferior test results from a placement exam; they could feel ashamed to take courses with "slower" peers (Kingan and Alfred, 1993). Students may also be reluctant to go through remediation because of the stigma associated with being identified as a weaker student (Moltz, 2009b; Bostian, 2008). However, if students with weaker academic skills are placed into classes commensurate with their level of preparation, then failing the placement test could make underprepared students feel better about entering college, thus, tempering some stigmatization (Palmer and Davis, 2012).

A third reason is that placement exams provide new information about one's ability to succeed in college. As noted by Manski (1989), some students have incomplete information about their skills and preferences for postsecondary schooling. Furthermore, because some students overestimate their academic preparation, finding out that remediation is required may come as a surprise (Venezia et al., 2003). For those students believing they were college-ready, the realization that they are in need of remediation might lead to "cooling out" their educational aspirations (Clark, 1960).

2.3 Institutional Details in Texas During the Study Period

During the study period, the Texas Academic Skills Program (or TASP) governed remediation placement policies in Texas. It required that all students pursuing academic degrees in public colleges and universities in the state – both two- and four-year institutions – enter remediation if they could not

[3] Along similar lines, students may attempt to avoid costly mandatory remedial requirements by enrolling at institutions with weaker or no requirements. During our study period, all public colleges face statewide policies regarding remedial placements, so this would be less of a concern. Placement exam performance might affect enrollment behavior by altering the decision to attend a 2-year versus 4-year school, since tuition is lower in 2-year colleges as compare to 4-year colleges (NCES, 2003). However, in results not reported using the approach described below, we find no significant effects of placement exam performance on the decision to attend a 2- versus 4-year school.

demonstrate college readiness. Students who scored well enough on the state's high school graduation exam or on the SAT or ACT college entrance exams were automatically deemed college ready. Other students had to demonstrate college readiness by meeting performance standards on a college placement exam, the most prominent of which was the TASP test.[4] Students could also take other exams (such as the College Board's Accuplacer) to comply with the placement testing requirement, but the TASP test was by far the most commonly taken test, representing over 70 percent of initial placement test attempts (THECB, 2000). While the TASP test was used to make assignments to remediation, the TASP policy explicitly stated that they were not to be used for admissions purposes.

The TASP test was offered six times per year through statewide administrations.[5] To take the test, students had to pay a nominal fee. They receive score reports in the mail within 2 weeks of taking the test (Boylan, 1996). The score reports inform students of minimum passing standards, whether the student passed or failed each section, and their exact score on each section. The TASP test consisted of math, reading, and writing exams, and students who failed one or more sections of the exam were required to enter remediation. Importantly, Texas policy during the study period only stipulated that students had to be in remediation for at least one subject, but did not specify the subjects for which a student would have to be in remediation (although individual institutions could make their own such requirements (THECB 2005)).[6] Furthermore, failing a particular subject did not necessarily have any implications about the number of remedial classes one must take. However, students were required to retake all failed sections of the TASP and were ineligible to receive an Associate's Degree until all sections of the test had been passed.

---

[4] About 23 percent of students enrolled in college received exemptions from the testing requirement by scoring high enough on the exit exam, the SAT or the ACT (THECB, 2000).

[5] Because the state administers the TASP only six times annually, some colleges offer a "Quick TASP" that can be given on short notice. Preliminary score reports for the Quick TASP for math and reading are available within 24 hours of the testing contractor receiving the test documents. Final results that include the writing section are available within 5 business days. Based on our calculations, the median time between the first and second attempts for students in our sample is 5 months. Very few (3%) took it less than 1 month after the first attempt, and only 8% took it more than 2 years after their first attempt. This suggests that although students had opportunities to retake the TASP within a narrow period of time, in practice, this was not typically done.

[6] Note that students could retake a section of the TASP test without taking remedial classes in that subject.

The policy in place during our study period established a minimum passing cutoff that applied statewide at all public institutions. For each section, raw scores were translated into a scale score that ranged from 100 to 300. The math and reading sections were multiple-choice exams with a passing score equal to 230. The main portion of the writing exam was an essay that was scored by two graders on a 1 to 4 scale. If the sum of the scores was less than 5, students failed automatically and if it was exactly 5, they must score sufficiently well on the writing multiple choice section to pass.[7] Although the minimum passing score was set by the state, individual institutions could use a higher cutoff to place students into remediation. However, all of our regression discontinuity analyses use the statewide passing cutoff since institution-specific passing cutoffs are only available for students who enroll in college.[8] Moreover, a survey conducted during our study period found that the vast majority of institutions did not set higher passing standards than those specified by the state (THECB 2005).[9]

The way the writing section was scored poses two challenges for our proposed regression discontinuity design. First, it introduces "lumpiness" in the score. Second, since graders know the scores needed to pass, students receiving the minimum passing score could be very different from students with the next lowest score. Therefore, as we describe in greater detail below, the regression discontinuity analysis uses the math and reading exams as "running variables" and our primary results use the sample of students that passed the writing exam.

[7] A useful way of interpreting these cutoffs is by examining the college completion rates of students scoring close to them. For students within 3 scale score points of passing all sections, the 6-year degree completion rate for students initially enrolling at a two-year college is about 30 percent (compared to 32 percent overall) and it is about 38 percent for students starting at a four-year college (compared to 48 percent overall; Martorell and McFarlin, 2011). This suggests that students at the margin for assignment to remediation have only moderately lower college completion rates than the overall population of Texas college-goers, which in turn reflects how widespread remediation is.

[8] To the extent that a student attends an institution that uses a higher passing cutoff than the state minimum, we may misclassify a student as not in need for remediation when in fact they were. Thus, our results should be interpreted as the "intent to treat" effect of being just above/below the cutoff set by the state for assignment to remediation. However, the empirical evidence we present below (in Figure 1) suggests that very few students scoring above the state passing standard went into remediation. In contrast, if the use of higher passing scores were widespread, then the rate of remediation just above the statewide minimum would be much higher than it actually was.

[9] Specifically, this survey finds that in 2000, 78 percent of institutions used the statewide cutoff for math; 94 percent did so for reading; and 90 percent did so for writing.

An important consideration for this study is the timing of the TASP exam. Unlike most tests administered in high school, there is no set schedule for taking the TASP. During our study period, Texas policy required that students take the TASP (or one of the other exams that could be used to demonstrate college readiness) before enrolling in college-level coursework. In practice, most of the students who we observe taking the TASP did so prior to entering college or did not enter college at all. However, some students (about 40 percent) took the TASP test for the first time after enrolling in college. One reason for this may be that students took and failed another placement test (which we do not observe) prior to entering college and then took the TASP after enrolling in college. A second possibility is that schools may not have been fully compliant with the regulation and allowed students to enroll as long as they took the TASP early in their first term.[10] Regardless of the reason, the outcome of the TASP test by definition cannot affect the enrollment behavior of students who took it after enrolling, so as we explain further below, we exclude students who were enrolled in college when they took the TASP test.

The consequences of performing poorly on a placement exam are significant for Texas students intending to pursue higher education. Obviously, one should expect failing the TASP to exert a sizeable effect on whether students enter remediation. This can be seen in Figure 1, which plots the fraction of students who entered remediation as a function of the initial-attempt TASP score.[11] Students who scored just below the minimum college readiness threshold on at least one subject (i.e., values of the horizontal axis of Figure 1 below zero) were about 55 percentage points more likely to enter remediation than were students who barely met the college readiness threshold (i.e., values of the horizontal axis of Figure 1 at or above zero). Figure 1 also shows that some students who scored below the college readiness threshold did not enter remediation. One reason is that this graph includes students who never entered college at

[10] Prior to fall 1998, students could enroll in 9 college-level semester credit hours before taking a placement exam.

[11] The score used on the horizontal axis is the minimum of TASP math and reading scale scores, where each score has been rescaled to be zero at the passing cutoff. Defined in this way, students with scores below zero are assigned to remediation and students with scores at or above zero are not. See the following section for details on the data underlying Figure 1 and Section 4.2 for details on the construction of the running variable used in the analysis.

all, or who dropped out of school before entering remediation.[12] Another reason is that students who initially failed the TASP could retake it and pass it, thereby avoiding the mandatory remediation requirement. Nonetheless the evidence in Figure 1 clearly shows that students that failed the initial-attempt of the TASP test faced a sharply higher probability of going into remediation if they were to enroll in college. In terms of "intensity" of the treatment, among enrollees, students barely failing the TASP take about 1.5 additional remedial classes compared to barely passers within the first year after they enroll in college (Martorell and McFarlin, 2011).

Taken together, these results suggest that students who do not demonstrate college readiness on the first TASP test attempt spend a substantial portion of the first semester in college taking remedial courses that do not count towards an academic degree. The costs of being in remediation in terms of time and tuition are tangible and directly felt by students, and for this reason the primary focus of this paper is on estimating the effect of being assigned to remediation in at least one subject relative to avoiding placement in remediation altogether. However, since the TASP test consists of multiple sections, there are different combinations of subjects which can be passed or failed, and each of these represents potentially interesting "treatments" (Papay et al., 2011a; 2011b). In particular, students who barely fail one subject but do well in the other two might not feel compelled to alter their college enrollment plans, whereas a negative effect on enrollment might be stronger for a student who does poorly in two subjects relative to only doing poorly in one (or none). Similarly, the signals students receive from performance on the TASP about the likelihood of succeeding in college could vary by the subjects that are passed and failed (e.g., failing math might convey more information than failing reading). We return to this issue below in Section 4.3.

3. Data and Summary Statistics

[12] When students who do not enroll in college are excluded from Figure 1, the discontinuity in the probability of entering remediation at the placement exam passing cutoff rises to about 57 percentage points. Martorell and McFarlin (2011) find no evidence that failing the placement exam affects short-term retention in college.

Data for this study comes from the Texas Schools Project (TSP), a collection of administrative records from state agencies in Texas. We make use of files from the Texas Education Agency (TEA) that oversees K-12 public schools and the Texas Higher Education Coordinating Board (THECB), which oversees the state's public colleges and universities. We also collected data on all TASP test score records from the THECB's testing contractor which made it possible to identify students who took the placement exam and when they did so for the first time.

We draw our sample from the high school graduation records in the TEA files, which contain basic demographic information as well as the date a student graduated. We use students who graduated between 1998 and 2001 because during this time period students were required to take the placement exam prior to enrolling in college, and because the TEA data we use in this study does not include students who graduated after 2001. To the high school graduation records, we merged information on TASP test scores and college enrollment outcomes through 2005 from the THECB. To observe enrollment outcomes over a sufficiently long follow-up period, we also restrict the sample to students who took that TASP test by 2001.[13]

We make a number of sample restrictions before conducting the analysis. First, we exclude students who took the TASP after they enrolled in college since by definition the results of the test could not affect their college enrollment behavior. Similarly, we exclude students who took the TASP so soon before they enrolled that they would not have received their scores in time for the results of the placement test to affect their enrollment decision. In particular, we exclude students who took the TASP test after they enrolled in college, or less than 14 days prior to enrolling in college (the 14 day period corresponds to the time between taking the TASP test and receiving the results (Boylan, 1996)).[14]

---

[13] An implication of this restriction is that the share of students from earlier high school graduation cohorts is higher than the share from later cohorts (see Table 1).

[14] We assume a student starts college on August 15th if they enter in the fall and January 1st if they start in the spring. There is some variation in when the academic term begins across schools and over time. We made inquiries with the registrar's office at several institutions and it seemed that the start dates we assume provide a reasonable approximation for true start dates.

Students who did not enroll in college are included irrespective of when they took the TASP test.[15] In order to test the sensitivity of our results to the choice of the buffer used to make this sample restriction, we also report results where we exclude students who took the TASP test 90 or more days prior to enrolling in college.

Second, we exclude students who do not have valid placement exam scores, since our identification strategy uses these scores as the running variable in the regression discontinuity analysis. Since we do not observe scores from tests other than the TASP, and since the TASP consists of three sections, we limit the sample to those who have non-missing scores on all sections on the first attempt of the TASP.[16] After making this restriction, our base sample has 93,537 observations, and the 90-day buffer sample has 32,909 observations. Many of our analyses also use the subset of students who pass the writing test (about 71 percent of the sample). As we describe in greater detail in the next section, this restriction is made because the aforementioned "lumpiness" in the writing score and its lack of suitability as a running variable in a regression discontinuity analysis. Among students who passed the writing test, overall passing status (i.e. passing all three sections of the test) is determined solely by performance on the math and reading sections.

When considering the generalizability of our findings, it helps to consider the types of students excluded from the analysis due to these sample restrictions. In particular we exclude those who (1) did not take a placement exam (either due to having no interest in college or by virtue of being exempt from the TASP testing requirement due to scoring well on the high school exit exam or the SAT or ACT), (2) took the TASP after enrolling in college (or concurrently with college enrollment), (3) took a placement exam other than the TASP test. The first two types of restrictions exclude students whose college-going behavior would not be affected by the placement exam requirement and who therefore have little policy

[15] An alternative sample restriction along these lines would be to exclude all students, including those who did not enroll in college, who either enrolled in college prior to taking the TASP test as well as any student who took the TASP test 14 days prior to the start of the fall or spring semesters. We obtain very similar results when making this alternative sample restriction.

[16] As we describe below in Section 4, we also exclude a small number (less than 2 percent of observations) with valid scores that are at the extreme end of the test score distribution.

relevance for the research questions we address in this paper. The third restriction is made because we do not observe scores for placement exams other than the TASP test. However, 70 percent of students take the TASP test when first taking a placement exam (THECB, 2000), and this mitigates the impact on external validity of restricting the analysis to TASP test takers.

Our main analysis focuses on whether the decision to enroll in higher education is affected by passing status for the TASP placement exam. Students are coded as enrolled in college if they enrolled in any public institutions in Texas within four years of their first TASP test attempt (to allow an equally long follow-up for all students in the sample), registered for one or more Coordinating Board approved courses, and did not withdraw prior to or on the official census date during the reporting period.

An important limitation of our data is that it only contains information on enrollment at public colleges in Texas. Since the TASP remediation requirements only apply to public institutions in Texas, it may be that a consequence of failing the placement test is that students go to school outside of the Texas public postsecondary system. This type of behavior would lead to an upward bias in the estimated effect of failing the placement test on college enrollment. However, since our estimates of the enrollment effect are small, it is likely that any such biases are inconsequential. Moreover, enrolling in college outside of the Texas public postsecondary system is relatively rare for lower-achieving Texas high school graduates (Kain and O'Brien, 2000). Therefore, having data only on public in-state postsecondary institutions is unlikely to impart substantial bias to our estimates.

Table 1 shows descriptive statistics. The first column shows results for the full sample and the remaining columns show results for our primary analysis sample (i.e., for the subset of students who passed the writing section). The results in columns 3-5 indicate that failing at least one section of the TASP test is much more common among black and Hispanic students, as well as among students from economically disadvantaged backgrounds. Overall, women comprise more than half of the sample, which is consistent with higher school graduation and college-going rates among women (Goldin et al., 2006). However, conditional on being in the sample, men are more likely to pass all sections of the test.

Because our sample consists of students who took the TASP (which is used only for placement purposes), college enrollment rates in our sample are high, with about 93 percent of students enrolling in either a two-year or four-year college. However, enrollment rates are nearly 4 percentage points lower for students who fail the TASP. In the 90-day buffer sample, the gap in enrollment rates is more than 7 percentage points. While these differences are unlikely to be unbiased estimates of causal effect of TASP performance, they do suggest that there is scope for performance on the TASP to affect enrollment behavior.

4. Econometric Strategy

4.1 Research Design

The sample means in Table 1 suggest that students who score above and below the remedial placement cutoffs differ in terms of baseline covariates like gender, race and by construction, test scores. Thus, simple comparisons by placement exam passing status, or even comparisons adjusting for observed covariates are unlikely to yield unbiased causal estimates. Our regression discontinuity (RD) approach exploits the fact that students on either side of the remedial education placement threshold differ in whether they are required to enter remediation if they go to college, but should be similar in other dimensions.

Formally, the model we estimate has the form:

(1) $Y = \theta P + f(S) + \varepsilon,$

where Y is a measure of enrollment, P is an indicator variable for whether a student scored high enough to avoid having to go into remediation, $f(S)$ is a flexible function of the placement test score, S, and ε is the residual. The parameter θ measures the effect of passing the test on enrollment. It reflects the reduced-form effect of passing that encompasses the effects operating through the potential mechanisms discussed above in Section 2.2, although we cannot determine which of these mechanisms is most important. For expositional purposes, the effect of interest, θ, is assumed to be constant; we return to the issue of heterogeneity in these effects below in Section 4.4.

The key assumption underlying this approach is that unobservable determinants of enrollment are similar for individuals on either side of the remedial placement cutoff, which can be expressed as $\mathrm{Cov}(P,\varepsilon)=0$. This condition will be met so long as students do not control their exact test score, and within a narrow region around the cut score, whether a student passes or fails is determined by randomness in the test score (Lee, 2008). If this condition holds, then a comparison of students who barely pass and fail the placement test will identify the effect of passing status on enrollment.

In all analyses, we use running variables that are based on scores from the first TASP attempt. This is done because retaking the TASP is "endogenous". In particular, students who pass the test have no reason to retake it. Thus, any regression discontinuity analysis that used, for instance, the highest TASP score as the assignment variable would potentially suffer from retesting bias. The drawback to using the score from the first attempt is that it may not be the score that actually determines assignment to remediation if students retake the test prior to entering remediation. However, the outcome of the initial attempt has a very strong effect on the probability of entering remediation (as seen in Figure 1). Furthermore, even if students can retake the test, the information conveyed in the first-attempt score might be enough to discourage college-going.

4.2 Estimation

The current context deviates from the conventional RD setup in that the covariate used to determine remedial placements is multi-dimensional since the TASP consists of three sections. Our main approach to dealing with this issue is to transform the multi-dimensional running variable into a one-dimensional variable. In particular, since our primary research aim is to estimate the effect of being placed into remediation in at least one subject relative to avoiding remediation entirely, and since students have to go into remediation if they fail at least one subject, the student's *lowest* score can be used as a conventional, one-dimensional, running variable in an RD analysis. This approach, referred to in the recent literature on multi-dimensional RD as the "binding-score RD" by Reardon and Robinson (2012) and the "centering approach" by Wong et al. (2011), identifies the average treatment effect of being required to enter remediation among students at the passing cutoff in at least one subject and whose scores in other subjects are at or above the passing cutoff (Wong et al., 2011).

This approach cannot be implemented using all three subject scores since, as explained above, the writing scores are assigned in such a way that standard RD assumptions are unlikely to hold.[17] We address this issue in several ways. First, we limit the sample to those students who passed the writing section of the TASP (about 71 percent of the sample) and use $S=\min(M,R)$ as the assignment variable, where M and R are the math and reading scores, respectively, that are both re-centered to zero at the passing cutoff. Defining the running variable and sample in this way, students are assigned to remediation if and only if $S<0$. This is because $S<0$ for students who fail either math or reading; $S\geq0$ for students who pass both math and reading, and all three sections when limiting the sample to students who pass the writing section.

Second, we estimate models where we include students who fail the writing section, and use "fuzzy" RD methods (Hahn, Todd and van der Klaauw, 2001) to account for students with $S\geq0$ being assigned to remediation because of failing the writing section. Specifically, we use a dummy variable for passing both math and reading (i.e., it takes the value 1 for $S\geq0$) as an instrumental variable for passing all three sections. The advantage of this approach is that it does not require limiting the sample to higher-ability students who pass the writing section. However, it is important to recognize that the estimates from this approach are "local" in the sense that they pertain to the subset of students for whom overall TASP passing status is determined by whether or not they pass both the math and reading subjects irrespective of their performance on the writing section (Angrist and Imbens, 1994). Consequently, these estimates are still most pertinent to those who pass the writing section because students who fail the writing test will be assigned to remediation irrespective of their performance on the math and reading sections. We also estimate a variant of this model that uses two dummy variables,

[17] Martorell and McFarlin (2011) discuss evidence inconsistent with the assumptions necessary for a valid RD research design when the writing score is used as the assignment variable.

passing the math and passing the reading section, as instrumental variables for passing all three sections, and includes separate control functions (i.e., the function *f(S)*) for the math and reading scores.

Since the effects of TASP performance could differ depending on whether one fails the math or reading section, we also estimate models that distinguish between the effects of passing the math and reading sections (holding passing status on the other subject constant). Specifically, we estimate a model that has the following form:

(2) $Y = \theta_M P_M + \theta_R P_R + g_M(M) + g_R(R) + v$

where P_M and P_R are dummy variables for passing the math and reading sections, respectively. In contrast to the earlier models that focused on the effect of overall passing status and did not differentiate between whether a student had difficulty passing the math or the reading test, this model allows for differential effects. Note that because students could pass math (or reading) and still be assigned to remediation because of poor performance on the other two subjects, the estimates from this model cannot be interpreted as the effect of assignment to remediation, but instead have an "intent to treat" interpretation (where the "treatment" is defined as assignment to remediation for at least one subject).

For all models, we follow the recommendations in Imbens and Lemieux (2008) and use two approaches for estimation. First, we use the full range of test scores and control for a "global polynomial" in the test score.[18] Because we do not know a priori what the correct functional form *f(S)* is, we report estimates using a quadratic and a cubic polynomial (where the coefficients on the polynomial terms are all allowed to be different on either side of the cut score). Second, we estimate a local linear regression where the bandwidth around the cut score is determined on the basis of the cross-validation procedure proposed in Imbens and Lemieux (2008). As shown below, the results are not very sensitive

[18] We trim observations with extreme test score outliers (roughly the top 0.3 percent and the bottom 1.2 percent) even for the global polynomial specifications because the regression fit in this range is poor and we do not want extreme test score values affecting the estimated discontinuities at the remedial placement cutoffs.

to these modeling choices. Due to the discrete nature of the TASP test scores, we follow the suggestion of Lee and Card (2008) and adjust the standard errors for clustering at the test score cell level.[19]

4.3 Estimating the Effect of Passing and Failing Different Combinations of Subjects

As noted in Section 2.3, there are several potentially interesting "treatments" corresponding to different combinations of passed and failed subjects. However, the models described in the preceding section only capture the effect of being assigned to remediation in at least one subject relative to none.[20] We address this issue in two ways. First, we implement the multi-dimensional RD estimator proposed by Papay et al. (2011a). The main strength of this approach is that it allows one to estimate multiple treatment effects simultaneously, such as the effect of failing both math and reading versus passing both subjects, or the effect of passing versus failing reading for students who passed math. However, as noted in Papay et al. (2011b), the procedure is very "data intensive" in that it requires sufficient sample size at several discontinuity points, and thus often produces imprecise estimates, and relies on strong functional form assumptions.[21] As an alternative to the Papay et al. (2011a) estimator, we also use S^{max}=max(M,R) as the running variable and a dummy for $S^{max}\geq 0$ as the "treatment". This approach examines the effect of failing both math and reading relative to passing at least one of these subjects, and allows us to investigate whether failing an incremental section (e.g., the effect of failing two sections versus just one section) affects enrollment decisions. In contrast to the Papay et al. (2011a) estimator, this approach is less data intensive (and produces estimates with smaller standard errors), however it answers fewer

---

[19] For models that use both the math and reading scores separately in the models, we use the two-way cluster adjustment procedure proposed by Cameron, Gelbach and Miller (2011).

[20] The model described in Equation (2) captures the effect of being assigned to remediation due to failing math or reading, but these effects do not depend on performance on other subjects.

[21] The specification of the model is given by the equation:

$$Y = \beta_0 + \beta_1 P_M + \beta_2 P_R + \beta_3 P_M * P_R + \beta_4 M + \beta_5 R + \beta_6 M * R + \beta_7 P_M * M + \beta_8 P_R * R + \beta_9 P_M * R + \beta_{10} P_R * M + \beta_{11} M * R * P_M$$
$$+ \beta_{12} M * R * P_R + \beta_{13} M * P_M * P_R + \beta_{14} R * P_M * P_R + \beta_{15} M * R * P_M * P_R + v$$

We use this model to look at five treatment effects evaluated at the passing cutoff (i.e., M=0 and R=0): the effect of failing both math and reading relative to passing both ($\beta_1+\beta_2+\beta_3$); the effect of passing vs. failing reading conditional on failing math (β_2) and conditional on passing math ($\beta_2+\beta_3$); and the effect of passing vs. failing math conditional on failing reading (β_1) and conditional on passing reading ($\beta_1+\beta_3$). The model assumes that the relationship between the outcome and the running variables is linear close to the passing cutoffs. We use the cross-validation procedure suggested in Papay et al. (2011a) to determine the bandwidth, but even within the bandwidth the model may be misspecified, and unlike the one-dimensional RD models, it is not possible to assess graphically how well the parametric model tracks the underlying data.

questions of interest (for instance, it does not produce estimates of the effect of failing both subjects versus failing zero subjects). Both of these approaches are implemented using the whole sample, as well as when stratifying by passing status to see if the effects of the TASP performance differ depending on whether one passes or fails the writing section.

4.4 Heterogeneous Effects

The preceding discussion assumed that a given treatment effect associated with placement exam performance is the same for all students. However, this is unlikely to be true in practice. Because our research design focuses on comparisons between students at the boundary for placement into remediation, our estimates are most informative about the effects for students whose initial score is close to the remedial placement cutoff. One reason is that students who score just under the passing cutoff might be more likely to pass the test on a retake and avoid remediation. However, as discussed above, barely failing the TASP does exert a very strong positive effect on the likelihood of entering remediation. Another reason is that even when students at the margin we examine (i.e., the margin between no remediation and any remediation) are placed into remediation, they are generally placed into the highest remedial tier. Thus, the informational "shock" regarding their lack of academic preparation, as well as any associated stigma, may be relatively less severe than it would be for students assigned to the lowest remedial tiers.

Although our research design does not allow us to identify the effects away from the remediation placement threshold, the effects our approach does capture have policy relevance for at least three reasons. First, many students in our data score relatively close to the remediation placement threshold; in our primary analysis sample, 25 percent of students score within one-third of a standard deviation of the passing cutoff. Second, the effects we estimate are relevant for thinking about the likely impact of a modest change in the remediation placement threshold. In fact, some college administrators in Texas expressed concern that the passing standard necessary to avoid remediation might reduce enrollment in Texas community colleges (Stewart, 1995). Third, policymakers presumably believed that remediation would be beneficial for students who scored just below the remediation placement threshold they set (Hardesty and Matthews, 1991), and our results address the question of whether or not there were unintended effects on enrollment for this group that was assigned to remediation.

5. Results

5.1 Testing the Validity of the Design

The identification assumptions required for the validity of the RD design have the testable implication that the distribution of the "running variable" (i.e., the minimum of the reading and math scores) does not exhibit any discontinuous behavior at the cutoff score. Appendix Figure 1 shows the histogram of S, and visual inspection reveals no evidence that students sort around the cutoff score in a manner that generates a discontinuous density at the cutoff score. Moreover, estimates of the discontinuity in the number of observations are not statistically significant, providing further support in favor of the smoothness of the distribution of the running variable.[22]

Another testable implication is that the baseline covariates should trend "smoothly" through the cutoff score. Examining whether this pattern holds is akin to testing whether baseline covariates are balanced between treatment and control groups in a randomized-control trial (Imbens and Lemieux, 2008; Lee, 2008). To implement this test, Table 2 reports estimated discontinuities at the cutoff score for the baseline variables used in the analysis. Columns 1-3 show estimates for the 14-day buffer sample and columns 4-6 show estimates for the 90-day buffer sample. Overall, the estimated discontinuities are small in magnitude and almost always statistically insignificant. One exception is for the maximum of the math and reading scores where barely passers have somewhat *lower* scores than barely failers in the 14-day

[22] Given how the scale scores are assigned, there is a "spike" in the distribution at S=-2 and zero observations at S=-1 (there is only one other whole number value, -11, between -80 and 50 with no observations). Note that this is not evidence of sorting around the cutoff score because both S=-2 and S=-1 are below the cut score. Thus, Appendix Figure 1 replaces the cell sizes at S=-1 and S=-2 with half of the S=-2 cell size and assigns this to S=-1.5. After making this transformation, the estimated discontinuity in the number of observations at the cut score (obtained by regressing the cell size on a cubic polynomial in the test score separately on either side of the passing cutoff, and weighting by cell size) is -313 with a standard error of 282. The discontinuity in the unadjusted cell size is also statistically insignificant when each observation is weighted equally. Because of this data "heaping" pattern just to the left of the cut score, we employed the "donut RD" estimator proposed by Barreca et al. (2010) in which we excluded observations at S=-2, S=-1 and obtained similar results to what we report here (similar results were obtained when we also excluded observations at S=0).

buffer sample. However, the point estimate is fairly small (about 1.5 compared to a standard deviation of 23), and the estimate is smaller and statistically insignificant in the 90-day buffer sample. Furthermore, the graphical evidence in Appendix Figure 2 shows little indication that the maximum test score behaves discontinuously at the cutoff, and when we control for baseline covariates, the main qualitative and quantitative results are not affected.

5.2 Main Results

To examine whether assignment to remediation affects the likelihood of enrolling in college, Figure 2 plots the fraction of students enrolled in college as a function of the minimum of re-centered math and reading scores. The open circles represent test score cell means and the curve is the regression fit from a regression of college enrollment on a cubic polynomial in the running variable (estimated separately on either side of the cutoff score).[23] The upper panel shows results for the 14-day buffer sample, and the lower panel shows results for the 90-day buffer sample. For both samples, the fraction of students enrolled in college increases in TASP performance (except for the very highest scores where there are relatively few observations). However, the likelihood of enrolling in college is very similar for students just above and below the cut score.

The estimates in Table 3 also suggest that failing the TASP has little effect on the likelihood of enrolling in college. For the 14-day sample (upper panel), the estimates in the top row that correspond directly to the results in Figure 2 are all very small (0.5 percentage points or less) and statistically insignificant. The point estimates are also very similar across the local linear, quadratic, and cubic polynomial specifications, which suggest that the choice of regression specification is not driving the results. Consistent with a valid regression discontinuity design, estimates in columns 1-3 are also very similar to the estimates from models that include baseline covariates (reported in columns 4-6).

---

[23] We use the results from a cubic polynomial specification in the figures since this specification is most flexible and least susceptible to bias from imposing an incorrect parametric specification. However, it is very important to look at the estimates from alternative specifications to see if the results are "fragile" or "robust" (i.e., if they hold up to alternative reasonable specifications or if they only appear in some of the specifications). When the results are not robust across parametric specifications, we have relatively less confidence that the results are not being driven by functional form assumptions.

The estimates from the fuzzy RD models that include students who failed the writing section of the TASP are reported in the next two sets of rows. For the model that uses a dummy for passing both math and reading as a single instrumental variable, the estimates are quite similar to the estimates from the sharp RD models. The effects found in models that use separate instrumental variables for passing math and passing reading are positive but fairly small in magnitude and statistically significant in only one out of six specifications. Finally, results from the model that allow the effect of failing math and reading to differ are reported in the last two sets of rows. These estimates are generally consistent with the results from models that focus on the effect of failing any section. For both subjects, there is no indication that passing this section affects the likelihood of enrolling. Moreover, we can never reject the hypothesis that the effect of passing math and the effect of passing reading are the same.

The estimates from the 90-day buffer sample, reported in the lower panel of Table 3, are broadly consistent with the results for the 14-day buffer sample. In particular, after excluding students who enroll in college within 2-3 months after taking the TASP test, we find no evidence that TASP performance affects the likelihood of enrolling in college. If anything, the estimates suggest that passing the TASP test reduces the likelihood of enrolling in college, although none of the negative estimates is statistically significant. We also examine whether effects differ for early or late TASP test takers for students who took the TASP 14-45 days prior to enrolling in college or 46-90 days prior to enrolling in college. The impetus for this analysis is that late test takers might be less committed to college-going, which could manifest larger discouragement effects. However, the results in Table 4 provide little evidence supporting this hypothesis.

An important consideration is whether our estimates are statistically insignificant only because we do not have enough statistical power. For the 14-day buffer sample, the upper bound of the confidence interval for the most precise estimates (those from the sharp RD models that restrict the sample to students who passed the writing test and that use the local linear and quadratic specifications) is about 1.3 percentage points. The estimates for the 90-day buffer sample are less precisely estimated, but again the most precise estimates (from the sharp RD models, columns 4 and 5) allow us to rule out effects of about 2.5 to 3 percentage points. Given that the baseline enrollment rate in the 14-day buffer sample is 93 percent and the raw difference in enrollment rates between students who passed the TASP and those who failed is 4 percentage points (80 percent and 7 percentage points, respectively, for the 90-day buffer sample), the results in Table 3 suggest that our most precise estimates allow us to rule out effects half as large as those suggested by naïve comparisons of means.

5.3 Subgroup Results

Although we found little evidence that TASP performance affected the likelihood of college enrollment, it may be that there are effects for certain subgroups of students. To examine this possibility we produced estimates separately by three classes of subgroups. First, we examined effects by race, since black and Hispanic students are more likely to be first-generation college students, and may be more susceptible to being dissuaded from going to college by poor performance on the TASP test. Second, we examined effects by gender since there is evidence that the impact of educational interventions such as assignment to academic probation have differential effects for men and women (Lindo et al., 2010). Finally, we produced estimates by a measure of economic disadvantage since the additional costs associated with being assigned to remediation could have larger negative effects for low-income students. Table 5 shows estimates by these subgroups from the sharp RD models adjusted for baseline covariates. For the most part, we find little evidence of heterogeneous effects on enrollment. For blacks, in some specifications we find *negative* effects of passing the TASP that are statistically significant in two of the six specifications (and large in magnitude for some of the 90-day sample results). For Hispanics, one estimate (cubic polynomial for the 14-day buffer sample) suggests passing increases enrollment by 2.5 percentage points and is statistically significant. However, finding some statistically significant estimates for these subgroup analyses may be due to the large number of estimates reported for the 7 subgroups and 6 specifications.

5.4 Estimates of the Effect of Passing and Failing Different Combinations of Subjects

The results in the preceding section all focus on the effect of the treatment of being assigned to some remediation relative to none, but there are other treatment effects that are interesting. Table 6 shows estimates of several treatment effects obtained by the multi-dimensional RD model proposed by Papay et al. (2011a). It is difficult to discern clear patterns because the estimates are fairly noisy. The point estimates suggest that passing both math and reading relative to failing both, and passing relative to failing math conditional on failing reading (or passing relative to failing reading conditional on failing math) may have a positive effect on enrollment. However, the estimates are imprecisely estimated and only one is statistically significant. Moreover, some estimates are negative although none is statistically significant. There is also not a clear pattern in the estimates with respect to passing status on the writing test.

Next, we considered whether failing two sections relative to passing at least one (as opposed to failing at least one section relative to passing all sections) affects enrollment. To examine the effect of failing both math and reading relative to passing at least one, we use the maximum of math and reading as the running variable. As seen in Figure 3, in the sample restricted to students who passed the writing test, the likelihood of enrollment increases in the maximum score, but there is no discernible discontinuity in the pattern at 0 for either the 14-day or 90-day buffer samples. This is consistent with the small and statistically insignificant estimates in the second row of the upper panel of Table 7. For the 90-day buffer sample, all of the estimates are statistically insignificant, and the sign switches between positive and negative. For the 14-day buffer sample, the estimate from the quadratic polynomial specification is statistically significant for the full sample and the sample conditioned on failing writing, but the estimates for the local linear and cubic polynomial specifications are smaller and statistically insignificant. We also examined whether the results for the models using the minimum score as the running variable (i.e., the specification used to examine the effect of passing both math and reading relative to failing at least one) differ by a student's performance on the writing test. These estimates are all small and statistically insignificant.

Overall, these results do not contradict the main result from the preceding sections that TASP performance does not affect college enrollment decisions. Some of the estimates suggest that there may be a cumulative effect of failing multiple sections that could dampen college enrollment, but we do not find robust and statistically significant evidence of such an effect.

6. Discussion and Conclusion

During our study period, Texas students were required to take a placement test prior to enrolling in college-level courses, and a non-trivial share of students who took the exam chose not to enroll in college. This paper examined whether failing a college placement exam – the primary way students are assigned to college remediation – acts to discourage college-going. Such an effect might arise because assignment to remediation significantly alters the college experience by increasing costs, raising individual awareness of academic underpreparedness, and potentially exposing students to "social stigma." As some policymakers suggest, poorly prepared students who face these consequences may become discouraged from pursuing postsecondary education.

There are important policy implications for the existence or absence of such discouragement effects on college going. If a societal objective is to have as many students go to college as possible, as President Obama recently suggested, and failing a placement exam reduces college-going, then it would suggest that remediation policies should be redesigned. For example, if tuition and time to degree are crucial factors in dissuading students with weaker academic preparation from attending college, then one prescription would be to make remedial courses (at least partly) count toward a degree. Another possibility would be to give students the option of going into remediation rather than making it mandatory. An alternative view is that policies aimed at maximizing postsecondary enrollment are not desirable since many students lack the skills necessary to complete college (Rosenbaum, 2002). According to this perspective, discouragement effects brought about from failing placement exams might be interpreted as evidence that placement test results correctly inform underprepared students that they may be better off not going to college.

Although we found large differences in college enrollment rates between placement test passers and failers, regression discontinuity estimates offer little indication that failing a placement exam has a negative causal effect on enrollment. In particular, we can rule out effects on enrollment larger than about half of the difference in average enrollment rates between students who pass and fail the exam. The conclusions are robust to several regression specifications and a variety of subgroups such as gender and race.

One interpretation for these findings is that being assigned to remediation does not have large enough financial and psychic costs, and does not provide enough new information about one's ability to succeed in college, to dissuade students from going to college. While our empirical strategy produces only "reduced-form" estimates of the effect of placement test performance, the fact that we find little evidence of negative enrollment effects suggests that none of the hypothesized mechanisms is important enough to generate effects on college enrollment. A second possibility is that students could feel relieved at being placed into remedial classes if they think that is appropriate for their level of preparation, and this may offset stigma/discouragement effects. A third explanation is that students may not be perfectly informed about the costs and benefits of remedial education. In particular, students may be unaware that remedial courses do not count toward a degree. In fact, in a recent survey of community college students (Rosenbaum et al., 2006), three-fourths of remedial students either believed that remedial courses count toward a degree or were unsure. Obviously, this type of misinformation about remedial education would undermine the extent to which failing a placement exam could have a discouraging effect on college-going, and it also suggests that students are making decisions about pursuing postsecondary education based on faulty information. If true, it would imply that colleges could be more transparent about the non-credit bearing nature of remedial coursework so that students are making decisions with complete information.

To be clear, our results do not rule out the possibility that the presence of a mandatory placement testing policy could dissuade students from entering college. That is, knowing that there is a chance that one might have to go into remediation may be enough to induce students to not go to college. Because all of our analysis is based on a period when there was mandatory placement testing, we cannot investigate whether such an effect exists.

Finally, it is also important to bear in mind that these findings are "local" to the students who score close to the remedial placement threshold. We argue that this group is policy relevant because a large share of students scores close to this threshold and also because the effects for this group are the most relevant when thinking about the likely impact that a change in the placement threshold would have. Nonetheless, it may be that performing very poorly on a college placement exam could have larger discouragement effects on enrollment behavior than what is reported here (for instance, being assigned to lower-tier remedial classes may seem more onerous than being placed in the highest remedial tier). An important avenue for future research is to examine whether there are any enrollment effects of being assigned to remediation across a broader range of placement exam scores and levels of academic preparation.

References

Abraham, Ansley. 1986. "College-Level Study: What Is It? Variations in College Placement Tests and Standards in the SREB States." In *Issues in Higher Education*, No. 22. Atlanta, GA: SREB.

Alliance for Excellent Education. 2006. *Paying Double: Inadequate High Schools and Community College Remediation*. Washington, DC: Alliance for Excellent Education.

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2009. "Incentives and Services for College Achievement: Evidence from a Randomized Trial," *American Economic Journal: Applied Economics* 1(1): 136-163.

Arenson, Karen. 2001. "The Pitfalls of Make-or-Break Tests," *The New York Times*. March 7, 2001.

Barreca, Alan and Melanie Guldi, Jason Lindo, and Glen Waddell. 2010. "Running and *Jumping* Variables in Regression Discontinuity Designs." Mimeo.

Becker, Gary. 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 3rd Ed. Chicago, IL: University of Chicago Press.

Bettinger, Eric and Bridget Long. 2009. "Addressing the Needs of Underprepared Students in Higher Education: Does College Remediation Work?" *Journal of Human Resources* 44(3): 736–771.

Bettinger, Eric, Bridget Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2009. "The Role of Simplification and Information in College Decisions: Results from the H&R Block FAFSA Experiment," National Bureau of Economic Research Working Paper No. 15361.

Breneman, David and William Haarlow. 1998. "Remedial Education: Costs and Consequences," *Remediation in Higher Education: A Symposium*. Washington, DC: Thomas B. Fordham Foundation.

Bostian, Brad. 2008. "Avoiding Remedial Education: Academic Effects on College Transfer Students," PhD Dissertation. University of North Carolina at Charlotte. 114 pages. AAT 3320959.

Boylan, Hunter. 1996. *An Evaluation of the Texas Academic Skills Program (TASP)*. Prepared under contract for the Texas Higher Education Coordinating Board.

Cameron, Colin, Jonah Gelbach, and Douglas Miller. 2011. "Robust Inference with Multi-way Clustering," *Journal of Business and Economic Statistics*. 29(2): 238-249.

Calcagno, Juan Carlos and Bridget Long. 2008. "The Impact of Postsecondary Remediation Using a Regression Discontinuity Approach: Addressing Endogenous Sorting and Noncompliance," NBER Working Paper No. 14194.

Clark, Burton. 1960. "The 'Cooling-Out' Function in Higher Education," *The American Journal of Sociology* 65(6): 569-576.

Collins, Michael. 2008. *It's Not About the Cut Score: Redesigning Placement Assessment Policy to Improve Student Success*. Boston, MA: Jobs for the Future.

Dounay, Jennifer. 2006. "Embedding College Readiness Indicators in High School Curriculum and Assessments," Education Commission of the States.

Fletcher, Jason and Mansur Tokmouline. 2010. "The Effects of Academic Probation on College Success: Lending Students a Hand or Kicking Them While They Are Down?" THEOP Working Paper

Goldin, Claudia, Lawrence Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap," *Journal of Economic Perspectives* 20(4): 133-156.

Hardesty, Robert and Joan Matthews. 1991. "Getting Started: The Work of the Committee on Testing," In *From Politics to Policy: A Case Study in Educational Reform*. Joan Matthews, Ronald Swanson, and Richard Kerker. New York, NY: Praeger Publishers.

Hahn, Jinyong, Petra Todd and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69(1): 201-209.

Howell, Jessica, Michal Kurlaender, and Eric Grodsky. 2010. "Postsecondary Preparation and Remediation: Examining the Effect of the Early Assessment Program at California State University," *Journal of Policy Analysis and Management* 29(4): 726-748.

Hughes, Katherine, and Judith Scott-Clayton. 2010. "Assessing Developmental Assessment in Community Colleges: A Review of the Literature," Community College Research Center Working Paper No. 19., Columbia University.

Imbens, Guido and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62(2): 467-475.

Imbens, Guido and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics* 142(2): 615-635.

Jepsen, Christopher, Peter Mueser, and Kenneth Troske. 2010. "Labor Market Returns to the GED Using Regression Discontinuity Analysis," Working Paper No. 1014., Department of Economics, University of Missouri.

Kain, John and Daniel O'Brien. 2000. "High School Outcomes and College Decisions of Texas Public School Students," TSP Working Paper.

Kingan, Mary, and Richard Alfred. 1993. "Entry Assessment in Community Colleges: Tracking or Facilitating?" *Community College Review* 21(3): 3-16.

Lee, David. 2008. "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics* 142(2): 675-697.

Lindo, Jason, Nicholas Sanders, and Philip Oreopoulos. 2010. "Ability, Gender, and Performance Standards: Evidence from Academic Probation," *American Economic Journal: Applied Economics* 2(2).

Long, Bridget and Erin Riley. 2007. "Sending Signals to Students: The Role of Early Placement Testing in Improving Academic Preparation," In *Minding the Gap: Why Integrating High School with College Makes Sense and How to Do It*. Nancy Hoffman, Joel Vargas, Andrea Venezia, and Marc Miller (Eds.), Cambridge, MA: Harvard Education Press.

Manski, Charles. 1989. "Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon," *Economics of Education Review* 8(4): 305-312.

Martorell, Paco and Isaac McFarlin. 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.

Moltz, David. 2009a. "Shopping for 'Cut Scores'," *Inside Higher Ed*. April 14, 2009.

Moltz, David. 2009b. "Black Men and Remedial Education," *Inside Higher Ed*. March 31, 2009.

National Center for Education Statistics. 2003. Remedial Education at Degree-Granting Postsecondary Institutions in Fall 2000, NCES 2004-010. Washington, DC: NCES.

Palmer, Robert and Ryan Davis. 2012. "'Diamond in the Rough:' The Impact of a Remedial Program on College Access and Opportunity for Black Males at an Historically Black Institution," *Journal of College Student Retention: Research, Theory & Practice*. 13(4): 407-430.

Papay, John, John Willett, and Richard Murnane. 2011a. "Extending the Regression-Discontinuity Approach to Multiple Assignment Variables," *Journal of Econometrics* 161(2): 203-207.

Papay, John, John Willett, and Richard Murnane. 2011b. "High School Exit Examinations and the Schooling Decisions of Teenagers: A Multi-Dimensional Regression-Discontinuity Analysis." NBER Working Paper 17112.

Reardon, S. F., & Robinson, J. P. 2012. Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness* 5(1): 83-104.

Rosenbaum, J., Deil-Amen, R. and A. Person. 2006. *After Admission: From College Access to College Success*. Russell Sage Foundation.

Stewart, Linda. 1995. "State Test Cited in Lower College Enrollments," *Dallas Morning News*. January 24, 1995.

Texas Higher Education Coordinating Board (THECB). 1998. *Annual Texas Academic Skills Program Report of Student Performance: Pass Rates by Race/Ethnicity and Test Area 1993-1997 High School Graduating Classes*.

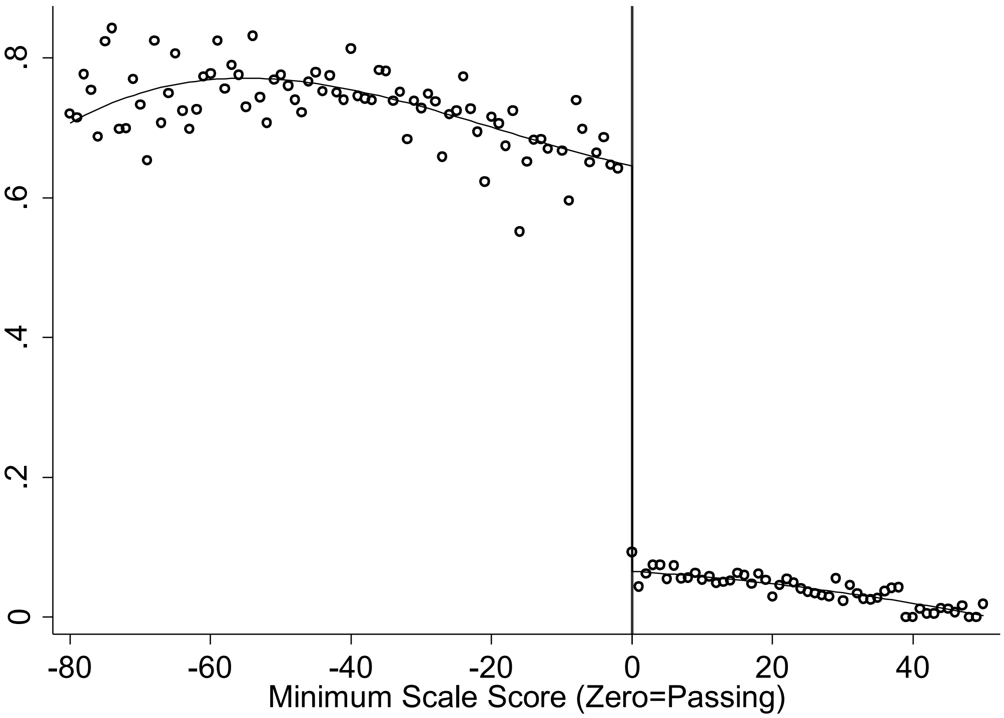
Texas Higher Education Coordinating Board (THECB). 2000. *Texas Academic Skills Program (TASP) Summary of TASP/Alternative Test Results: Academic Year 1998-99*.

Texas Higher Education Coordinating Board (THECB). 2005. *Developmental Education in Texas Higher Education: A Comparison of Policies and Practices*, Part 1

Venezia, Andrea, Michael Kirst, and Anthony Antonio. 2003. *Betraying the College Dream: How Disconnected K-12 and Postsecondary Education Systems Undermine Student Aspirations*. Stanford, CA: Bridge Project, Stanford Institute for Higher Education Research.

Wong, Vivian, Peter Steiner, and Thomas Cook. 2011. "Analyzing Regression-Discontinuity Designs with Multiple Assignment Variables: A Comparative Study of Four Estimation Methods," Institute for Policy Research Working Paper 10-02, Northwestern University.

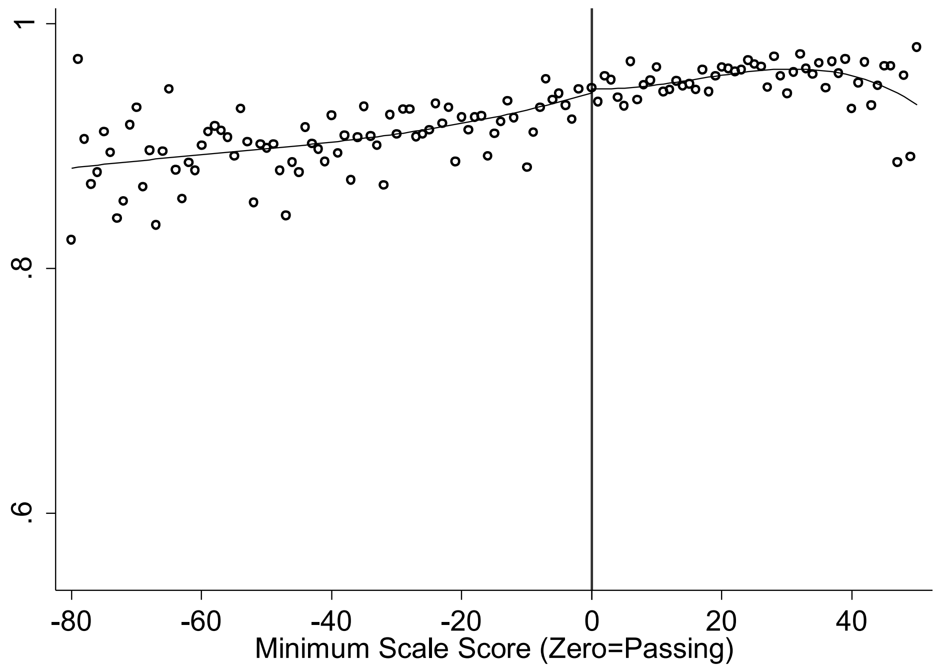
Figure 1: Fraction of Students in Remedial Education by Min(Math, Reading)



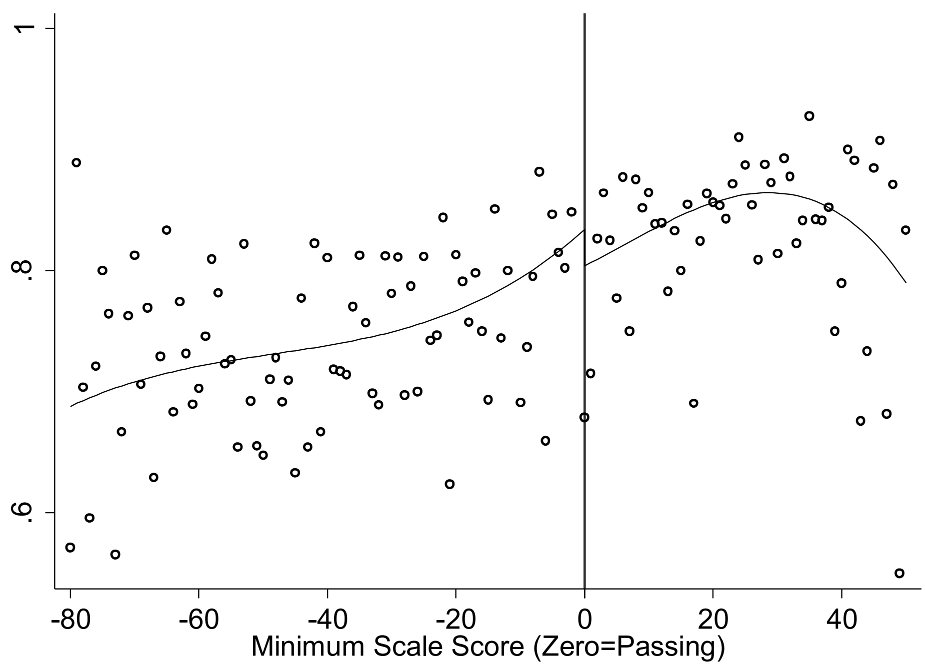
Note: Sample includes students who took the TASP at least 14 days prior to enrolling (or did not enroll) in college, and is limited to students who passed the writing test. The horizontal axis is the minimum of the math and reading scores, where each has been re-centered to be equal to zero at the passing cutoff. Thus cells to the left of zero are for students who failed the TASP test and cells to the right or at zero are for students who passed the test. Open circles represent the fraction of students in a test score cell who go into remedial education; curve represents fitted values from a cubic polynomial (estimated separately on either side of the cut score). See text for additional details.

Figure 2: Fraction Enrolling in College After Taking TASP by Min(Math, Reading)

a) 14-Day Buffer Sample



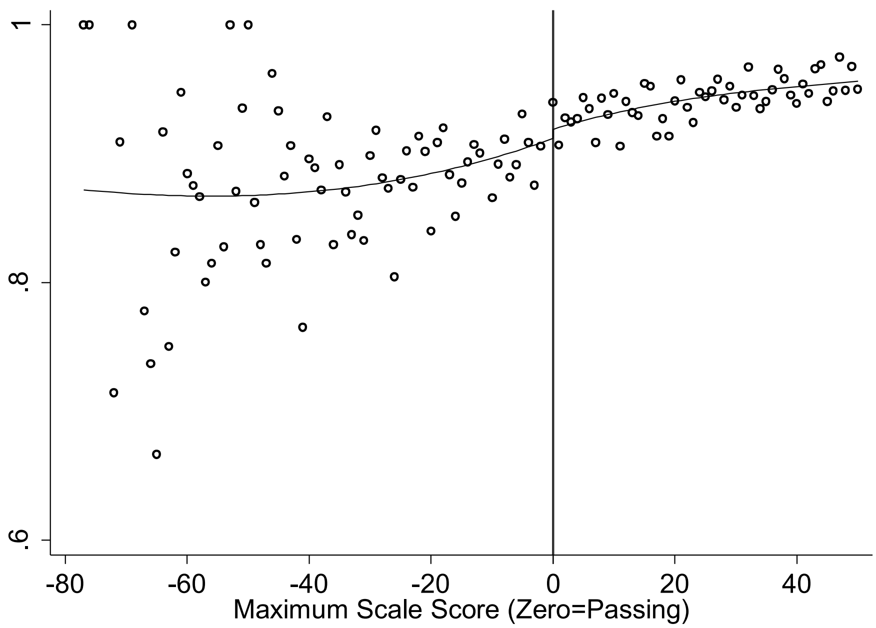
b) 90-Day Buffer Sample



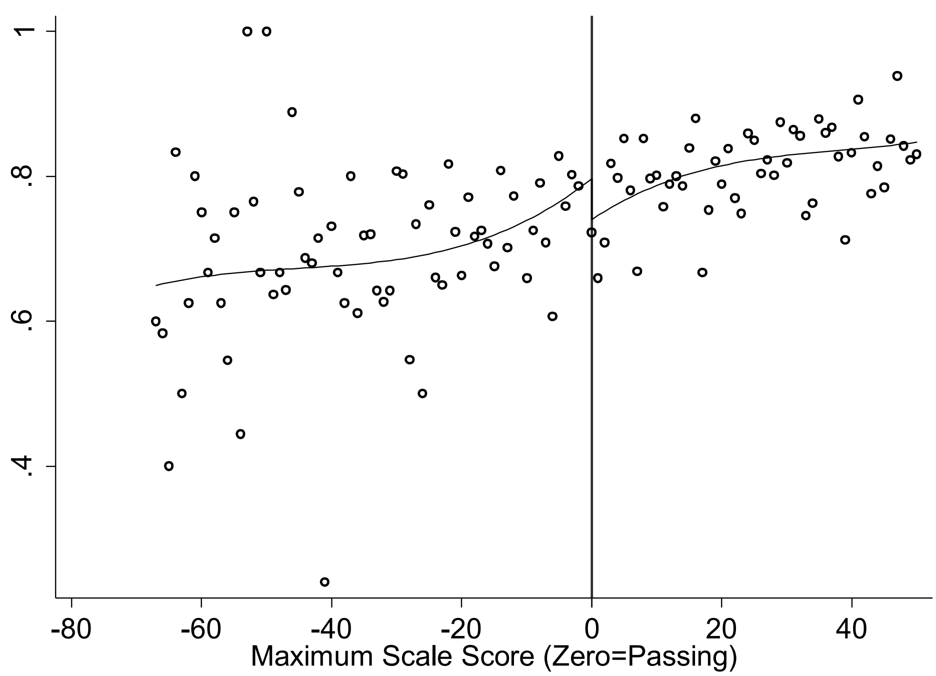
Note: Sample includes students who took the TASP at least 14 or 90 days (panel (a) and (b), respectively) prior to enrolling (or did not enroll) in college, and is limited to students who passed the writing test. Open circles represent scale score cell means; curve represents fitted values from a cubic polynomial (estimated separately on either side of the cut score). See text and notes to Figure 1 for additional details.

Figure 3: Fraction Enrolling in College After Taking TASP by Max(Math, Reading)

(a) 14-Day Buffer Sample

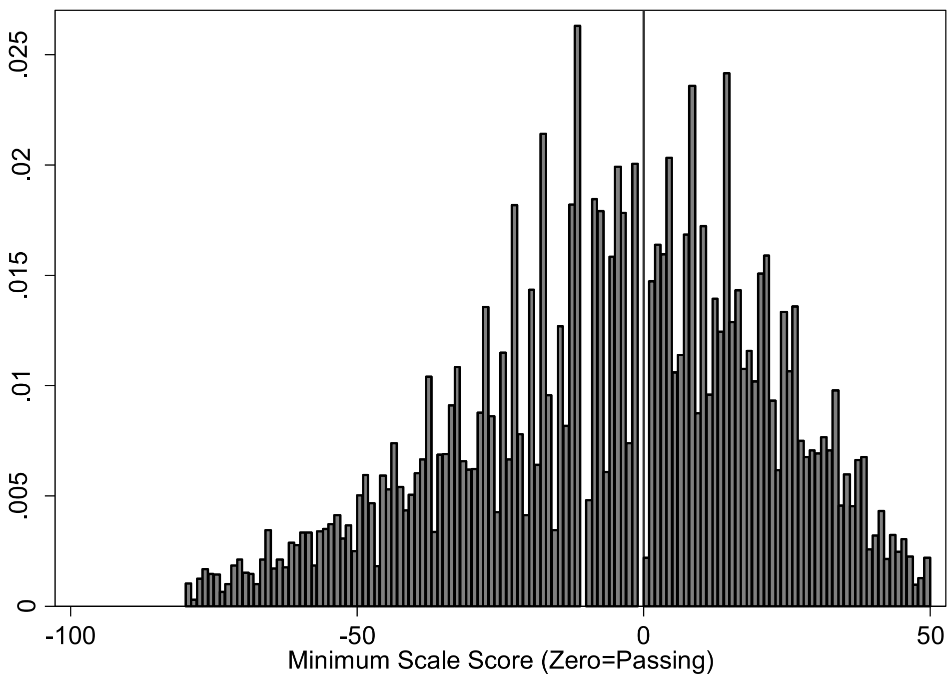


(a) 90-Day Buffer Sample



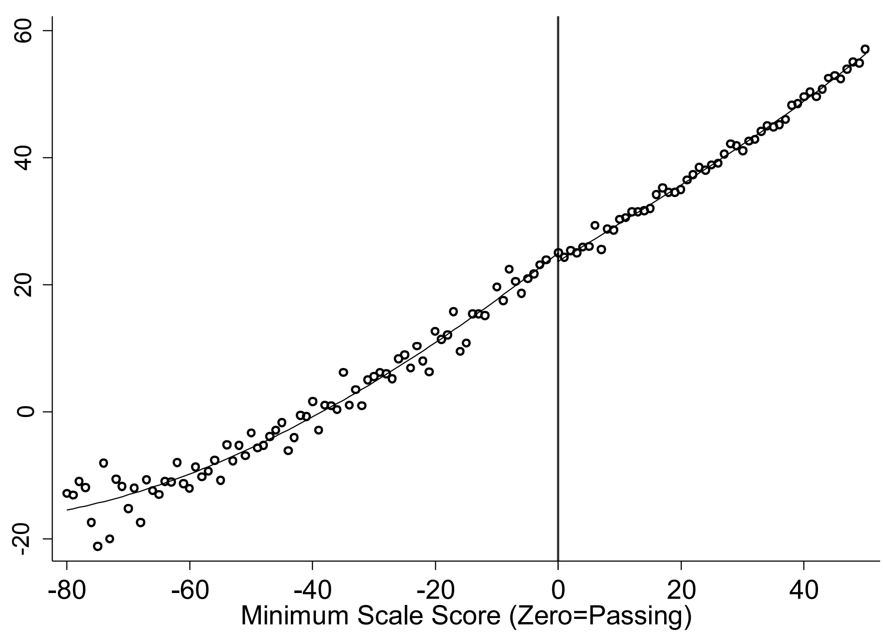
Note: Sample includes students who took the TASP at least 14 days (panel a) or 90 days (panel b) prior to enrolling (or did not enroll) in college, and is limited to students who passed the writing test. Open circles represent scale score cell means; curve represents fitted values from a cubic polynomial (estimated separately on either side of the cut score). See text and notes to Figure 1 for additional details.

Appendix Figure 1: Distribution of Min(Math, Reading) Score



Note: Sample consists of students who took the TASP at least 14 days prior to enrolling in college (or did not enroll) and passed the writing section. This graph "smooths" the density at S=-2 and S=-1 by replacing the values at S=-2 and S=-1 with half the density at S=-2 and assigning a value of S=-1.5. This is because in the unadjusted data, there is a large spike at *S*=-2 and no observations at S=-1. See text for additional details.

Appendix Figure 2: Average Max(Math,Reading) Score by Min(Math, Reading) Score



Note: Sample consists of students who took the TASP at least 14 days prior to enrolling in college (or did not enroll) and passed the writing section. Open circles represent scale score cell means; curve represents fitted values from a cubic polynomial (estimated separately on either side of the cut score). See text for additional details.

Table 1: Sample Means

	Full Sample	Sample that Passed the Writing Section			
		All	Pass all sections	Fail at least one section	Pass - Fail
Demographics					
Male	0.460	0.416	0.470	0.371	0.098**
White	0.529	0.590	0.662	0.530	0.131**
Hispanic	0.279	0.248	0.208	0.281	-0.073**
Black	0.160	0.135	0.093	0.170	-0.077**
Economically Disadvantaged	0.232	0.186	0.142	0.223	-0.080**
High School Graduate in 1998	0.376	0.396	0.389	0.402	-0.013
High School Graduate in 1999	0.277	0.270	0.271	0.269	0.003
High School Graduate in 2000	0.195	0.191	0.191	0.191	0.000
High School Graduate in 2001	0.152	0.143	0.148	0.138	0.010
Test Scores					
Max (Math, Reading)	13.667	20.376	34.511	8.481	26.030**
	(26.976)	(23.391)	(13.746)	(23.215)	(2.002)
Math Scale Score	-4.217	1.459	23.030	-16.692	39.722**
	(30.034)	(28.086)	(14.843)	(23.370)	(2.877)
Writing Scale Score	-3.858	22.300	23.498	21.291	2.208**
	(57.342)	(13.302)	(15.255)	(11.304)	(0.229)
Reading Scale Score	5.396	13.166	29.140	-0.276	29.416**
	(31.087)	(27.427)	(15.246)	(28.181)	(2.587)
% Pass Writing	0.714				
Outcomes					
Enroll in College (14-day buffer sample)	0.912	0.934	0.954	0.917	0.037**
Enroll in College (90-day buffer sample)	0.749	0.796	0.841	0.766	0.074**
N (14-day buffer sample)	93537	66831	30539	36292	
N (90-day buffer sample)	32909	21656	8768	12888	

Notes: Sample consists of students in the 1998-2001 high school graduating classes who took the TASP test at least 14 days prior to enrolling (or did not enroll) in college. Standard deviations are reported in parentheses. *Denotes statistically significant difference between TASP passers and failers at the 5% level. ** denotes significant difference at the 1% level.

Table 2: Estimated Discontinuities in Baseline Covariates

	14-Day Buffer Sample			90-Day Buffer Sample		
Male	0.007	-0.000	0.005	0.002	-0.016	-0.002
	(0.011)	(0.011)	(0.014)	(0.015)	(0.015)	(0.021)
White	0.003	-0.006	-0.012	0.012	0.000	0.003
	(0.015)	(0.015)	(0.020)	(0.019)	(0.019)	(0.025)
Hispanic	0.000	0.010	0.016	-0.006	0.009	0.007
	(0.009)	(0.008)	(0.011)	(0.015)	(0.015)	(0.021)
Black	-0.008	-0.008	-0.010	-0.008	-0.008	-0.005
	(0.006)	(0.007)	(0.009)	(0.013)	(0.014)	(0.019)
Economically Disadvantaged	0.006	0.011	0.014	0.021	0.029	0.034
	(0.009)	(0.009)	(0.012)	(0.016)	(0.018)	(0.025)
High School Graduate in 1998	-0.007	0.018	-0.002	0.019	-0.001	-0.026
	(0.053)	(0.050)	(0.059)	(0.098)	(0.101)	(0.129)
High School Graduate in 1999	-0.027	-0.030	-0.046	-0.046	-0.038	-0.066
	(0.064)	(0.060)	(0.071)	(0.109)	(0.112)	(0.142)
High School Graduate in 2000	0.067	0.046	0.084	0.038	0.053	0.089
	(0.053)	(0.050)	(0.063)	(0.040)	(0.042)	(0.057)
High School Graduate in 2001	-0.033	-0.035	-0.037	-0.011	-0.014	0.003
	(0.035)	(0.032)	(0.036)	(0.019)	(0.019)	(0.024)
Max(Math, Reading)	-1.509*	-1.414*	-1.612**	-0.658	-0.837	-0.674
	(0.636)	(0.583)	(0.583)	(0.822)	(0.753)	(0.805)
Writing Section Score	0.202	0.338	0.116	0.629	0.445	0.193
	(0.526)	(0.523)	(0.696)	(0.453)	(0.462)	(0.558)
Polynomial Specification	Local Linear	Quadratic	Cubic	Local Linear	Quadratic	Cubic
Number of Observations	38019	66831	66831	13289	21656	21656

Notes: Cell entries are estimated discontinuities at the remediation placement cut score. The running variable is the minimum of the math and reading scale score (both recentered to be zero at the passing cutoff) and the sample is limited to students who passed the writing section. Standard errors are adjusted for clustering at the test score level in parentheses. *Statistically significant at the 5% level, **Statistically significant at the 1% level.

Table 3: Effect of TASP Performance on College Enrollment

	14-Day Buffer Sample					
Effect of Passing all Sections						
Sharp RD (sample restricted to passing writing section)	0.001	-0.001	0.003	0.002	0.001	0.005
	(0.006)	(0.006)	(0.008)	(0.006)	(0.006)	(0.008)
Fuzzy RD (IV = pass math & reading)	0.001	-0.001	0.006	0.004	0.002	0.009
	(0.009)	(0.008)	(0.012)	(0.008)	(0.008)	(0.011)
Fuzzy RD (IV1=pass math; IV2=pass reading)	0.007	0.017	0.016	0.017*	0.016	0.016
	(0.007)	(0.010)	(0.013)	(0.009)	(0.009)	(0.012)
Effect of Passing Math, Reading						
Math	0.005	0.010	0.011	0.009	0.008	0.010
	(0.006)	(0.007)	(0.009)	(0.005)	(0.006)	(0.008)
Reading	0.004	0.005	0.000	0.009	0.008	0.003
	(0.010)	(0.008)	(0.010)	(0.010)	(0.009)	(0.011)
	90-Day Buffer Sample					
Effect of Passing all Sections						
Sharp RD (sample restricted to passing writing section)	-0.020	-0.029	-0.030	-0.016	-0.023	-0.019
	(0.028)	(0.029)	(0.039)	(0.023)	(0.024)	(0.032)
Fuzzy RD (IV = pass math & reading)	-0.017	-0.033	-0.036	-0.010	-0.024	-0.020
	(0.041)	(0.043)	(0.061)	(0.034)	(0.036)	(0.051)
Fuzzy RD (IV1=pass math; IV2=pass reading)	-0.004	-0.023	-0.049	0.005	-0.015	-0.036
	(0.032)	(0.043)	(0.059)	(0.033)	(0.037)	(0.050)
Effect of Passing Math, Reading						
Math	0.005	-0.001	-0.009	0.009	0.002	-0.003
	(0.020)	(0.020)	(0.027)	(0.016)	(0.017)	(0.023)
Reading	-0.027	-0.034	-0.056	-0.019	-0.029	-0.049
	(0.047)	(0.046)	(0.058)	(0.041)	(0.039)	(0.049)
Polynomial Specification	Local Lin.	Quadratic	Cubic	Local Lin.	Quadratic	Cubic
Baseline Controls?	No	No	No	Yes	Yes	Yes

Notes: Cell entries are the estimated effect of passing the TASP on the probability of college enrollment. For the 14- and 90-day buffer samples, estimates in the first row ("Sharp RD") are linear probability models where the sample is limited to students who passed the writing test and the running variable is the minimum of the recentered math and reading section scores. Estimates in the second row are 2SLS estimates where a dummy for passing both math and reading is used as an instrumental variable for passing the overall exam and the running variable is the minimum of the math and reading scores. Estimates in the third row are 2SLS estimates where a dummy for passing math and a dummy for passing reading are used as instrumental variables and separate control functions for both the math and reading scores are included. For the fourth and fifth rows, estimates of the effect of passing math and of passing reading are generated from a single linear probability model with separate control functions for the math and reading scores. Estimates in columns 1 and 4 are from local linear models where the sample is limited to observations within a bandwidth of the cut score that is chosen by a cross-validation procedure (Imbens and Lemieux, 2008). Estimates in the remaining columns are from global polynomial models of degree 2 (columns 2 and 5) or 3 (columns 3 and 6) where the polynomial terms are interacted with a passing dummy. Standard errors in parentheses are adjusted for clustering at the test score level (models that have control functions in both math and reading use the multi-way clustering adjustment proposed by Cameron et al., 2011). Estimates in columns 4-6 include controls for variables listed in Table 2. See text for details. * Significant at the 5% level ** Significant at 1% level.

Table 4: Sensitivity to Buffers Between TASP and Enrollment of Different Length

Dep Variable: Enrollment			14-45 Day Buffer			
Sharp RD (sample restricted to passing writing section)	0.015	0.007	0.021	0.016	0.009	0.024
	(0.017)	(0.016)	(0.017)	(0.015)	(0.016)	(0.017)
Fuzzy RD (IV = pass math & reading)	0.011	0.009	0.024	0.015	0.011	0.029
	(0.021)	(0.019)	(0.022)	(0.018)	(0.018)	(0.021)
			46-90 Day Buffer			
Sharp RD (sample restricted to passing writing section)	0.006	0.008	0.019	0.008	0.013	0.023
	(0.029)	(0.028)	(0.035)	(0.031)	(0.031)	(0.039)
Fuzzy RD (IV = pass math & reading)	0.012	0.008	0.035	0.022	0.017	0.045
	(0.038)	(0.036)	(0.048)	(0.039)	(0.038)	(0.050)
Polynomial Specification	Local Lin.	Quadratic	Cubic	Local Lin.	Quadratic	Cubic
Baseline Controls?	No	No	No	Yes	Yes	Yes

Notes: Cell entries are the estimated effect of passing the TASP on the probability of college enrollment.The estimates in the first row ("Sharp RD") are linear probability models where the sample is limited to students who passed the writing test and the running variable is the minimum of the recentered math and reading section scores. Estimates in the second row are 2SLS estimates where a dummy for passing both math and reading is used as an instrumental variable for passing the overall exam and the running variable is the minimum of the math and reading scores. Estimates in columns 1 and 4 are from local linear models where the sample is limited to observations within a bandwidth of the cut score that is chosen by a cross-validation procedure (Imbens and Lemieux, 2008). Estimates in the remaining columns are from global polynomial models of degree 2 (columns 2 and 5) or 3 (columns 3 and 6) where the polynomial terms are interacted with a passing dummy. Standard errors in parentheses are adjusted for clustering at the test score level. Estimates in columns 4-6 include controls for variables listed in Table 2. See text for details. * Significant at the 5% level ** Significant at 1% level.

Table 5: Effect of TASP Performance by Subgroup on College Enrollment

	14-Day Buffer Sample			90-Day Buffer Sample		
Black	-0.016	-0.023	-0.001	-0.107**	-0.130**	-0.048
	(0.013)	(0.013)	(0.018)	(0.039)	(0.040)	(0.056)
Hispanic	0.010	0.009	0.025*	0.002	-0.005	0.030
	(0.009)	(0.010)	(0.012)	(0.031)	(0.035)	(0.047)
White	0.002	0.003	-0.003	-0.009	-0.011	-0.032
	(0.008)	(0.009)	(0.010)	(0.026)	(0.027)	(0.034)
Male	-0.003	0.002	0.005	-0.041	-0.035	-0.032
	(0.009)	(0.010)	(0.013)	(0.033)	(0.037)	(0.054)
Female	0.005	0.000	0.005	0.003	-0.014	-0.011
	(0.006)	(0.006)	(0.007)	(0.022)	(0.021)	(0.024)
Econ. Disadvantaged	0.000	0.002	0.001	-0.007	-0.011	-0.025
	(0.012)	(0.014)	(0.017)	(0.038)	(0.041)	(0.051)
Not Econ. Disadvantaged	0.002	0.001	0.005	-0.018	-0.026	-0.018
	(0.006)	(0.007)	(0.009)	(0.024)	(0.026)	(0.037)
Polynomial Specification	Local Lin.	Quadratic	Cubic	Local Lin.	Quadratic	Cubic

Notes: Cell entries are estimates of the effect of passing the TASP on the probability of enrolling in college. Estimates are from the "Sharp RD" specification that use baseline covariates; see notes to Tables 3 and the text for additional details. Standard errors in parentheses are adjusted for clustering at the test score level. * Statistically significant at the 5% level. ** statistically significant at the 1% level.

Table 6: Estimates of Effect of TASP Performance by Number and Subject of Sections Passed

Parameter	Full Sample	Pass Writing	Fail Writing
Pass math&reading vs. fail both	0.085	0.102	0.028
	(0.049)	(0.057)	(0.073)
Pass vs. fail reading, given fail math	0.056	0.057	0.040
	(0.049)	(0.057)	(0.079)
Pass vs. fail reading, given pass math	-0.011	0.015	-0.077
	(0.027)	(0.031)	(0.067)
Pass vs. fail math, given fail reading	0.096*	0.086	0.105
	(0.049)	(0.058)	(0.081)
Pass vs. fail math, given pass reading	0.029	0.044	-0.012
	(0.028)	(0.028)	(0.068)

Notes: Table shows results from the estimator proposed by Papay et al. (2011a) for a regression discontinuity estimator with multiple running variables. The estimation equation is:

$$Y = \beta_0 + \beta_1 P_M + \beta_2 P_R + \beta_3 P_M{*}P_R + \beta_4 M + \beta_5 R + \beta_6 M{*}R + \beta_7 P_M{*}M + \beta_8 P_R{*}R + \beta_9 P_M{*}R + \beta_{10} P_R{*}M + \beta_{11} M{*}R{*}P_M + \beta_{12} M{*}R{*}P_R + \beta_{13} M{*}P_M{*}P_R + \beta_{14} R{*}P_M{*}P_R + \beta_{15} M{*}R{*}P_M{*}P_R + v$$

Where the variables are as defined in the text. The effect of failing both math and reading relative to passing both is equal to β1+ β2 + β3; the effect of passing vs. failing reading conditional on failing math is equal to β2, and conditional on passing math it is β2 + β3; the effect of passing vs. failing math conditional on failing reading is given by β1 and conditional on passing reading is β1 + β3. As in Papay et al. (2011b), optimal bandwidth computed using a generalization of the cross-validation procdure described in Imbens and Lemieux (2008) and standard errors adjusted for clustering on clusters defined by the combination of math and reading scores. All estimates use the 14-day buffer used in Tables 3 and 5. Estimates in the first column are based on all students irrespective of performance on the writing section. The second column is restricted to students who passed the writing section, and the third column is restricted to students who failed the writing section. See text for additional details. * Statistically significant at the 5% level, ** Significant at the 1% level.

Table 7: Effect of TASP Performance on Enrollment Depending On The Number of Sections a Student Failed

	14-Day Buffer			**90-Day Buffer**		
	Effect of failing both math and reading vs. passing at least one subject [Running Variable=Max(Math,Reading)]					
Full Sample	0.014	0.017*	0.009	-0.001	-0.006	-0.031
	(0.008)	(0.007)	(0.009)	(0.024)	(0.025)	(0.031)
Restricted to Passing Writing	0.009	0.010	0.008	-0.017	-0.037	-0.045
	(0.009)	(0.008)	(0.011)	(0.028)	(0.030)	(0.039)
Restricted to Failing Writing	0.023	0.029*	0.013	0.029	0.035	-0.007
	(0.013)	(0.011)	(0.015)	(0.025)	(0.028)	(0.036)
	Effect of failing math or reading vs. passing both subjects [Running Variable=Min(Math,Reading)]					
Full Sample	0.004	0.002	0.006	-0.008	-0.020	-0.018
	(0.006)	(0.006)	(0.008)	(0.026)	(0.027)	(0.037)
Restricted to Passing Writing	0.002	0.001	0.005	-0.016	-0.023	-0.019
	(0.006)	(0.006)	(0.008)	(0.023)	(0.024)	(0.032)
Restricted to Failing Writing	0.017	0.014	0.004	0.032	0.003	-0.030
	(0.015)	(0.015)	(0.019)	(0.052)	(0.052)	(0.066)
Polynomial Specification	Local Lin.	Quadratic	Cubic	Local Lin.	Quadratic	Cubic

Notes: Cell entries in the upper panel are from models that use the maximum of the math and reading scores as the running variable, and are RD estimates of the coefficient on a dummy for passing either math or reading. This parameter captures the effect of passing math or reading vs. failing both. Results in the lower panel are from models that use the minimum of the math an reading scores are RD estimates of the coefficient on a dummy for passing both math and reading, and capture the effect of passing math and reading vs. failing at least one of these two subjects (note that the estimates in the second row of the lower panel are identical to those reported in Table 3). All estimates are from models that include the baseline covariates listed in Table 2. Standard errors in parentheses adjusted for clustering on the running variable. See text for additional details. *-Significant at the 5% level **-Significant at 1% level.