Remedial Writing and Collegiate Performance: Using Grader Toughness to Identify Causal Effects

Melissa Famulari

Department of Economics
University of California – San Diego
La Jolla, CA 92093-0508
mfamulari@ucsd.edu

October 14, 2009

Abstract

The University of California, San Diego (UCSD) uses remedial writing classes to teach underprepared students to write in a manner that effectively communicates with faculty. A key essay exam that determines student placement in remedial writing is the Analytic Writing Placement exam (AWP). UCSD's freshman class of 2000 are equally likely to be assigned to a tough AWP grader, but getting a tough AWP grader significantly increases the likelihood of being placed in remedial writing. Standard errors are large on the IV estimates, but I do find that students who were required to take a remedial class because they got a tough grader take significantly fewer classes at UCSD in their first year and there is some evidence that their grades are higher on average in these classes. For non-Asian students at the placement margin, the likelihood of remediation is highly influenced by the toughness of their AWP grader and the likelihood of graduating from UCSD in five or six years is significantly reduced if placed in remedial writing. In a separate analysis of UCSD students whose AWP graders disagreed on remedial placement, I find significant negative effects of the remedial writing requirement on graduation probabilities and grades. However, there remains the possibility of downward bias on the estimated effect of remediation since observed measures of writing ability are lower for students eventually placed in remediation. I conclude that large positive effects of remediation for UCSD students at the placement margin are unlikely.

I. Introduction

Many colleges and universities admit students that are not considered ready *by the institution* to take their classes. These students are typically required to take one or more remedial (also called basic, entry level, or developmental) classes to prepare them for the institution's courses. According to the National Center for Education Statistics (2003), 68% of public 4-year degree-granting institutions offered a remedial writing program in 2000. As states seek to expand access to higher education to a wider range of students, such as through Texas' 1997 "Top 10% Rule" and Berkeley's 2002 and UCLA's 2007 adoption of "comprehensive review" to evaluate all freshman applications, the question of whether remedial classes address limitations in student skills at the college level is likely to become more important.

The key difficulty in identifying the effect of remediation is that only students with low skills take remedial classes. Worse collegiate and labor market outcomes for students who take remedial classes are likely attributable to the lower skills of remediated students and says little about the effectiveness of remedial classes. Several recent papers have tried to resolve this endogeneity issue for remedial classes in math, writing, and/or reading at the collegiate level by using a small randomized trial at the margin of placement (Aiken et. al. (1998)), instrumental variables (Bettinger and Long (2009)), and regression discontinuity techniques (Butcher et. al. (2009), Calcagno and Long (2008), Leake and Lesik (2007), Lesik (2006, 2007), Martorell and McFarlin (2007), and Moss and Yeaton (2006)). The estimated effects of remedial classes on college outcomes are mixed. Bettinger and Long generally find positive effects of remediation for students at the margin of placement in Ohio. Among the three large RD studies, Butcher et. al. for several years of Wellesley students, Calcagno and Long for Florida community college students and Martorell and McFarlin for Texas students generally find no effects of remediation on college outcomes. Martorell & McFarlin is the only study to examine the labor market effects of college remediation, and they find no effect of remediation on future labor market earnings.

The main empirical strategy uses the random assignment of students to essay exam graders with different propensities to fail students to identify causal effects. This strategy follows that of Kling (2006) and Doyle (2007, 2008). Kling uses judge toughness as an instrumental variable to identify the causal effects of incarceration length on the employment and earnings while Doyle uses case workers' propensity to place children in foster care as an

instrumental variable to identify the causal effects of foster care on child outcomes and adult crime. The University of California's Analytical Writing Placement exam, or AWP, is a key determinant of whether a University of California student is required to take remedial writing. Unlike a math test where graders are likely to agree on the number of correct answers, the AWP is an essay exam and the score is likely to be influenced by the grader's writing standards or her opinions on the value of taking remedial writing classes. If there are tough and easy graders of the AWP then, for students whose writing abilities are at border of passing and not passing the AWP, getting a hard grader will place the student in the remedial class while getting an easy grader will exclude the student from the remedial class. If graders are randomly assigned to students, then the causal effects of the remediation requirement can be estimated for students whose writing skills are at the margin of placement. As pointed out by Doyle (2007, 2008), this may be a particularly useful policy margin as it informs the UC whether or not they should encourage AWP graders to grade more like easy or like hard graders.

The data used in this study are the 3,117 students in the University of California, San Diego's (UCSD) freshman class of 2000. These data have detailed student admissions data, demographic and family characteristics and so I can condition on an extensive set of observed student characteristics. Further, these data report every course attempted and grade assigned over a six year period which allows me to examine many collegiate outcomes. In addition, the University of California requires that all students pass two college level-writing classes and performance in these college writing classes is expected to be directly impacted by any improvements in writing skills. Finally, UCSD is a highly selective institution and only one other study, Butcher et. al. (2009) study of Wellesley students, has examined the local average treatment effects of remediation at such a high ability margin. The effect of remediation at this ability margin is policy relevant to the extent that selective colleges and universities value diversity and tradeoff preparedness for diversity, particularly given there is some evidence suggesting that low skilled students may be less likely to graduate from selective colleges (Loury and Garman (1995) and Light and Strayer (2008)).

Out of the 1815 students who took the AWP exam, there are 857 who have at least one grader ID reported on their AWP exam. In these data there is an important tradeoff between more precisely identifying the grader's failure rate and the number of students over which to estimate the effects of the remedial writing requirement. To

clarify the nature of these tradeoffs, I estimate results for two samples: students whose grader is identified on at least 6 other exams (640 students and 82 graders) and students whose grader is identified on at least 9 other exams (450 students and 45 graders). I find that getting a hard grader is not significantly related to observed student characteristics. However, students scored by a hard grader are significantly more likely to be required to take remedial classes than students graded by an easy grader. For students at the margin of placement, being required to take a remedial class significantly reduced the credit hours taken in the first year at UCSD by 7.8-8.5 hours and there is some evidence that first year GPA in UCSD classes is higher. For students at the placement margin, I find no significant effects of being required to take a remedial class on the GPA in the first college-level writing class, the cumulative GPA of graduates, the probability of dropping out in 2 or 5 years or of graduating in 4, 5, or 6 years. However, the effect of remedial placement is imprecisely estimated. I then focus on two groups whose likelihood of remedial placement is strongly influenced by their AWP grader's toughness: non-Asian students and students whose writing abilities are more likely to be at the margin of placement. Standard errors are still large, but non-Asian students at the placement margin who are required to take remedial classes are significantly less likely to graduate in 5 or 6 years.

I also examine the effect of the remedial writing requirement across students whose graders disagreed on whether the student passed the AWP exam. These exams are sent to a third, more experienced grader who determines whether the student passes. Out of the 1815 students who took the AWP exam, there are 271 students where the first two graders disagreed. There are few significant observed differences across these students who are and are not placed in remedial classes except, importantly, measured writing ability (SAT I, SAT II and AP scores) are lower for those who are required to take remedial writing. Though I condition on these measures of writing ability, I find many negative effects of remedial placement for the students whose first two graders disagreed: they are significantly less likely to graduate in 5 years, they have lower first year GPAs, graduates have lower cumulative GPAs, and they have lower grades in the first UC writing class. These negative results may be attributable to unobserved differences in the abilities of students required to take and not take remedial classes. Given the IV estimates based on grader strictness and the results for students where initial graders disagree on

placement, the UC remedial writing requirement is unlikely to have substantial positive effects on the academic outcomes of UCSD students at the margin of placement.

The paper proceeds as follows. Section two discusses the grader failure rate and the conditions under which it is a valid empirical approach to identify the causal effects of remedial writing. Section three discusses the UC Entry Level Writing Requirement and what determines whether students are placed in a remedial class. Section four describes the student data and the grader data that will be used in the analysis. Section five presents the empirical results and section six concludes.

II. Empirical Approach

Following Doyle (2007, 2008), let academic outcome, Y, such as grade point average, be a function of observable student characteristics, X, and an indicator for being required to take remedial writing, R, for student i:

(1) $\quad Y_i = X_i \beta + \alpha_i R_i + \varepsilon_i$

(2) $\quad R_i = 1 \; if \; (Z_i \gamma + X_i \delta + u_i > 0)$

where Z_i is a variable that influences placement in remediation, but does not influence academic outcomes directly. α_i will be positive if the benefits of remediation exceed the costs[1]. The primary econometric problem to estimating (1) is that R_i may be correlated with ε_i. An omitted variable, such as attending a low quality high school or taking less rigorous classes in high school, may both increase the likelihood a student is remediated and lead to lower grades in college. Omitting these variables would then bias downward α_i.

A plausible candidate for Z is the propensity of the grader of the placement exam to fail students. Consider two types of graders: easy and hard. The difference in student outcomes across these two graders would estimate

---

[1] Remedial classes could reduce academic outcomes. Remedial classes take time that the student could devote to other classes, which may lead to lower grades or increased time to degree. In order to avoid increasing the time to degree, students may take a heavier course load in her first quarter of college which may impact her grades. There may be negative peer effects from being placed in a remedial class with other low writing ability students and possibly continuing with this cohort into college-level writing classes.

local average treatment effects: the average effect of being required to take remedial writing across students where the strictness of their grader determined placement in remediation. Let $Z_i=1$ if the student is assigned to a tough grader and $Z_i=0$ if assigned to an easy grader then:

$$(3) \qquad a^{LATE} = \frac{E(Y \mid Z = 1) - E(Y \mid Z = 0)}{P(R = 1 \mid Z = 1) - P(R = 1 \mid Z = 0)}$$

Doyle (2007, 2008) identifies the four conditions to interpret (3) as a local average treatment effect (LATE) as

(i) $E(R|Z=z)$ is a nontrivial function of z. This is a requirement that getting a tough grader does affect the likelihood of being required to take remedial classes. I test this requirement below.

(ii) Z and ε_i are independent. This is a requirement that the failing tendencies of the AWP graders do not directly impact the student's academic outcomes at UCSD. Since AWP graders are not UCSD instructors, this condition is likely to be met.

(iii) Z and u_i are independent. This is a requirement that the failing tendencies of AWP graders are uncorrelated with unobserved determinants of remedial placement. If graders are randomly assigned to students, this condition is likely to be met. I examine this requirement below.

(iv) $\gamma > 0$. Below I test whether an increase in grader failure rates increases the likelihood of remedial placement. The fact that γ is the same for all grader types means that grader's propensity to fail students monotonically affects the likelihood the student is remediated. Thus, there can be no students who would be more likely to fail if assigned to an easy grader than if assigned to a hard grader. This is difficult to test but Doyle (2007, 2009) points out that the bias introduced by violations of monotonicy goes to zero if the effect of remedial placement is the same for students induced into treatment by an easy grader as it is for students induce into treatment by a hard grader.

If the student's writing abilities are so weak (strong) that both easy and tough graders would fail (pass) the student, the effect of remediation cannot be identified. The effect of remediation is only identified for students whose writing abilities lead easy and hard graders to disagree about whether the student passed the exam.

Finally, in these data, grader strictness is continuous and not dichotomous. Thus, the estimates in this paper are marginal treatment effects, the effect of the remedial writing requirement as we move from easier to harder graders. The marginal treatment effect is, again using Doyle's (2007, 2008) notation:

$$(4) \qquad \alpha^{MTE} = \frac{\partial E(Y \mid (P(R=1) \mid Z=z))}{\partial P(R=1 \mid Z=z)}$$

III. The University of California's Entry-Level Writing Requirement (ELWR)

All University of California (UC) students must demonstrate a minimal proficiency in English composition.

"…Each student must be able to understand and to respond adequately to written material typical of reading assignments in freshman courses. This ability must be demonstrated in student writing that communicates effectively to University faculty"

Academic Senate Regulation 600

Students may not take university-level writing classes until the Entry Level Writing Requirement (ELWR) is fulfilled. If a UC student does not fulfill the ELWR within her first year, then she is barred from enrolling in any class at the university.[2] Demonstrating writing proficiency can be done prior to matriculation by scoring sufficiently high on a standardized writing test[3] or by passing with a C or better an acceptable college writing class.

---

[2] At some UC campuses there is an exception for students who are identified as non-native English speakers. At UCSD for example, non-native speakers are placed into English as a Second Language classes (ESL). They have three quarters to pass the ESL class and then an additional three quarters to pass the remedial writing class.

[3] In 2000, students could meet the ELWR with a minimum score of 680 on the SAT II-Writing Test, a 3 on either Advanced Placement composition test, a 5 on the International Baccalaureate Higher Level English A exam, or an 8 on the University of California Analytical Writing Placement Exam. 26,549 freshman enrolled in a University of California in 2000 and 67.6% met the entry-level writing requirement prior to matriculation (10.6% with SAT II, 22.3% with AP, 28.0% with AWP and 6.7% with IB or a college composition class). http://uasother.ucop.edu/cgi-bin/awpe/results.pl

Once a student matriculates to a specific UC campus, then she must fulfill the ELWR as specified by the campus. UCSD students must pass with a C grade or better a remedial writing class designed by UCSD but taught by local community college instructors. In addition, students must pass an essay exam that is similar in all respects to AWP exam. Surprisingly, a passing grade on the AWP exam is an 8 while a passing score on UCSD's Exit exam is a 6. Thus, a UCSD student need not increase her score on an essay exam to fulfill the ELWR.

If the AWP graders find signs that English is a second language (ESL) for the student, then UCSD requires the student to pass an ESL class before taking the remedial writing class. Students have up to three quarters to pass the ESL class and if she does not, then the student has not met the ELWR and must leave UCSD. If the student passes the ESL class within three quarters, then she takes the remedial writing class as discussed above.

IV. Data

IV.A. Student Data

For each of the 3,117 freshman who entered UCSD in 2000, I have every UCSD class taken and the grade earned from fall quarter of 2000 through the spring quarter of 2006 as well as the grade and course title for every English class taken at other institutions that UCSD accepted for credit. I data on high school GPA, location (state and city) and name of the last high school attended, race, sex, parent's income, education level of mother and father, and the student's scores on the following pre-college exams: SAT I-Verbal, SAT I-Math, SAT II-English Subject, SAT II-Math Subject, a third SAT II Achievement Test in various subjects, AP English Language and Composition, AP English Literature and Composition, and the University of California's Analytical Writing Placement Exam (AWP).

Over a quarter of UCSD's freshman did not test out of the ELWR and, while most of these students (82%) chose to take remedial classes at UCSD, a substantial fraction (13%) chose to take a remedial class before matriculation. There are seven students that take a remedial class when they tested out and there are 39 students

who were placed in remedial writing but may not have taken a remedial class.[4] Thus, the estimates I report in this paper are intent-to-treat effects.

There is some evidence that students who take remedial writing after matriculation have lower skills than students who take the remedial classes before matriculation. Since I have no plausible instrument to address the endogenous choice of whether to take remedial classes before or after matriculation to UCSD, I cannot identify the effects of UCSD's remedial program separate from other remedial programs. Thus, the estimates in this paper are interpretable as the effect of the remedial writing requirement for students at the margin of placement, as we go from more lenient to more strict graders.

IV.B. Analytic Writing Placement Exam (AWP) Grader ID and Score

UC requires all admitted students to take the AWP exam in May if they have not fulfilled the ELWR by April. Two graders read each AWP essay. Each grader grades the essay from 1 to 6. The student's AWP score is the sum of the two grader's scores unless (a) the scores across the two graders are two or more points apart or (b) one reader scores the student as a 3 (a just failing grade) and the other reader scores the student as a 4 (a just passing grade). In these cases a third, more experienced grader scores the exam and the final exam grade is 2 times the third reader's score (roughly three-quarters of the time the reason the exam goes to the third grader is that one grader scored a 3 while the other grader scored a 4). In Figure 1, the yellow dots indicate when the exam goes to a 3rd grader. 1815freshman took the AWP exam and I have the physical exam for 1747 students. On the exam is recorded each grader's score and, on occasion, the grader's ID. While each grader may assign a score from 1 to 6, over 78% of all grades assigned are either a 3 or a 4. Recall that to pass the AWP Exam, the sum of the two grader's scores must be at least an 8. Thus, most graders assign a just passing grade (4) or a just failing grade (3).

Nearly 90% of the students who took the AWP exam did so in May and the other 10% took the exam at UCSD, generally in September. The recording of grader IDs differs across May and September. In May,

---

[4] UCSD has a code for how each student fulfilled the ELWR and I base whether the student took a college level class that fulfilled the ELWR on this code. However, the exact date the code was accurate is unclear and 27 of these 39 students took an English class at another institution that UCSD accepted for credit and so these 27 students may have fulfilled the ELWR. Of the 12 students that have no transfer English class, no remedial writing classes and did not grade out of the requirement, 9 drop out of UCSD.

approximately 200 graders gather at UC-Berkeley to grade the AWP exams. Graders are given batches of 20 randomly selected exams to grade at a time and several graders sit at one table with a table leader, an experienced AWP grader, who monitors and provides advice to the other graders at the table. Each grader writes her score on the exam in invisible ink. "Runners" take the bundles of 20 exams to a separate room for a second grader to score. Runners also record grader IDs on the front of the exam. Though runners are supposed to mark the grader IDs on each exam, only 682 of the 1,572 May exams (or 43.4%) have one or more grader IDs recorded. Runners do write the first two grader IDs on the exams whenever the exam is given to a third reader for grading (346 exams) or when the student is identified as a non-native speaker (170 exams).

If a student misses the May AWP, then she takes the exam on campus, typically right before fall classes start. There were nine graders at UCSD in 2000 and all marked his or her score and initials on each exam. Therefore, I know all grader ID's for the 175 AWP exams graded at UCSD. Note also that the second grader knows both the first grader's ID and score. Most graders at UCSD are also graders for the May exam but I do not know their May ID number and so the failure rates are for a grader at a point in time. Unlike May graders who score the test from 1-6 in whole numbers only, UCSD graders occasionally assigned plus, minus, and split scores such as "2/3".

There are 857 students that took the AWP exam where at least one grader ID is identified on their exam. It is for this group of students that I am potentially able to calculate grader failure rates as described below.

V. Estimates

V.A. Grader Failure Rate

The key instrumental variable is the failure rate of the first grader assigned to score the student's exam. Suppose there are n students. Let FR_{ij} be the failure rate of grader j for student i

$$(5) \quad FR_{ij} = \frac{\sum_{i=1}^{n} G_j F_{ij} - G_j F_{ij}}{\sum_{i=1}^{n} G_j - 1}$$

Where $G_j = 1$ if the grader is j and 0 otherwise and $F_{ij} = 1$ if the grade assigned to student i by grader j is a failing grade and is zero otherwise. [5] Note that student i's own exam score is not used in the calculation of grader j's failure rate.

The first grader is identified on 857 exams, the second grader is identified on 558 exams and the third grader is identified on 186 exams. In all cases but one, second grader IDs are reported on a subset of the exams where the first grader is identified and third grader IDs are on the subset of exams where both the first and second graders are identified. While AWP grader assignment is likely to be random, the *recording* of AWP grader IDs is not. Grader ID's are more likely to be reported when one grader passes the student and the other grader fails the student, when the two graders scores differ by 2 or more points, when the first grader thinks the student is a non-native speaker and when the student takes the exam at UCSD. The first two are plausibly a function of the random matching of easy and hard graders. However, if tough graders are more likely to identify students as having English as a second language, and students identified as non-native speakers have worse outcomes at UCSD, then again the effects of remediation will be biased downwards. Further, it is possible that students taking the AWP exam in September (when I always observe grader ID) are different from students who take the AWP in May. I test for these possibilities below

If I had data on all grader IDs, then I would use the failure rates of the first two graders of the student's exam as instruments. Since the first grader's ID is most frequently recorded, I use the failure rate of the first grader as the instrumental variable for being required to take remedial writing. Student i's first grader may be identified on another student's exam as either a first, second, or third grader. Thus, FR_{ij} can be calculated using exams that j graded as the (1) first grader, (2) the first or second grader or (3) the first, second or third grader. Methods (2) and (3) increase the number of exams used to calculate the grader's failure rate and may increase the number of students used in the subsequent analysis of collegiate outcomes. How sensitive is the grader's failure rate to whether she grades first, second, or third? There are 63 graders who are identified as a first grader at least 5 times. Across methods 1-3, the mean and standard deviation of the failure rate across these 63 graders is .578 (.226), .572 (.215)

---

[5] There are 9 students who have a grader that assigned a "3/4". Since I cannot determine whether the grader passed or failed the student, these exams are not used in the calculation of the grader's failure rate. The students are still used in the analysis because the student's own exam score is not used in determining the grader's failure rate.

and .570 (.213) respectively. I cannot reject the equality of the mean failure rate calculated across the 3 different methods and the correlation of the failure rate calculated using method (1) with method (2) and (3) is .896 and .886 respectively. Since I find no evidence that a grader's propensity to fail students differs depending upon whether she is a first, second or a third grader, I determine the first grader's failure rate based on all exams where the grader is identified (method 3).

Another key question for this study is the minimum number of students over which to calculate the grader's failure rate. There are 857 students whose first grader is identified and the average number of exams where their grader is identified is 18 with a standard deviation of 22. The distribution is highly skewed: for the median student, their first grader's ID is on 10 other exams. There is an important tradeoff in these data: if I require more exams per grader (to reduce the standard deviation of the grader's failure rate), then I substantially reduce the number of students for whom I can estimate the effects of remedial placement. I test the sensitivity of the results to this tradeoff by conducting the analysis when requiring 7 exams per grader (which results in a sample of 617 students and 82 graders) and when requiring 10 exams per grader (which results in a sample of 450 students and 48 graders). There is substantial variation in the failure rate across graders in both samples: the mean grader failure rate is 54.7% with a standard deviation of 19.8% when using 6 or more exams to calculate the grader's failure rate and is 55.3% with a standard deviation of 18.2% when using 9 or more exams to calculate the grader's failure rate.

Table 1 compares sample means across all students taking the AWP exam, students whose AWP grader is observed on at least 6 other exams and students whose AWP grader is observed on at least 9 other exams. A key difference is that students whose first grader is observed multiple times are 11.8%-12.7% more likely to be required to take a remedial class and 10.7-12.1% more likely to be identified as having English as a second language (ESL). This is not surprising since all 250 students where a grader identified the student as ESL have a grader ID and all except 6 of these students failed the AWP. Student demographic characteristics, family background, high school GPA, and admissions test scores are otherwise quite similar across the three samples of students.

V.B. Is there a Relationship between Grader Failure Rates and Student Characteristics?

Given the grading process described above, I expect grader assignment to be random across students. However, it is unclear whether the process that leads to *identifying* a grader is random across students. To test the hypothesis that the identification of graders is random, I regress grader failure rates on observed student demographic characteristics, family background and test scores.[6] I estimate student-level regressions and allow for correlation in errors across students graded by the same individual by clustering the standard errors at the grader level. Results are presented in Table 2. There are few student characteristics significantly related to grader toughness. For the 617 student sample, SAT I: Verbal scores and being of Asian descent are significantly positively related to having a tough grader and parent's income is negatively related to having a tough grader. For the 450 student sample, having a mother who is a college graduate is significantly positively related to having a tough grader. For both samples, F-tests do not reject the hypothesis that all the coefficients are jointly equal to zero. In regressions not reported here, I also ran these same regressions at the grader-level, weighted by the number of exams over which the failure rate is calculated. F-tests reject the hypothesis that grader failure rates are significantly related to the average characteristics of the students they grade.

Finally, I examine whether grader failure rates are higher for students for whom English is a second language (ESL) than for native English speakers. This would occur if hard graders are more likely to identify a student as ESL. AWP graders mark on the exams whether they have found any evidence that the student is a non-native English speaker. I find no significant differences in the mean first grader failure rate for students identified and not identified as ESL in either the 617 or the 450 student samples.

V.C. Grader Failure Rates and the Requirement to take Remedial Writing

How important is the toughness of the student's first AWP grader in determining whether the student is required to take remedial writing? There is unlikely to be a one-to-one relationship since the toughness of the

---

[6] Demographic characteristics are dummy variables for female, Asian, Hispanic, other race. Family background variables are parent's income and a dummy variable for whether mom is a college graduate and a dummy variable for whether dad is a college graduate. Test score data are scores on SAT I Math, SAT I Verbal, SAT II Math, SAT II Writing, AP Literature and Composition, AP Language and Composition, and high school GPA. Missing dummy variables for each variable are also included so that all students could be included in the analysis.

student's second grader and third grader (if the exam goes to a third grader) also determines placement in remedial writing. Further, for students taking AP tests, an AP exam score may exempt the student from the requirement even if she fails the AWP exam. The AWP exam appears to be easier to pass than either AP composition test because 84.2% of students just failing the AWP must take remedial writing while 39.2% of students just below the AP Language and Composition cutoff and 46.9% of students below the AP Literature and Composition cutoff are required to take remedial writing). Thus, AWP grader toughness is likely to be an important predictor of being required to take a remedial class.

I use an Epanechnikov kernel in Figure 2 to illustrate the distribution of first grader failure rates across students who were required and not required to take remedial writing. For both samples, students who were required to take remedial writing had harder graders than students who were not required to take remedial writing. For students whose grader is identified on at least 10 exams, the mean grader failure rate is .062 higher for those required to take remedial writing and this difference is significant at the 1% level. For students whose grader is identified on at least 7 exams, the mean grader failure rate is .049 higher for those required to take remedial writing and this difference is significant at the 1% level. Kolmorgov-Smirnov tests also confirm that, for both samples, the distribution of grader failure rates is significantly higher for those required to take a remedial writing class. Finally, I plot the likelihood a student is required to take a remedial class given her grader's failure rate in Figure 3 (I require that there be at least 5 students who have a first grader with the specified failure rate). I find that the likelihood a student is required to take remedial writing is generally increasing in the grader's failure rate. I examine this relationship more formally by estimating the following linear probability model (though results are similar if I use a probit model),

(6) $R_{ij} = a_0 + bX_i + dF_{ij} + e_{ij}$

Where R_{ij} is 1 if remediation is required for student i assigned to grader j, F_{ij} is the failure rate for grader j, (student i's first AWP grader), X_i is a vector of student i's characteristics and e_{ij} is an error term that is correlated across graders. I cluster the standard errors by grader to address this correlation. Table 3 reports the effect of the grader's

failure rate on the likelihood that the student is required to take remediation unconditionally and conditional on X_i (student test scores, high school GPA, demographic and family background characteristics). The estimated effect of the grader's failure rate is significantly positive in all regressions and does not change much depending upon whether I condition on the X_i since, as shown in the previous section, the grader's failure rate is unrelated to student characteristics. Getting a grader that is 20% more likely to fail students, which is about a one standard deviation increase in grader toughness, is estimated to increase the likelihood that a student must take remedial writing by 9.9-6.9 percentage points.

In using instrumental variables to identify causal effects, not only must the instrument be exogenous but it must also be strongly correlated with the endogenous variable. Stock, Wright and Yogo (2002) identify the critical values of the F statistic for the set of instruments from the first stage regression for testing the null hypothesis of weak instruments. In the case of one endogenous variable with one instrumental variable, the F-statistic is 8.98 and so the t-statistic on the grader's failure rate must be 2.98 to reject the hypothesis that the grader's failure rate is a weak instrument. From Table 3, the t-statistic on the grader's failure rate is 3.44 and 4.52 in the regressions conditional on student characteristics for the 7 exams per grader and 10 exams per grader samples respectively and so I reject the hypothesis that the grader's failure rate is a weak instrument for remedial placement.

V.D. Remedial Writing Placement and UCSD Outcomes

Table 4 reports IV estimates of the effect of being required to take a remedial writing class on academic performance at UCSD for students at the margin of placement. I also report OLS estimates for comparison. In both OLS and IV regressions, I cluster the standard errors at the grader level. The academic outcomes I examine are the probability that students drop out in 2 or 5 years, the probability that students graduate in 4, 5 or 6 years, the total number of credit hours taken at UCSD (remedial coursework is excluded), and the cumulative GPA of students who graduate within 6 years. I also focus on several first year outcomes including the total number of UCSD credit hours taken in the first year (remedial coursework is excluded), the total number of class hours taken in the first year (including remedial coursework), and the first year GPA. I focus on first year outcomes because they are less subject to the bias introduced by students dropping out of UCSD than outcomes in later years.

The most direct measure of the effects of the remedial writing program is likely to be grades in college writing classes. The University of California requires that all graduates must pass two quarters of college-level writing classes with a C- or better. While I expect performance in these two college-level writing classes to be higher if the remedial writing classes improve student's ability to write, there are some important limitations in using grades in college writing to measure the effectiveness of the remedial classes. First, 3.7% or 114 UCSD students never take a college level writing class and students required to take a remedial class are significantly more likely to be in this group. Second, 130 students start the college writing sequence but do not complete it and again, students required to take remedial classes are significantly more likely to be in this group. Finally, all UCSD students took the writing classes for a letter grade except Warren College students. Warren is one of UCSD's six colleges and the writing classes are unique to each of the six colleges. Warren graded their standard writing classes pass/no pass. Therefore, I exclude all Warren College students from the analysis and I focus on grades earned in the first college-level writing class. Thus, I examine the effect of the remedial writing program on college writing conditional on having taken a college level class.

In each regression I condition on test scores (SAT I Math and Verbal, SAT II Math and English, the two AP English Tests), high school GPA, sex, race/ethnicity, parent's income, whether mom graduate from college, and whether dad graduated from college. I include missing dummy variables for each variable so that all students are included in each regression. Column 2 of Table 4 reports OLS estimates for both samples. Even after conditioning on an extensive set of observed characteristics, OLS estimates indicate that students required to take a remedial class have worse academic outcomes than similar students who are not required to take a remedial class. In particular, the probability that students placed in remedial writing drop out in 2 years is significantly higher by 5.4-6.0 percentage points which is substantial given the average two year dropout rate is only 7.5-8.0% across the two samples. Students placed in remediation are significantly less likely to graduate in 5 years than those not placed in remediation by 9.9-11.9 percentage points. OLS estimates small negative effect of remediation on first year GPA and the cumulative GPA of graduates.

As noted earlier, there may be unobserved differences across students required and not required to take remedial writing and so the OLS estimates are not the causal impacts of remediation. The IV estimates reported in

column (3) are the estimated marginal effects of remedial placement as we go from less strict to more strict graders for students at the placement margin. Standard errors of the IV estimates are large which contributes to the finding of few significant effects of the remedial writing requirement. Students at the placement margin who are required to take remedial writing take significantly fewer credit hours, -8.5 to -7.8 hours, at UCSD in their first year (i.e., excluding the remedial writing classes). This translates roughly into two classes less at UCSD in the student's first year. Does the drop in UCSD credit hours match the increase in hours spent in remedial writing? There is no statistically significant difference in total credit hours taken in the first year (remedial plus UCSD) across the remediated and not remediated students though the point estimates are negative. There are no significant effects of the remedial writing requirement on the likelihood the student drops out in 2 or 5 years nor on the probability of graduating in 4, 5, or 6 years. There is some evidence that the first year GPA is higher for remediated students at the margin of placement: point estimates are positive for both samples and the estimate is statistically significant at the 10% level in the 10 exams per grader sample. The 95% confidence interval for the effect of remediation on first year GPA for students at the placement margin is -.054 to .965 average grade points. Higher grades may result for several possible reasons: (1) improved writing skills as a result of the remedial class improves performance at UCSD, (2) a lighter UCSD course load if UCSD classes are harder on average than the remedial writing classes, (3) students take easier classes at UCSD while trying to meet the remedial writing requirement. In results not presented in Table 4, I examine whether students required to take remedial classes are less likely to complete a year of calculus in their first year to provide some evidence on whether remediation leads students to take an easier course load. I find no statistically significant effects of the remedial writing requirement on the likelihood of taking a year of calculus for students at the margin of placement.

To reduce the standard errors of the IV estimates, I examine several groups to determine if there is a group of students whose likelihood of placement is strongly influenced by their first grader's propensity to fail students. I examine several groups including men, women, Asians, non-Asians and several different writing skill groupings based on SAT I Verbal scores and present the results in Table 5.[7] Note that since all of these groups are

---

[7] Other groups I examined that are not reported in Table 5, but are available from the author upon request, are students not taking AP exams, non-Asian and non-Hispanic, May test takers, and ESL students. For none of these groups was the grader's failure rate significant in the "required to take remedial writing" regressions.

predetermined when the AWP exam is taken and graded, selecting the sample on the basis of these pre-determined variables will not bias the results but may increase the explanatory power of the first stage regression and reduce the standard errors of the IV estimates.

Given the smaller sample sizes when separately estimating equation 6 for men, women, Asians and non-Asians, I focus on students whose graders are identified at least 5 times (recall only 4 graders will be used in the calculation of the failure rate) or 6 times (recall only 5 graders will be used in the calculation of the failure rate). I report these results in Table 5. I find that grader failure rates are positively, but generally insignificantly related to remedial writing placement for both Asian students and female students.[8] However, for men and non-Asian students, getting a tough grader significantly increases the likelihood of placement. Only for non-Asian students can I reject the hypothesis that the grader failure rate is a weak instrument. Excluding students whose SAT I Verbal scores are in the top or, especially, the bottom of the distribution also resulted in a sample of students whose likelihood of taking remedial classes is highly influenced by their grader's failure rate. Therefore, I re-estimate the IV effects of remedial placement for students whose SAT I scores are not in the bottom 15% and non-Asians separately and report these results in Table 6.

The standard errors of the IV estimates are still large. Again, I find that both non-Asian students and students with SAT I scores not in the bottom 15% take significantly fewer UCSD credit hours, but not fewer total credit hours, if they are required to take remedial classes and are at the placement margin. For neither group do I find a statistically significant effect of being required to take remedial writing on first year GPA. For non-Asian students at the placement margin, being required to take remedial classes significantly reduces the probability that a student graduates in 5 or six years.

---

[8] A possible explanation for this finding for Asian students is that they are more likely to have English as a second language and both easy and hard graders are likely to place ESL students in remediation. In UCSD's freshman class, women are 9.1 percentage points more likely to take an AP test than men and this added test taking (and not scores conditional on taking the test) results in women being 8.4 percentage points more likely to fulfill the ELWR with AP scores than men. The higher fraction of women taking AP tests and grading out of the ELWR with AP scores likely contributes to why AWP grader failure rates are less important for women than for men in determining remedial placement.

V.E. Robustness Check: Students Whose AWP Graders Disagreed on Placement

As a check on the estimates using grader failing tendencies, I examine the effect of being required to take a remedial class among students whose writing abilities are likely to be quite similar: those students where the first two graders disagreed on whether the student should pass. In these data there are 271 students who earned a total 7 of based on their first two graders because one grader assigned a 3 (a just failing grade) and one grader assigned a 4 (a just passing grade). These essays are given to a third, more experienced, grader who makes the final determination on whether the student passes the AWP. Table 5 compares students required and not required to take a remedial class both in the full sample and in the sample of students whose first two graders disagreed. Across all UCSD freshman, students required to take remedial writing have significantly lower test scores, lower high school GPAs, have parent's whose income and education is significantly lower, and are significantly more likely to be male, Asian and Hispanic. However, if I look only at the 271 students whose writing abilities were so similar that their first two graders disagreed on whether to pass the student, there are few significant differences across those required and not required to take remedial writing. This is not simply due to the smaller sample size increasing the standard error of the difference: differences across the two groups are generally much smaller than in the full sample. Though measured writing ability is still significantly lower for students placed in remedial writing among these students whose first graders disagreed, differences are 71%-36% smaller than in the overall sample. Given students whose first graders disagreed are a much more homogenous group on the basis of observable characteristics and because I can condition on observed differences in writing ability, I expect the downward bias in the estimated effect of remedial writing to be smaller in this sample of students whose first graders disagreed. Again note that I will be estimating a local average treatment effect because I estimate the effect of the remediation requirement only for students whose essays were so good that at least one grader thought the student should pass but not so good that both grader thought the student should pass.

Table 8 reports the estimated effect of being required to take remedial writing on performance at UCSD on the full sample and across students whose AWP graders disagreed. I report unconditional OLS estimates and estimates conditional on student demographic characteristics to see how much simply conditioning on observed student characteristics reduces the negative estimated effects of remediation. In the full sample, column (2) provides a

straight comparison of mean outcomes across students placed in remedial classes and those not and indicates significant negative effects of remedial placement on all UCSD outcomes I examine. Conditioning on observed characteristics (column 3) does generally reduce the negative estimated effects of remedial placement for the full sample. This is particularly true for grades: while there are still statistically significant negative effects of remediation on first year GPA and GPA in the first college writing class, point estimates are only about one-third the size of the unconditional estimates. Since observed test scores, student demographic characteristics, and family background variables do not account for much of the variation in drop out probabilities or graduation rates by year, controlling for these covariates does not have a substantial impact on the estimated effect of remediation.

In columns 5-8, I examine the effects of remedial placement for the smaller, but more homogenous group of students who had AWP graders that disagreed on whether to place the student in remediation. For this group of 271 students, the estimated effects of remedial placement are (1) similar whether or not I condition on observed student characteristics and (2) are similar to the estimated effects from the entire sample that are conditional on demographic characteristics. However, standard errors are much larger in the smaller sample and so there are many fewer significant effects. Nevertheless, students placed in remediation have significantly lower grades, particularly in the first year and in the first UCSD writing class (-.19 and -.18 respectively which is about a third of a standard deviation), 4.8 fewer credit hours taken in the first year and 12.4 percent less likely to graduate in 5 years.

Finally, I exclude students taking either of the AP composition tests from the analysis. Students taking AP writing tests are more likely to have taken advanced writing classes in high school and are less likely to be placed in remediation. Since I do not have information about the rigor of the student's high school curriculum, this omitted variable would bias downward the effect of remedial writing placement and excluding AP test takers may help mitigate this source of bias. Fifty-six percent of UCSD freshman take an AP composition test. Among the 1365 freshman who do not take an AP composition test, the difference in observed writing ability across those placed and not placed in remediation is somewhat smaller than in the overall sample: the difference in SAT I Verbal and SAT II Writing scores is -85.56 (4.13) and -96.58 (4.28) respectively. As reported in Table 9, the estimated effects of remediation are still negative among students not taking AP composition tests, but fewer coefficient estimates are significant both because of small declines in the point estimates and small increases in standard errors. There are

161 students who did not take an AP test and whose AWP graders disagreed on placement. Again, observed measures of writing ability are significantly worse for students placed in remediation: SAT I Verbal scores are lower by 43.7 points and SAT II Writing scores are lower by 30.6 points. The 107 students who were placed in remedial writing are significantly more likely to drop out in the first two years, are significantly less likely to graduate in 4 or 5 years, take significantly fewer UCSD classes in their first year and have significantly lower first year GPAs than the 54 students who did not have to take remedial writing.

VI. Conclusions

Using data from UCSD's freshman class of 2000, I estimate the effect of being required to take remedial writing for students who are at the margin of placement. I find that students taking the AWP exam are equally likely to get a hard grader and that getting a hard grader significantly increases the likelihood a student is required to take a remedial writing class. For students at the margin of placement, placement in remedial writing has few significant effects on academic outcomes except to reduce the number of UCSD credit hours taken in the first year and, at least in one specification, to increase first year GPA in the UCSD classes they do take. AWP grader toughness strongly influences non-Asian students' placement in remedial classes and, for non-Asian students at the margin, placement in remediation reduces the likelihood of graduating in five and six years.

With these data it is possible to identify another group of students at the margin of being placed in remedial writing: students whose AWP graders disagreed about placement. Family background, student demographic and math test scores are not significantly different for those placed in remedial classes and those not among the 271 students whose AWP graders disagreed on placement. Differences in observed writing skills are much smaller than in the overall sample but are still significantly lower for those placed in remedial writing. Though I condition on observed writing skills, I find significant negative effects of remedial placement on 5 year graduation rates, first year GPA, grade in the first college writing course, and the cumulative GPA of graduates.

Given the imprecision of the IV estimates and the fact that, even among students whose AWP graders disagreed on placement, observed measures of writing ability are lower for those placed in remedial classes, the results of this analysis must be interpreted with care. Nevertheless, it is unlikely that there are large positive effects to requiring

UCSD's freshman to take a remedial class if they are at the margin of placement. What could account for this finding? It is possible that the remedial classes offered at UCSD and at community colleges are ineffective. Second, consider a model where the student produces academic outcomes using inputs such as the student's time, textbooks, attending class, office hours, etc. Remedial classes could simply reduce the student's use of other inputs in the academic production function with no change in academic outcomes. Finally, the effect of the remedial class is likely to be heterogeneous with low writing ability students more likely to benefit than students with high writing ability. Clearly the UC believes that benefits to high ability students are low because high ability students are not required to take a remedial writing class. It is possible that the ability threshold UC has set is too high and, while students at this margin do not benefit, students of lower ability may. A randomized trial that excuses students at all levels of ability from the remedial writing requirement as well as collecting information about hours spent studying and other inputs into students' academic production functions would shed light on these issues.

Bibliography

Aiken, Leona S., Stephen G. West, David E. Schwalm, James L. Carroll and Shenghwa Hsuing (1998), "Comparison of a Randomized and Two-Quasi-Experiments in a Single Outcome Evaluation: Efficacy of a University-Level Remedial Writing Program." *Evaluation Review*, 22(2), 207-244.

Bettinger, Eric P. and Bridget T. Long (2009) "Addressing the Needs of Under-prepared College Students: Does College Remediation Work?" *Journal of Human Resources*, 44(3):736-771.

Bettinger, Eric P. and Bridget T. Long (2007), "Institutional Responses to Reduce Inequalities in College Outcomes: Remedial and Developmental Courses in Higher Education," *Economic Inequality and Higher Education: Access, Persistence and Success*. Stacy Dickert-Conlon and Ross Rubenstein, eds. New York: Russell Sage Foundation Press.

Breneman, David W. and William N. Haarlow (1997), "Remedial Education: Costs and Consequences," *Remediation in Higher Education*. Washington, D.C.: Thomas B. Fordham Institution.

Calcagno, Juan Carlos and Bridget T. Long (2008). "The Impact of Postsecondary Remediation Using a Regression Discontinuity Approach: Addressing Endogenous Sorting and Noncompliance," Working Paper 14194, National Bureau of Economic Research.

Doyle, Joseph J. "Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care" *Journal of Political Economy* 116(4). August 2008: 746-770

Doyle, Joseph J. "Child Protection and Child Outcomes: Measuring the Effects of Foster Care." *American Economic Review* 97(5). December 2007: 1583-1610.

Kling, Jeffrey R., "Incarceration Length, Employment, and Earnings." *American Economic Review* 96(3). June 2006: 863-876.

Lagerlöf, Johan N. M. and Andrew J. Seltzer (2008), "The Effects of Remedial Mathematics on the Learning of Economics: Evidence from a Natural Experiment". *Journal of Economic Education*, forthcoming.

Leake, M. and Sally Lesik (2007), " Do Remedial English Programs Impact First Year Success in College? An Illustration of the Regression Discontinuity Design," *International Journal of Research and Method in Education*, Vol 30. No. 1, pp. 89-99.

Lesik, Sally (2006), "Applying the Regression Discontinuity Design to Infer Causality with Non-Random Assignment." *The Review of Higher Education*, Vol 30. No. 1, pp. 1-19.

_____(2007), "Do Developmental Mathematics Programs have a Causal Impact on Student Retention? An Application of Discrete-Time Survival and Regression Discontinuity Analysis" *Research in Higher Education*, Vol. 48, No 5, pp. 483-608.

_____(2008), "Studying the Effectiveness of Programs and Initiatives in Higher Education Using the Regression-Discontinuity Design" J.C. Smart (eds.) Higher Education: Handbook of Theory and Research, Springer Netherlands, pp. 277-

Light, Audrey and Wayne Strayer (2000), "Determinants of College Completion: School Quality or Student Ability?" *Journal of Human Resources*, Vol. 35, No. 2, pp. 299-332.

Loury, Linda and David Garman (1995) "College Selectivity and Earnings" *Journal of Labor Economics*, Vol. 13, No. 2, pp. 289-308.

Martorell, Paco and Isaac McFarlin Jr. (2007). "Help or Hindrance? The Effects of College Remediation on Academic and Labor Market Outcomes," Unpublished manuscript, Rand and the University of Texas at Dallas.

Moss, Brian G. and William H. Yeaton (2006), "Shaping Policies Related to Developmental Education: An Evaluation using the Regression-Discontinuity Design." *Educational Evaluation and Policy Analysis*, Vol. 28, No. 3, pp. 215-229.

Pozo, Susan and Charles A. Stull (2006), "Requiring a Math Skills Unit: Results of a Randomized Experiment". *American Economic Review, Papers and Proceedings*, Vol. 96(2), pp. 437-441.

U.S. Department of Education, National Center for Education Statistics. *Remedial Education at Degree-Granting Post Secondary Institutions in Fall 2000*. NCES 2004-010, by Basmat Parsad and Laurie Lewis. Project officer: Bernard Greene. Washington D.C. 2003.

Table 1: Sample Means for Students taking the AWP exam: All Students, Students Whose Grader is Identified on at Least 7 Exams and Students Whose Grader is Identified on at Least 10 Exams

	All Students taking AWP		AWP takers: Grader on 7+ exams		AWP takers: Grader on 10+ exams	
Remediation Required	.404	(.491)	.522	(.500)	.531	(.500)
Take UCSD Remedial Class	.364	(.481)	.486	(.500)	.507	(.501)
Take UCSD or Transfer Remed	.393	(.489)	.520	(.500)	.531	(.500)
ESL Identified	.142	(.349)	.263	(.440)	.249	(.433)
Did not take SAT I exams	.028	(.164)	.031	(.173)	.033	(.180)
SAT I Verbal (takers)	568.034	(78.455)	558.328	(81.188)	563.081	(77.455)
SAT I Math (takers)	640.969	(78.484)	644.629	(81.724)	647.816	(78.869)
Did not take SAT II-Writing	.068	(.251)	.068	(.252)	.069	(.254)
SAT II Writing (takers)	565.195	(69.521)	555.461	(72.747))	559.141	(69.487)
Did not take SAT II-Math	.069	(.254)	.070	(.255)	.071	(.257)
SAT II-Math (takers)	629.213	(82.960)	633.554	(88.036)	637.536	(83.470)
Did not take AP: Lang and Comp	.870	(.336)	.865	(.341)	.862	(.345)
AP Lang & Comp (takers)	2.386	(.782)	2.181	(.647)	2.194	(.623)
Did not take AP: Lit and Comp	.646	(.478)	.713	(.453)	.722	(.448)
AP Lit & Comp (takers)	2.805	(.882)	2.605	(.800)	2.560	(.787)
Missing High School GPA	.006	(.078)	.005	(.070)	.007	(.081)
High School GPA (reporters)	3.972	(.237)	3.958	(.240)	3.959	(.242)
Fraction Female	.515	(.500)	.511	(.500)	.500	(.501)
Missing Race	.080	(.272)	.079	(.271)	.084	(.278)
Faction Asian (reporters)	.477	(.500)	.482	(.500)	.481	(.500)
Fraction Hispanic (reporters)	.138	(.345)	.136	(.343)	.107	(.310)
Fraction Other (reporters)	.031	(.174)	.037	(.189)	.044	(.205)
Fraction White (reporters)	.354	(.478)	.345	(.476)	.369	(.483)
Missing Mom's Ed	.043	(.203)	.050	(.219)	.058	(.234)
Fraction Mom Coll Grad (reporters)	.438	(.496)	.439	(.497)	.469	(.500)
Missing Dad's Ed	.069	(.254)	.060	(.238)	.067	(.250)
Fraction Dad Coll Grad (reporters)	.507	(.500)	.439	(.497)	.550	(.498)
Missing Family Inc.	.181	(.385)	.169	(.375)	.178	(.383)
Family Inc (reporters)	68,192	(68,811)	68,724	(77,346)	72,722	(77,829)
Sample Size	1815		617		450	

Table 2: Are AWP Graders Randomly Identified? First Grader Failure Rates and Student Characteristics

Dependent Variable: Failure Rate of Student's First AWP Grader
OLS regressions with standard errors clustered at the grader level.

	7 or more Exams Per Grader		10 or more Exams per Grader	
	Coefficient (Standard Error)		Coefficient (Standard Error)	
SAT I: Math/1000	-0.030	(0.188)	0.051	(0.211)
SAT I: Verbal/1000	0.245*	(0.137)	0.174	(0.125)
SAT II: Writing/1000	-0.146	(0.147)	-0.032	(0.173)
SAT II: Math/1000	-0.206	(0.198)	-0.290	(0.250)
High School GPA	0.013	(0.030)	0.013	(0.034)
AP Lit & Comp	0.006	(0.018)	0.013	(0.021)
AP Lang & Comp	0.021	(0.039)	0.019	(0.048)
Female	0.015	(0.020)	-0.001	(0.023)
Asian	0.029*	(0.017)	0.024	(0.020)
Hispanic	0.012	(0.030)	0.007	(0.033)
Other Race/Ethnic	0.013	(0.051)	0.002	(0.057)
Mom College Grad	0.025	(0.022)	0.042*	(0.025)
Dad College Grad	0.016	(0.021)	-0.010	(0.017)
Parent's income/1m	-0.325**	(0.125)	-0.176	(0.131)
Took May AWP	-0.081	(0.060)	-0.082	(0.066)
Missing SAT I	0.159	(0.141)	0.164	(0.144)
Missing SAT II Writing	-0.021	(0.112)	0.026	(0.137)
Missing SAT II Math	-0.175	(0.143)	-0.242	(0.176)
Missing HS GPA	-0.003	(0.133)	0.011	(0.150)
Missing AP Lit	0.024	(0.046)	0.055	(0.052)
Missing AP Lang	-0.008	(0.099)	-0.019	(0.121)
Missing Race/Ethnicity	0.054*	(0.030)	0.052	(0.032)
Missing Mom Ed	0.007	(0.037)	-0.008	(0.045)
Missing Dad Ed	-0.000	(0.047)	0.002	(0.059)
Missing Parent's Income	-0.034	(0.025)	-0.012	(0.028)
Constant	0.605**	(0.172)	0.569**	(0.206)
R-Square	0.067		0.079	
F-test for Regression	$F_{(25, 81)}=1.32$		$F_{(25, 47)}=1.38$	
Prob > F	0.176		0.166	
Number of Students	617		450	
Number of Graders	82		48	
Mean of Dependent Var.	0.547		0.553	
Stand. Dev. Dep. Var.	0.198		0.182	

Notes: *=10% significance and **=5% significance.

Table 3: First Grader Failing Tendencies and the Requirement to Take Remedial Writing

Dependent Variable: Dichotomous variable that equals 1 if remediation is required.
Linear probability model with standard errors clustered at the grader level.

	7 or more Exams per Grader		10 or more Exams per Grader	
	Coeff. (St. Error) Coeff. (St. Error)		Coeff. (St. Error) Coeff. (St. Error)	
1st Grader Failure Rate	0.315** (0.12)	0.344** (0.10)	0.466**(0.13)	0.497** (0.11)
SAT I: Math/1000		0.416 (0.35)		0.340 (0.41)
SAT I: Verbal/1000		-1.524** (0.27)		-1.399** (0.33)
SAT II: Writing/1000		-1.639** (0.26)		-1.666** (0.24)
SAT II: Math/1000		0.362 (0.36)		0.298 (0.46)
AP Lit & Comp		-0.188** (0.04)		-0.195** (0.05)
AP Lang & Comp		-0.131** (0.06)		-0.128* (0.07)
High School GPA		-0.117 (0.09)		-0.121 (0.12)
Female		0.068** (0.03)		0.083** (0.03)
Asian		0.072* (0.04)		0.092** (0.04)
Hispanic		0.085 (0.07)		0.049 (0.09)
Other Race/Eth		-0.070 (0.09)		-0.079 (0.11)
Mom College Grad		-0.034 (0.04)		-0.005 (0.04)
Dad College Grad		0.017 (0.04)		0.021 (0.04)
Parent's income/1m		0.214 (0.28)		-0.068 (0.37)
Missing SAT I		-0.704** (0.27)		-0.553* (0.28)
Missing SAT II Writing		-0.801** (0.17)		-0.835** (0.17)
Missing SAT II Math		0.284 (0.24)		0.212 (0.29)
Missing AP Lit		-0.214* (0.13)		-0.236 (0.16)
Missing AP Lang		-0.269* (0.14)		-0.233 (0.17)
Missing HS GPA		-0.547 (0.49)		-0.643 (0.59)
Missing Race/Eth		-0.063 (0.07)		-0.047 (0.07)
Missing Mom Ed		-0.209* (0.12)		-0.253 (0.15)
Missing Dad Ed		0.160 (0.10)		0.174 (0.13)
Missing Parent's Inc.		0.115** (0.05)		0.107* (0.05)
Constant	0.349** (0.07)	2.525** (0.41)	0.274**(0.08)	2.486** (0.52)
R-square	0.016	0.369	0.029	0.356
F-test	F(1,81)=7.04	F(25,81)=38.99	F(1,47)=12.35	F(25,47)=66.12
#Students	617	617	450	450
# Graders	82	82	48	48
Mean Dep. Var.	.522	.522	.531	.531

Notes: *=10% significance and **=5% significance.

Table 4: The Remedial Writing Requirement and UCSD performance: OLS and IV estimates

| Dependent variable | Grader Identified on 7 or more exams (82 graders) | | | | Grader Identified on 10 or more exams (48 graders) | | | |
	Mean Dependent Variable (1)	OLS Remediation Required Coef (2)	IV Remediation Required Coef (3)	Sample Size (4)	Mean Dependent Variable (5)	OLS Remediation Required Coef. (6)	IV Remediation Required Coef. (7)	Sample Size (8)
Dropped Out in 2 years	.075	.054* (.029)	-.034 (.185)	617	.080	.061* (.034)	-.026 (.184)	450
Dropped Out in 5 years	.147	.045 (.037)	.082 (.216)	617	.151	.066 (.043)	-.091 (.224)	450
Earned BA or BS in 4 years	.476	-.077 (.060)	-.057 (.287)	617	.476	-.059 (.069)	.122 (.232)	450
Earned BA or BS in 5 years	.781	-.099** (.047)	-.323 (.257)	617	.776	-.119** (.048)	-.021 (.243)	450
Earned BA or BS in 6 years	.825	-.053 (.042)	-.245 (.231)	617	.824	-.072 (.046)	.016 (.219)	450
Total UCSD Credit Hours	179.958	-12.129** (4.450)	-50.991 (34.953)	617	178.871	-14.349** (5.235)	-23.170 (29.440)	450
UCSD Credit Hours Year 1	43.040	-4.658** (.476)	-8.521** (4.096)	617	43.112	-4.472** (.477)	-7.837** (3.189)	450
UCSD+Remed Hours Year 1	45.989	.107 (.487)	-4.520 (3.819)	617	46.090	.195 (.519)	-4.052 (3.035)	450
First Year GPA UCSD classes	2.925	-.100** (.050)	.334 (.288)	617	2.935	-.111* (.059)	.456* (.253)	450
GPA in First UC writing class	2.887	-.085 (.063)	.286 (.489)	436	2.888	-.070 (.079)	.633 (.457)	317
Cum GPA of Students Graduating in 6 yrs	3.142	-.061* (.035)	.192 (.210)	509	3.143	-.055 (.039)	.186 (.192)	371

Notes: Each element of columns (2), (3), (5), and (6) represents a separate regression with the "Remediation Required" coefficient (and standard error) reported. In all regressions, standard errors are clustered at the grader level. The other regressors are SAT I-Math, SAT I-Verbal, SAT II-Writing, SAT II-Math, high school GPA, AP Lit & Comp, AP Lang & Comp, parent's income, dummy variables for female, Asian, Hispanic, mom is a college grad, dad is a college grad, and missing dummy variables for each variable. In the regressions for GPA in 1st UCSD writing class, I exclude Warren college students because Warren's writing class graded pass/no pass. I also excluded students taking honors writing and students who withdrew from the class. *=10% significance, **=5% significance

Table 5: First Grader Failing Tendencies and the Requirement to Take Remedial Writing

Dependent Variable: Dichotomous variable that equals 1 if remediation is required.
Linear probability model with standard errors clustered at the grader level.

	# Students	# Graders	Failure Rate Coeff. (Stand. Error)	T-statistic
I. Women				
A. Failure rate calculated over at least 4 exams	374	112	.240 (.124)	1.94
B. Failure rate calculated over at least 5 exams	344	95	.202 (.133)	1.52
II. Men				
A. Failure rate calculated over at least 4 exams	360	110	.284 (.108)	2.63
B. Failure rate calculated over at least 5 exams	334	94	.299 (.110)	2.71
III. Asians				
A. Failure rate calculated over at least 4 exams	345	111	.093 (.093)	1.00
B. Failure rate calculated over at least 5 exams	306	95	.105 (.089)	1.17
IV. Non-Asians				
A. Failure rate calculated over at least 4 exams	389	106	.418 (.138)	3.04
B. Failure rate calculated over at least 5 exams	372	94	.251 (.097)	2.76
V. Trimmed SAT I Verbal Scores (Failure rate calculated over at least 6 exams)				
A. All students with non-missing SAT I scores & reporting HS GPA	596	82	.332 (.102)	3.25
B. Exclude bottom 5%	570	82	.351 (.108)	3.26
C. Exclude bottom 10%	541	81	.367 (.110)	3.33
D. Exclude bottom 15%	523	81	.392 (.107)	3.67
E. Exclude bottom 20%	484	80	.410 (.116)	3.54
F. Exclude top 5%	565	82	.353 (.101)	3.50
G. Exclude top 10%	536	82	.318 (.100)	3.19
H. Excluding top 15%	501	82	.296 (.104)	2.84
I. Exclude top 20%	445	82	.297 (.103)	2.89
J. Exclude top and bottom 5%	539	82	.374 (.107)	3.50
K. Exclude top and bottom 10%	481	81	.353 (.109)	3.25
L. Exclude top and bottom 15%	428	80	.360 (.107)	3.36
M. Exclude top and bottom 20%	333	78	.417 (.118)	3.52

Notes: Each failure rate coefficient is from a regression that conditions on SAT I-Math, SAT I-Verbal, SAT II Writing, SAT II-Math, high school GPA, AP Lit & Comp, AP Lang & Comp, parent's income, dummy variables for female, Asian, Hispanic, Other Race/ethnicity, mom is a college grad, dad is a college grad, and missing dummy variables for each variable. The male and female regressions exclude the female dummy variable and the Asian and non-Asian regressions exclude the Asian dummy variable.

Table 6: The Remedial Writing Requirement and UCSD performance: Separate Results for Non-Asian Students and Higher Writing Ability Students

Dependent variable	Excluding Students with SAT I Verbal Scores in bottom 15%				Non-Asian Students			
	Mean Dependent Variable (1)	OLS Remediation Required Coef (2)	IV Remediation Required Coef (3)	Sample Size (4)	Mean Dependent Variable (5)	OLS Remediation Required Coef. (6)	IV Remediation Required Coef. (7)	Sample Size (8)
Dropped Out in 2 years	.069	.066** (.026)	-.080 (.182)	523	.090	.048 (.039)	.196 (.160)	389
Dropped Out in 5 years	.141	.050 (.035)	-.009 (.190)	523	.152	.040 (.044)	.319 (.211)	389
Earned BA or BS in 4 years	.488	-.059 (.060)	.013 (.281)	523	.488	-.129* (.072)	-.268 (.305)	389
Earned BA or BS in 5 years	.788	-.100** (.043)	-.206 (.211)	523	.779	-.112** (.047)	-.511* (.285)	389
Earned BA or BS in 6 years	.834	-.057 (.038)	-.126 (.187)	523	.817	-.062 (.044)	-.453* (.259)	389
UCSD Credit Hours Year 1	43.810	-4.499** (.534)	-9.433** (3.666)	523	43.297	-4.360** (.575)	-5.891* (3.173)	389
UCSD+Remed Hours Year 1	46.188	.214 (.548)	-5.933 (3.698)	523	45.487	.014 (.575)	-2.899 (3.263)	389
First Year GPA UCSD classes	2.940	-.091* (.049)	.361 (.263)	523	2.962	-.106** (.052)	-.220 (.297)	389
GPA in First UC writing class	2.905	-.110 (.070)	.608 (.585)	373	2.958	-.014 (.054)	.533 (.466)	287
Cum GPA of Students Graduating in 6 yrs	3.142	-.053 (.037)	.087 (.206)	436	3.186	-.060 (.040)	.004 (.230)	318

Notes: Each element of columns (2), (3), (6) and (7) represents a separate regression with the "Remediation Required" coefficient (and standard error) reported. Standard errors are clustered at the grader level. The other regressors are SAT I-Math, SAT I-Verbal, SAT II Writing, SAT II-Math, high school GPA, AP Lit & Comp, AP Lang & Comp, parent's income, dummy variables for female, Hispanic, mom is a college grad, dad is a college grad, and missing dummy variables for each variable. The failure rate is calculated over at least 4 exams in the non-Asian student regressions. In the regressions excluding the bottom 15% of SAT I Verbal scores, failure rates are calculated over at least 6 exams, I include an Asian dummy variable and exclude the student who does not report high school GPA. In the regressions for GPA in the 1st UCSD writing class, I exclude Warren college students because Warren's writing class graded pass/no pass. I also excluded students taking honors writing and students who withdrew from the class. Finally, *=10% significance, **=5% significance.

Table 7: Differences in Means across Students who ARE and ARE NOT Required to take Remedial Writing: Across the Full Sample and Across Students Whose AWP Graders Disagreed on Placement

Variable	All Students (n=3117)			Students Whose Graders Disagreed (n=271)		
	Remediation Required (St.Dev.)	Remediation NOT Required (St.Dev.)	Difference (St.Error)	Remediation Required (St.Dev.)	Remediation NOT Required (St.Dev.)	Difference (St.Error)
Take UCSD Remedial Class	.827 (.378)	0	.827** (.008)	.932 (.253)	0	.932** (.023)
Take UCSD or Transfer Remed	.951 (.216)	.003 (.055)	.948** (.005)	.979 (.142)	0	.979** (.013)
Missing SAT I scores	.030 (.171)	.022 (.145)	.009 (.006)	.041 (.199)	.024 (.154)	.017 (.022)
SAT I Verbal (non-missing)	528.333 (79.877)	632.141 (70.808)	-103.809** (3.048)	546.643 (67.490)	592.213 (65.989)	-45.570** (8.273)
SAT I Math (non-missing)	634.767 (87.066)	659.097 (69.640)	-24.329** (3.099)	639.857 (88.179)	638.361 (77.121)	1.496 (10.307)
Missing SAT II-Writing	.081 (.274)	.050 (.219)	.031** (.010)	.068 (.253)	.024 (.154)	.044* (.026)
SAT II Writing (non-missing)	531.678 (71.332)	639.078 (76.828)	-107.400** (3.219)	548.309 (63.974)	579.672 (50.156)	-31.363** (7.214)
Missing SAT II-Math	.076 (.266)	.050 (.217)	.027** (.009)	.082 (.276)	.032 (.177)	.050* (.029)
SAT II-Math (non-missing)	626.499 (89.894)	647.822 (77.510)	-21.323** (3.438)	631.791 (89.282)	630.083 (88.652)	1.708 (11.159)
Missing AP (36): Lang and Comp	.901 (.299)	.697 (.460)	.204** (.017)	.856 (.352)	.800 (.402)	.056 (.046)
AP Lang & Comp (nonmissing)	1.873 (.335)	3.275 (.826)	-1.402** (.094)	1.952 (.218)	2.440 (.651)	-.489** (.149)
Missing AP (37): Lit and Comp	.823 (.382)	.422 (.494)	.402** (.019)	.829 (.378)	.496 (.502)	.333** (.054)
AP Lit & Comp (nonmissing)	1.816 (.389)	3.479 (.872)	-1.664** (.074)	1.880 (.332)	2.937 (.644)	1.057** (.136)
Missing HS GPA	.003 (.050)	.005 (.064)	-.002 (.002)	.007 (.083)	.008 (.089)	-.001 (.010)
HS GPA (non-missing)	3.953 (.233)	3.993 (.245)	-.040** (.010)	3.964 (.239)	3.980 (.230)	-.016 (.029)
Fraction Female	.503 (.500)	.552 (.497)	-.049** (.020)	.548 (.499)	.568 (.497)	-.020 (.061)
Missing Family Inc.	.158 (.365)	.230 (.421)	-.072** (.017)	.151 (.359)	.192 (.395)	-.041 (.046)
Family Inc/1000 (non-missing)	55.988 (66.906)	86.353 (68.594)	-30.364** (3.084)	55.775 (42.904)	65.350 (50.447)	-9.575 (6.224)

Table 7 (Continued): Differences in Means across Students who ARE and ARE NOT Required to take Remedial Writing: Across the Full Sample and Across Students Whose AWP Graders Disagreed on Placement

Variable		All Students (n=3,117)			Students whose Graders Disagreed (n=271)		
		Remediation Required (St.Dev.)	Remediation NOT Required (St.Dev.)	Difference (St.Error)	Remediation Required (St.Dev.)	Remediation NOT Required (St.Dev.)	Difference (St.Error)
Missing Race		.073 (.260)	.099 (.299)	-.026** (.012)	.068 (.253)	.072 (.260)	-.004 (.031)
Among Reporters	Fraction Asian	.562 (.496)	.386 (.487)	.176** (.021)	.434 (.497)	.457 (.500)	-.023 (.063)
	Fraction Hispanic	.150 (.357)	.095 (.293)	.055** (.013)	.184 (.389)	.190 (.394)	-.006 (.049)
	Fraction Other	.028 (.166)	.033 (.179)	-.005 (.008)	.015 (.121)	.043 (.204)	-.028 (.021)
	Fraction White	.259 (.439)	.486 (.500)	-.227** (.021)	.368 (.484)	.310 (.465)	.057 (.060)
Missing Dad's Ed		.070 (.256)	.068 (.251)	.002 (.010)	.089 (.286)	.016 (.126)	.073** (.028)
Among Reporters	Dad HS Drop Out	.175 (.380)	.052 (.222)	.123** (.012)	.158 (.366)	.106 (.309)	.052 (.042)
	Dad HS Grad	.170 (.376)	.089 (.285)	.081** (.013)	.135 (.343)	.130 (.338)	.005 (.043)
	Dad Some College	.208 (.406)	.196 (.397)	.011 (.017)	.188 (.392)	.228 (.421)	-.040 (.051)
	Dad Coll Grad	.447 (.498)	.663 (.473)	-.215** (.020)	.519 (.502)	.537 (.501)	-.018 (.063)
Missing Mom's Ed		.038 (.190)	.050 (.218)	-.012 (.009)	.068 (.253)	.040 (.197)	.028 (.028)
Among Reporters	Mom HS Drop Out	.219 (.414)	.063 (.243)	.156** (.012)	.199 (.400)	.133 (.341)	.065 (.047)
	Mom HS Grad	.177 (.382)	.118 (.322)	.060** (.014)	.103 (.305)	.117 (.322)	-.014 (.039)
	Mom Some College	.243 (.429)	.240 (.427)	.003 (.018)	.316 (.467)	.300 (.460)	.016 (.058)
	Mom Coll Grad	.361 (.481)	.579 (.494)	-.219** (.021)	.382 (.488)	.450 (.500)	-.067 (.062)
Sample Size		798	2,319		145	125	

Notes: *=10% significance, **=5% significance

Table 8: The Remedial Writing Requirement and UCSD performance: Results for Full Sample and for Students Whose AWP Graders Disagreed about Remedial Writing Placement

|	FULL SAMPLE				STUDENTS WHOSE AWP GRADERS DISAGREED			
Dependent variable	Mean Dependent Variable (1)	Remediation Required Coefficient (2)	Remediation Required Coef. (Conditional) (3)	Sample Size (4)	Mean Dependent Variable (5)	Remediation Required Coefficient (6)	Remediation Required Coef. (Conditional) (7)	Sample Size (8)
Dropped Out in 2 years	.071	.038** (.011)	.045** (.013)	3117	.077	.040 (.033)	.049 (.038)	271
Dropped Out in 5 years	.125	.059** (.014)	.055** (.017)	3117	.129	.017 (.041)	.026 (.049)	271
Earned BA or BS in 4 years	.526	-.143** (.020)	-.078** (.025)	3117	.494	-.122** (.061)	-.115 (.073)	271
Earned BA or BS in 5 years	.805	-.088** (.016)	-.075** (.020)	3117	.797	-.080 (.049)	-.124** (.058)	271
Earned BA or BS in 6 years	.847	-.067** (.015)	-.057** (.018)	3117	.834	-.056 (.045)	-.059 (.054)	271
Total UCSD Credit Hours	179.282	-1.826 (1.912)	-7.425** (2.372)	3117	181.190	.078 (5.667)	-5.258 (6.731)	271
UCSD Credit Hours: Year 1	44.927	-5.834** (.238)	-4.617** (.295)	3117	43.561	-5.181** (.609)	-4.848** (.729)	271
UCSD+Remedial Hours: Year 1	46.172	-.978** (.227)	-.337 (.282)	3117	46.011	-.633 (.579)	-.735 (.688)	271
First Year GPA UCSD classes	2.996	-.277** (.023)	-.113** (.026)	3117	2.924	-.129* (.067)	-.187** (.073)	271
GPA in 1st UCSD writing class	2.993	-.319** (.031)	-.107** (.036)	2314	2.936	-.156* (.085)	-.175* (.100)	192
Cum GPA of Students Graduating in 6 yrs.	3.211	-.230** (.017)	-.073** (.020)	2639	3.126	-.117** (.052)	-.106* (.061)	226

Notes: Each element of columns 2, 3, 6 and 7 reports the coefficient and standard error of the "Remediation Required" variable from separate OLS regressions. The conditional regressions in columns (3) and (7) include SAT I-Math, SAT I-Verbal, SAT II Writing, SAT II Math, AP Lit and Comp, AP Lang & Comp, high school GPA, parent's income, dummy variables for female, Asian, Hispanic, whether mom is a college grad, whether dad is a college grad, and missing dummy variables for each variable. In the regressions for GPA in 1st UCSD writing class, I exclude Warren College students because Warren's writing class graded pass/no pass. I also excluded students taking honors writing and students who withdrew from the class. Finally, *=10% significance, **=5% significance.

Table 9: The Remedial Writing Requirement and UCSD Performance Among Students who did not take an AP Writing Test:
All Students and Students whose AWP Graders Disagreed about Remedial Writing Placement

| Dependent variable | No AP Writing Test | | | | No AP Writing Test & AWP Graders Disagreed | | | |
	Mean Dependent Variable (1)	Remediation Required Coefficient (2)	Remediation Required Coef. (Conditional) (3)	Sample Size (4)	Mean Dependent Variable (5)	Remediation Required Coefficient (6)	Remediation Required Coef. (Conditional) (7)	Sample Size (8)
Dropped Out in 2 years	.081	.035** (.015)	.038** (.018)	1365	.087	.075 (.047)	.088* (.050)	161
Dropped Out in 5 years	.137	.050** (.019)	.036 (.023)	1365	.118	.038 (.054)	.057 (.058)	161
Earned BA or BS in 4 years	.482	-.072** (.027)	-.044 (.032)	1365	.497	-.144* (.083)	-.187** (.089)	161
Earned BA or BS in 5 years	.785	-.059** (.022)	-.060** (.027)	1365	.807	-.123* (.066)	-.144** (.070)	161
Earned BA or BS in 6 years	.830	-.049** (.020)	-.038 (.025)	1365	.832	-.085 (.062)	-.098 (.067)	161
Total UCSD Credit Hours	179.766	-6.145** (2.658)	-8.424** (3.191)	1365	181.609	-8.224 (7.801)	-8.965 (8.632)	161
UCSD Credit Hours: Year 1	43.666	-5.350** (.319)	-3.889** (.375)	1365	42.925	-5.490** (.871)	-5.197** (.950)	161
UCSD+Remedial Hours Hours: Year 1.	45.855	-.503* (.292)	-.013 (.348)	1365	46.006	-.854 (.817)	-.998 (.874)	161
First Year GPA UCSD classes	2.959	-.206** (.031)	-.089** (.035)	1365	2.968	-.210** (.097)	-.190** (.090)	161
GPA in First UC writing class	2.930	-.287** (.039)	-.133** (.045)	976	2.958	-.126 (.119)	-.123 (.129)	106
Cum GPA of Students Graduating in 6 yrs.	3.170	-.189** (.023)	-.086** (.025)	1133	3.144	-.095 (.068)	-.097 (.073)	134

Notes: Each element of columns (2), (3), (6) and (7) represents a separate OLS regression with the coefficient on the "Remediation Required" variable reported. The conditional regressions in columns (3) and (7) include SAT I-Math, SAT I-Verbal, SAT II Writing, SAT II Math, AP Lit and Comp, AP Lang & Comp, high school GPA, parent's income, dummy variables for female, Asian, Hispanic, whether mom is a college grad, whether dad is a college grad, and missing dummy variables for each variable. In the regressions for GPA in 1st UCSD writing class, I exclude Warren College students because Warren's writing class graded pass/no pass. I also excluded students taking honors writing and students who withdrew from the class. Finally, *=10% significance and **=5% significance

Figure 1: The Scoring of the University of California Analytic Writing Placement Exam (AWP)
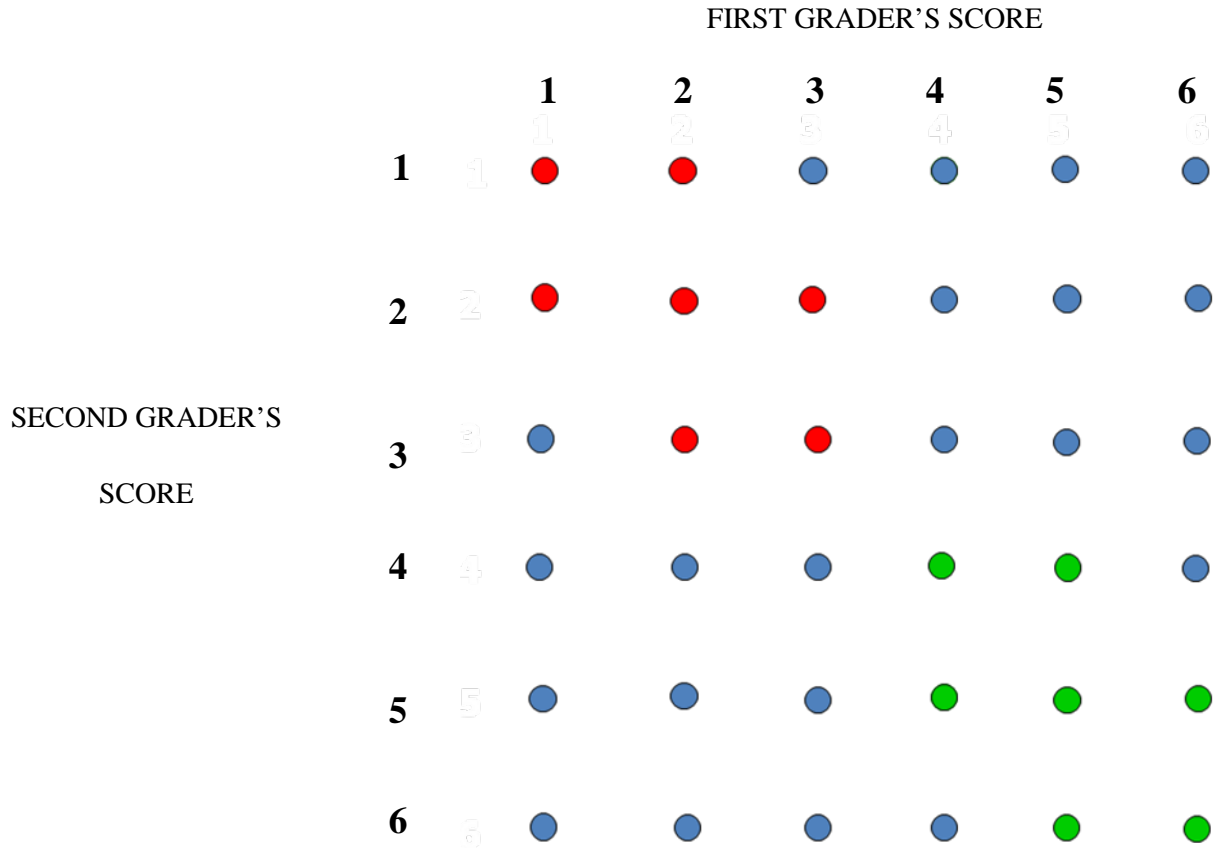
RED DOTS: Student fails based on first two grader's scores (sum of two scores is 6 or less)

GREEN DOTS: Student passes based on first two grader's scores (sum of two scores is 8 or more)

BLUE DOTS: Student's exam goes to a 3rd grader and student's score is 2*3rd grader's score.

FIGURE 2: The Distribution of Grader Failure Rates Across Students Required and Not Required to Take Remedial Writing

First Grader's Failure Rate
Calculated across 9 or more Exams

Remediation NOT Required   Remediation Required

First Grader's Failure Rate
Calculated across 6 or more Exams
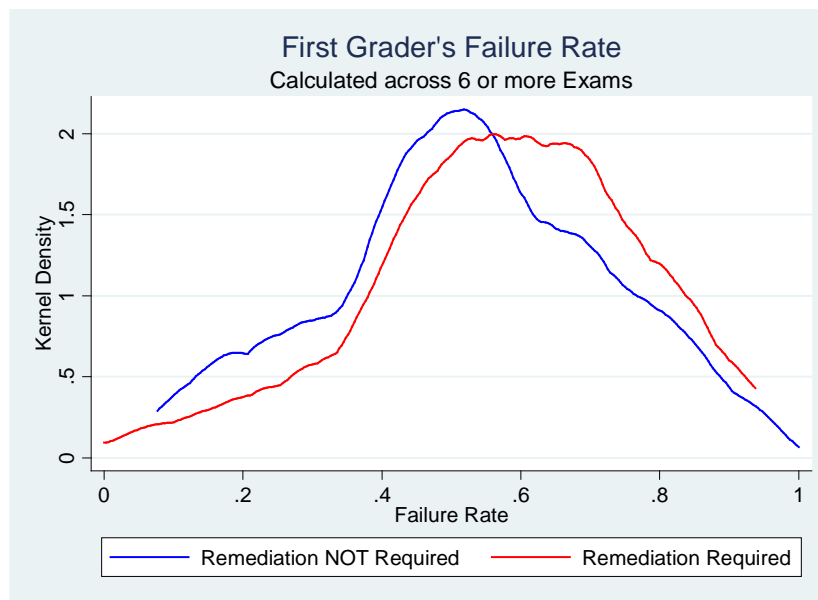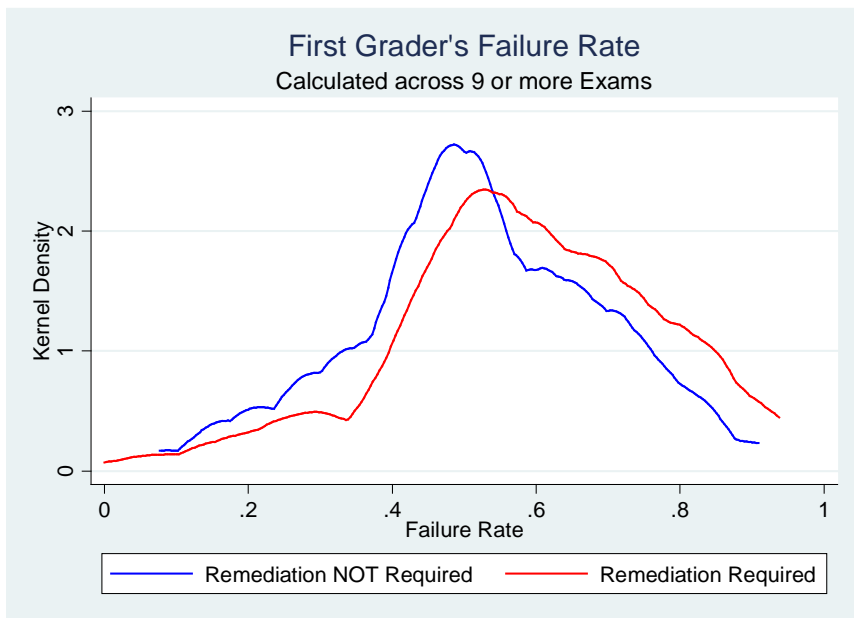
Remediation NOT Required   Remediation Required

FIGURE 3: The Likelihood of Remediation Given Grader Failure Rate

Likelihood of Remediation Given Grader Toughness
(5 or more students per failure rate)

Grader Failure Rate (Over 9+ Exams)

Likelihood of Remediation Given Grader Toughness
(5 or more students per failure rate)

Grader Failure Rate (Over 6+ Exams)