Assistance or Obstacle? The Impact of Different Levels of English Developmental Education on Underprepared Students in Community Colleges

Di Xu[1]

Developmental education is the most common approach used by community colleges to assist underprepared students for college-level course work. Yet there is limited evidence regarding this strategy on students assigned to the lowest level of the developmental sequence. This paper extends current knowledge on this critical question by examining the impacts of different levels of developmental reading and writing on students' academic outcomes. The results suggest that the impacts are generally insignificant for students on the margin of needing developmental course work, yet the estimates are negative on students assigned to the lowest level of the developmental sequence. The results therefore support the increasing national push to reform these programs.

Keywords: community colleges; econometric analysis; educational policy; educational reform; policy analysis; program evaluation; regression discontinuity

Due to their lower cost and flexibility, community colleges have increasingly become an entry point to higher education. Considering that students enrolled in community colleges are more likely than other postsecondary students to be minorities, low-income, and first in their families to attend college, community colleges assume a critical role in addressing the national equity agenda by providing access to higher education to students from traditionally underrepresented groups. Yet access without progress is no more than an empty promise. Many new college students arrive on campus lacking the preparation to successfully pursue their postsecondary education (Greene & Foster, 2003). Nationally, about two thirds of community college students are considered academically underprepared for college-level course work (Bailey, Jeong, & Cho, 2010). The most common approach that colleges use to address this widespread phenomenon has been to provide students who enter colleges with weak academic skills the opportunity to strengthen those skills and bring them up to an adequate level for further college-level course work, which is often termed as "developmental" or "remedial" education.

Despite the high hopes around developmental education, however, there is considerable uncertainty surrounding the effectiveness of this tactic. Although numerous studies have explored the impacts of developmental course work on students' academic outcomes, the majority of these studies drew inferences only on students scoring near the developmental course work assignment cutoff scores—that is, students who are on the margin of needing developmental education. In contrast, the impact of developmental education on students who are assigned to the lowest level of developmental sequence, those who are least academically prepared and who are most in need for academic support, has been largely left unexplored.

This study examines the causal impacts of different levels of reading and writing developmental course work on student academic outcomes. Taking advantage of the fact that during the time period of the study, the Virginia Community College System (VCCS) used standardized tests to place students into different levels of developmental course work, I used a regression discontinuity design to isolate the causal effects of different levels of developmental course work on various short-term and long-term outcomes.

[1]University of California, Irvine, CA

Educational Researcher, Vol. 45 No. 9, pp. 496–507
DOI: 10.3102/0013189X16683401

The results of the study therefore contribute to the ongoing discussion about the effectiveness of developmental education in several important ways. First, using a large administrative data set, this study adds evidence on developmental education in another particular state context. VCCS has recently undergone substantial reforms to its developmental education sequences, and the time period of this study occurred prior to those reforms, when the state's developmental education system was quite similar to those of other states nationwide (see Hodara, Jaggars, & Karp, 2012). As a result, the findings of the current study have important policy implications nationwide, especially with regard to the national push for developmental education reform, even when the policies under examination are no longer in place in Virginia. More importantly, this study is one of the first attempts to compare the academic outcomes of students assigned to long sequence of developmental education to students with similar academic skills but assigned to shorter developmental sequence. Therefore, results from the current study can inform ongoing national discussions on specific developmental education reform strategies, especially in terms of shortening the developmental sequence or speeding it up.

Theoretical Discussion and Empirical Evidence on the Impacts of Developmental Education

An educational intervention can affect students through multiple mechanisms, and some of them may be unintended. Scott-Clayton and Rodriguez (2015) propose three potential mechanisms through which developmental education courses may influence students' college outcomes in positive and negative ways: (a) improving college performance by developing important skills required for successful learning in college-level course work, (b) discouraging student persistence and progression by labeling a student as a poor performer, and (c) diverting students away from college courses.

A growing volume of studies have recently (e.g., Bettinger & Long, 2005; Boatman & Long, 2010; Calcagno & Long, 2008; Dadgar, 2012; Hodara, 2012; Lesik, 2007; Martorell & McFarlin, 2011; Scott-Clayton & Rodriguez, 2015) used quasiexperimental designs to draw causal inferences about the impact of developmental course work. Overall, these studies fail to find consistent positive impacts of developmental course work on students' academic outcomes. However, the majority of these studies focused on evaluating the effectiveness of developmental education on students at the margin of needing it and did not consider the effect of developmental education on students who were identified to have very low skills. As I will explain in more detail in the Data and Setting section, traditional developmental programs typically consist of a set of multiple courses that students must enroll in sequentially. As a result, students at the lowest levels are often required to complete at least three semesters of developmental course work for the corresponding subject area, compared to fewer developmental courses (or shorter sequence) for students assigned to the higher level. Yet whether increased dosage through lengthy developmental sequence could indeed benefit students with very low skills is an open question.

Only three studies to date (Boatman & Long, 2010; Dadgar, 2012; Hodara, 2012) explored the causal impact of developmental course work on students with much lower levels of preparation, and all used regression discontinuity designs. Using a large administrative data set from an anonymous community college system, Hodara (2012) focused on language-minority students and found that for students who just place into English remediation, developmental education does not help their college success, but for students with lower levels of academic preparation, assignment to two developmental English subjects versus one has a limited positive impact on their college outcomes. However, language-minority students constitute only a small proportion of community college students who are assigned to reading and writing developmental sequence, and therefore it is unclear whether the results are generalizable to the majority of students enrolled in English developmental courses. Dadgar (2012) used college administrative data from VCCS and observed null to negative impacts of assignment to longer math developmental sequence on a variety of student academic outcomes compared to assignment to shorter developmental math sequence. However, she did not explore the impacts of English developmental education in her study.

Boatman and Long (2012) is the only study to date that looked at all students who are assigned to higher-level and lower-level developmental English sequence. Focusing on undergraduates beginning at a public 2-year or 4-year college in Tennessee in fall 2000, they identified large negative effects on the margin of needing developmental course work but smaller and sometimes positive effects on students placed in lower-level developmental sequence compared to higher-level developmental sequence. Yet, they did not allow nonlinear terms in the model specification. As explained in detail in Lee and Lemieux (2010), when the true functional form between the outcome variable and the running variable (under either treatment or control) is nonlinear, running a simple linear local regression may result in large bias in the regression discontinuity (RD) estimates of the treatment effects. Therefore, it is recommended that researchers try more flexible specifications as robustness checks by either adding polynomials of the running variable in the local regression model or using nonparametric models to relax the linearity assumption, especially when the true function form between the running variable and the outcome variable is unknown. It is also questionable whether the positive findings regarding the long developmental sequence versus short sequence on least prepared students can be generalizable based on evidence from one single state. The current study responds to the national calls for more evidence on the impacts of developmental education by adding evidence on the impacts of different levels of reading and writing developmental sequence in a different state.

Data and Setting

Developmental Education in VCCS

Virginia is one of the several states that administer a multitiered statewide placement system to assign students to different levels of math, reading, and writing developmental courses. VCCS has undergone substantial reforms to its developmental education sequences beginning in 2013. The time period of this study occurred prior to the redesign of the developmental assessment and curriculum. Therefore, during the time period of this study, VCCS still used fairly traditional developmental education sequences similar to those of other states nationwide. Specifically,

the VCCS mandated that all of Virginia's 23 community colleges implement and use COMPASS as the primary tool to place students into multiple levels of reading, writing, and math courses with three possible options each: lower-level developmental courses, higher-level developmental courses, and college-level math or English courses. As a result, students put into the lower level of developmental course work need to complete a sequence of courses to proceed through the lower-level developmental course work and higher-level developmental course work before they can take the first college-level course in that subject area.

The criteria of developmental course work assignment vary from college to college, but the VCCS provides a narrow range for each level of developmental reading and writing within which colleges can set their own cut scores. Colleges may exempt student from taking COMPASS if the student demonstrates his or her readiness for college-level work by SAT scores or Advanced Placement scores. As a result, approximately one fifth of the students are missing placement test scores. These exempt students were excluded from the analysis.

Depending on the level of the developmental course work, courses are typically offered for credit in VCCS, but rarely are they counted toward graduation requirements. The lowest-level reading and writing courses covered basic writing and reading skills, whereas the highest-level reading and writing courses developed competencies in reading and writing necessary to succeed in college English. Although varied across schools, developmental courses are generally suffering from a high dropout rate and low completion rate; of those beginning in developmental mathematics, only 36% enrolled in a gatekeeper, college-level math course within 4 years (VCCS, Developmental Education Task Force, 2009). These high rates of failure in developmental courses and low likelihood of moving on to a college-level course lead many to wonder whether the developmental sequence might actually create barriers rather than assistance to students despite its initial intent. However, the descriptive results presented in these reports are confounded with students' academic capacity, thus disabling a conclusion regarding the causal impact of developmental course work.

Data Description

The primary data set includes approximately 46,000 students across 23 community colleges in Virginia who had valid information on college placement test scores on reading and writing. First-time students who initially enrolled during the summer or fall of 2004, 2005, and 2006 were tracked until the spring semester of 2011. The data include information on student demographics, institutions attended, developmental placement scores for reading and writing, transcript data on courses taken and grades received, and information on degree attainment. The 23 Virginia community colleges vary widely from one another in terms of institutional characteristics. The system comprises a mix of large and small schools as well as institutions located in rural, suburban, and urban settings. The appendix provides characteristics of the study sample in this state community college system compared to characteristics of a nationally representative sample of community college students. Overall, however, Virginia community colleges seem to represent a rural, low-income, underfunded, and African American student population based on statistics reported to the Integrated Postsecondary Education Data System. The largest difference across the samples is developmental education enrollment. Nearly 70% of the national 2003 cohort took at least one developmental education course, compared to 60% of the Virginia cohorts from 2004 to 2006.

Among all the 46,632 students who took both reading and writing COMPASS tests, 72% were college ready in reading whereas less than 60% were college ready in writing. Overall, 17% students were assigned to higher-level reading versus 10% to lower-level reading; 22% were assigned to higher-level writing versus nearly 20% to lower-level writing. The assignment distribution is fairly consistent across cohorts for both subjects, which is consistent with the report on developmental education in VCCS (VCCS, Developmental Education Task Force, 2009) that the developmental course work policy remained constant between 2003 and 2006.

Outcome Measures

The outcome measures explored in this study are divided into short- and long-term outcomes. One of the major criticisms leveled against developmental course work is that it imposes both economic and academic burdens on students by requiring a substantial amount of time spent on developmental course work that does not count toward a degree. Accordingly, of particular interest in the current literature on college developmental course work is whether these requirements are so burdensome that students become discouraged and drop out of college in early stages of their college career. I examined this possibility by including 1st-year dropout as a short-term outcome measure. Given that the purpose of developmental course work is to prepare students for college-level courses, I also explored whether receiving reading and writing developmental education would lead to higher probability of enrollment in the first college-level English course (ENG111) and, among those who enrolled in ENG111, whether taking developmental course work helped them pass the course. In terms of long-term outcomes, I examined the impacts of developmental course work on the total number of any types of credits earned within 5 years, total number of college-level credits earned in 5 years, and whether a student earned any degree or certificate or transferred to a 4-year school in 5 years.

One potential problem with the distal outcome measures is that developmental students will naturally spend more time in school due to extra remedial requirements, which therefore raises concerns of whether the time frame for the long-term outcomes is fair to developmental students. However, descriptive statistics suggest that this is a lesser concern in this particular setting: By the end of the 5-year tracking period, the majority of the students (more than 75%) included in the sample have already left college. That being said, in a separate robustness check, I limited the sample to the 2004–2005 cohort only and examined the long-term outcomes with different follow-up windows (5 years, 6 years, 7 years), and the estimates were fairly consistent.

Method

Addressing Ability Sorting: RD Strategy

In order to draw a causal inference regarding the different levels of reading and writing developmental course work on educational

outcomes, I used an RD design. Specifically, I exploited the fact that during the time period of the study, VCCS used standardized tests to place students into different levels of developmental course work. These mandatory cutoff scores created a discontinuity in the probability of receiving different levels of developmental treatment. The basic implementation of the RD design identifies the impact of developmental course work by comparing outcomes of students who score barely above the cutoff scores with those who score barely below; these students sharply differ in developmental course work assignment yet are otherwise very similar. As a result, the regression coefficient can be then interpreted as the causal impact of the intervention for students on the margin of passing the cutoff (Levin & Calcagno, 2008).

In terms of the function form for the local regression, I followed the recommendation by Lee and Lemieux (2010) by including not only placement test score in either reading or writing into the model but also polynomials to relax the nonlinearity assumption. Specifically, I include a squared term of the corresponding placement test score to address the possibility that the true relationship between the outcome variable and the running variable (either writing score or reading score) is nonlinear, an interaction term between placement test score and whether the student was directly assigned to college course work in that subject area (as opposed to higher-level developmental course work) to address the possibility that the relationship between the outcome measure and the running variable may not be the same for the treatment and control group, and an interaction term between the squared term of the placement test score and college course assignment to allow the quadratic term to vary between the treatment and control groups. In addition, I also included all available individual baseline characteristics into all the local regressions, which intends to improve the precision of the regression estimator by decreasing the standard error. Finally, I also included college fixed effects into the model (and clustered standard error at the institution level) to account for nonrandom clustering of students within a college.

Although it is straightforward to estimate the linear regressions within a given window of bandwidth around a cutoff point, a critical question is the selection of the bandwidth within which the analysis should be conducted. Lee and Lemieux (2010) specified the trade-off between precision and bias when finding an optimal bandwidth. On one hand, using a larger bandwidth yields more precise estimates; on the other hand, the linear specification is less likely to be accurate when a larger bandwidth is used, which can bias the estimate of the treatment effects. To identify the optimal bandwidth, I used the cross-validation procedure developed by Imbens and Lemieux (2008). The basic idea behind this procedure is to identify a bandwidth within which the functional form fits the data in an optimal way. Specifically, I estimated a linear regression to predict a given outcome variable within a set of different bandwidths. The bandwidth that minimizes the summation of the squared residuals then represents the best fit of the regression model to the data. The preferred bandwidth that I obtained using this particular procedure ranges depending on the cutoff explored and the outcome used, where most of them are around ±5 points. Accordingly, I reported results using a ±5-points bandwidth but conducted sensitivity analysis using half of the bandwidths (±2 points) and twice the bandwidths (±10).

Addressing Noncompliance: Fuzzy RD Design

The traditional RD method, known as "sharp RD," assumes full compliance with recommendations based on the test cutoff. In Virginia, however, not all students followed the assignment during the time period under study (Roksa, Jenkins, Jaggars, Zeidenberg, & Cho, 2009). Accordingly, the average probability of enrollment in developmental courses is less than 1 below the cutoff and more than 0 above the cutoff. To deal with potential bias associated with noncompliance, I followed existing literature for a "fuzzy RD" design (see Imbens & Lemieux, 2008, for a detailed discussion), using developmental course work assignments as instrumental variables for actual enrollment and employing a two-stage least squares strategy to provide a consistent estimate of the developmental course work on academic outcomes.

First-stage instrumental variable analyses revealed significantly positive relationships between being recommended to and actually enrolling in a given level of developmental course work for each subject area. The correlations between recommendation and enrollment are generally strong (0.60 for higher-level reading, 0.76 for lower-level reading, 0.74 for lower-level writing, 0.46 for higher-level writing), and the F tests for the first stage are all above 10, which rules out the problem of a weak instrument.

One important aspect of RD design, however, is that it provides estimates of the "local average treatment effects" (Imbens & Angrist, 1994) for a subpopulation around the cutoff points. Using an instrumental variable strategy, the fuzzy-RD design further restricts the relevant subpopulation to that of compliers of the developmental course work assignment. Although estimates from an RD design have been criticized for its limited external validity, the local average effects estimated by the current study are informative about the students whom the developmental course work policy intends to address. The results are therefore clearly relevant to remediation policies.

Validity of the RD Design

There are three testable assumptions underlying the validity of the RD design: (a) the probability of developmental enrollment should be discontinuous at the passing cutoff, (b) the expectations of pretreatment covariates should be continuous at the passing cutoff, and (c) there should be no discontinuity in density around the cutoff, which would otherwise suggest manipulation of assignment (Lee, 2008).

Figure 1 plots the likelihood of being in developmental course work in a certain level of either reading or writing as a function of the corresponding college pretest scores (centered to be zero at the passing cutoff). The four graphs clearly show a discontinuity at each passing cutoff for both reading and writing.

To test the second assumption, I conducted t tests to examine whether baseline characteristics exhibited significant difference around the cutoff, a test for random assignment around the discontinuity point (Imbens & Lemieux, 2008; Lee, 2008). If there

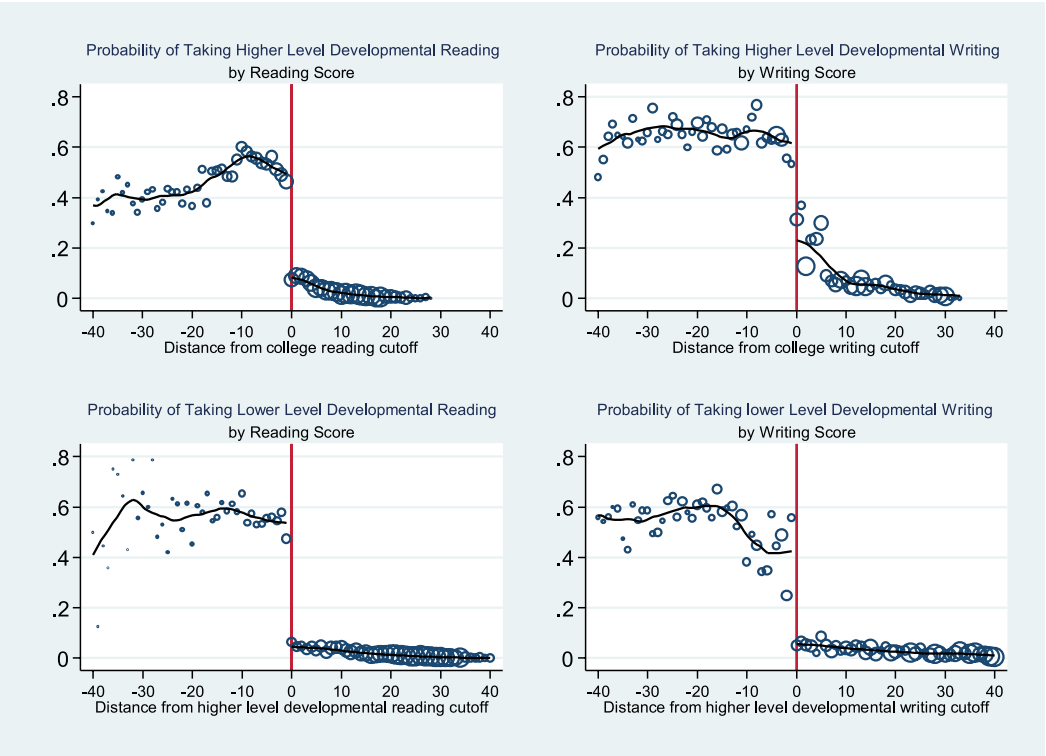

FIGURE 1. *Probability of developmental course work enrollment by placement test scores*
The graphs show possible discontinuity of students' probability of taking different levels of remedial reading and writing course work by students' placement reading and writing scores. All four graphs show clear discontinuity at the college cutoff score or at the higher-level remedial cutoff score.

were systematic sorting at the cutoff, one should observe significant differences in individual characteristics between students on the two sides of the score. The descriptive statistics are presented in Table 1. Comparisons between students just above (within a 5-point bandwidth) and just below (within a 5-point bandwidth) the cutoff revealed no significant differences in terms of most pretreatment characteristics for either reading or writing. Figure 2 further visually presents the distribution of pretest scores by baseline covariates around the cutoff score for college-level reading as an example. None of the eight baseline variables showed discontinuity around the cutoff score, and figures on other cutoff scores showed similar patterns.

Figure 3 examines assumption (c) by showing the distribution scores around each cutoff, and none of them showed clear discontinuity in distribution of students above and below the score. I also conducted a McCrary test (McCrary, 2008) to statistically check whether there is a discontinuity in the density of observations at the cutoff, and the test failed to reject the null. These results support the employment of the RD design in examining the causal impact of different levels of reading and writing developmental course work on educational outcomes.

Finally, one potential challenge to using RD design in this particular context is that students might be assigned to multiple developmental subject areas, including reading, writing, and math. Following the suggestion by Kane (2003) and Reardon and Robinson (2010), I adopted a "frontier RD" approach and subset the data by status on all but just one of the rating scores. For example, to estimate the treatment effect of higher-level remedial reading versus college-level reading, I limited the sample to students assigned to college-level writing and college-level math. The analysis based on this restricted sample would then address potential issues of receiving multiple developmental treatments. The results were fairly consistent with the results presented in Table 2 and Table 3.

Results

Impacts of Higher-Level Developmental Education (vs. No Developmental Education)

Tables 2 reports the impacts of higher-level reading (upper panel) and higher-level writing (bottom panel) on a variety of outcomes within ±5-points bandwidths. In each panel, the first row reports the intent-to-treat estimate, which measures the average differences between individuals above and those below the cutoff score controlling for baseline characteristics (covariates included in the model are listed below Table 2); the second row reports estimates using the fuzzy-RD design.

The intent-to-treat estimates are generally nonsignificant, and among the few that show significantly negative impacts of higher-level reading or writing on student outcomes, these differences generally vanish after using fuzzy RD to address noncompliance. Focusing on the estimates based on the fuzzy RD, higher-level developmental course work in both subjects has nonsignificant impact on any of the six academic outcomes except for earning a degree or transfer to a 4-year university in 5 years, where enrollment in higher-level reading developmental course work reduces

Table 1
Descriptive Statistics of the Sample Around Cutoffs: Group Means and Group Differences

	Reading						Writing					
	Band Around College-Ready Cutoff: ±5 points			Band Around Higher-Level Developmental Course Work Cutoff: ±5 Points			Band Around College-Ready Cutoff: ±5 Points			Band Around Higher-Level Developmental Course Work Cutoff: ±5 Points		
Variable	Below	Above	Difference	Below	Above	Difference	Below	Above	Difference	Below	Above	Difference
Demographic characteristics												
Age at college entry	20.976	20.847	0.130	21.156	21.055	0.101	20.830	20.586	0.244	21.371	21.616	0.245
Female	0.599	0.587	0.012	0.615	0.622	0.007	0.542	0.527	0.015	0.542	0.542	<0.001
White	0.536	0.577	0.041**	0.440	0.463	0.023	0.612	0.634	0.022	0.506	0.507	0.001
Black	0.323	0.279	0.044**	0.420	0.391	0.029	0.219	0.204	0.015	0.319	0.359	0.040
American Indian	0.007	0.005	0.002	0.003	0.005	0.002	0.005	0.004	0.001	0.006	0.004	0.002
Asian	0.061	0.065	0.004	0.077	0.066	0.011	0.069	0.080	0.011	0.086	0.053	0.032**
Hispanic	0.073	0.075	0.002	0.073	0.062	0.011	0.082	0.086	0.004	0.082	0.077	0.005
Academic characteristics												
Dual enrolled prior to entry	0.059	0.067	0.008	0.037	0.054	0.017**	0.059	0.047	0.011	0.042	0.038	0.004
Transfer track (vs. career-technical)	0.569	0.585	0.016	0.536	0.558	0.023	0.591	0.590	0.001	0.553	0.584	0.031
Eligible for financial aid	0.381	0.369	0.013	0.460	0.439	0.021	0.322	0.298	0.024	0.418	0.471	0.053**
Observations	3,661	5,378	9,039	1,215	1,907	3,122	2,219	2,927	5,146	842	1,185	2,027

**$p < .05$.

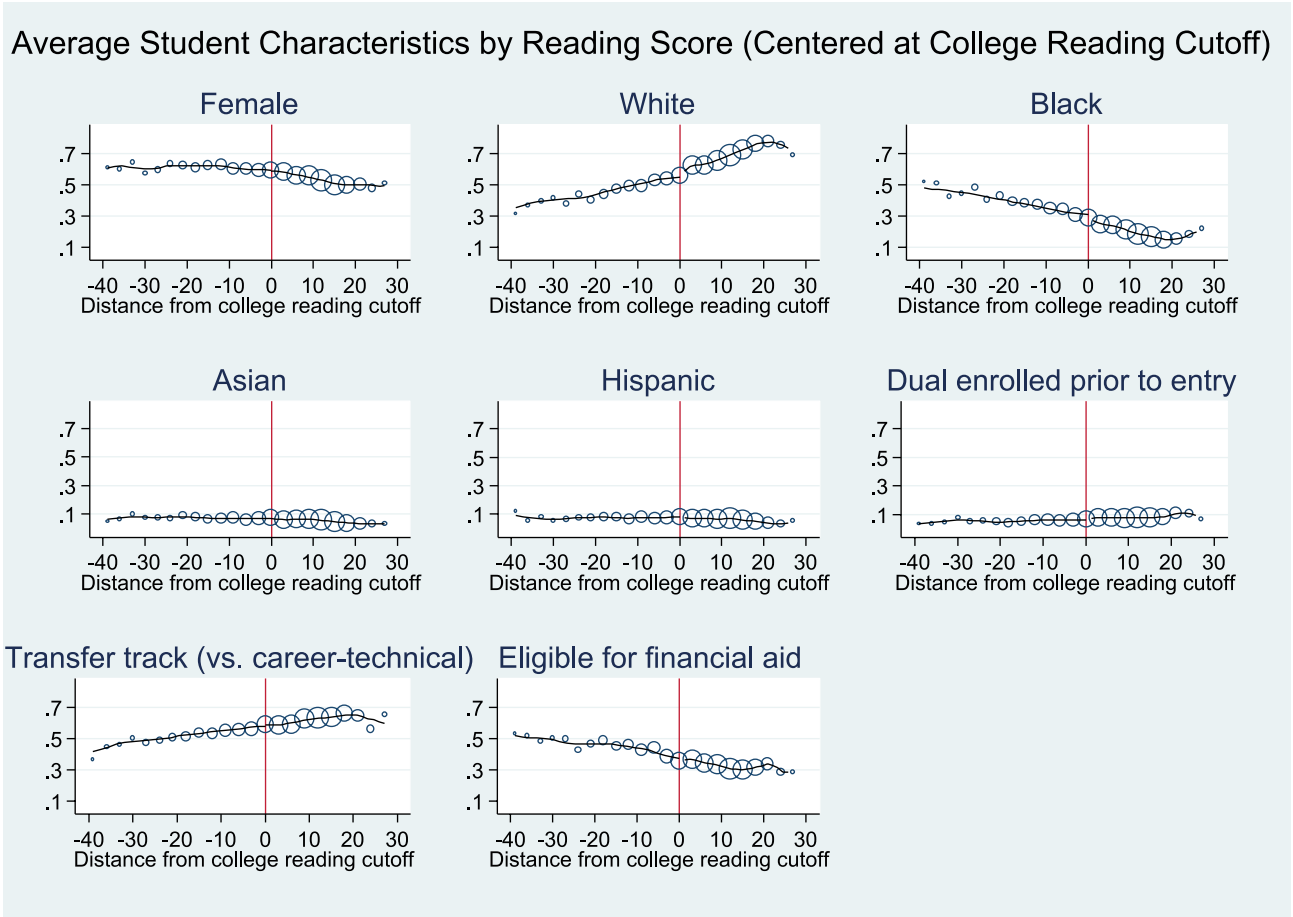

FIGURE 2. *Regression discontinuity validity check: Reading placement test score distribution by pretreatment individual characteristics* The graphs show possible discontinuity of students' baseline characteristics at college reading cutoff, including gender (female), race dummies (White, Black, Asian, Hispanic), whether dual enrolled prior to college, in a transfer-track program versus a career-technical program, and whether eligible for need-based financial aid. None of the characteristics show clear discontinuity at the cutoff.

the probability of achieving this outcome by around 10 percentage points. However, this negative impact is significant only at the 0.1 level, and the corresponding estimate is not significant even at the 0.1 level for high-level writing. Accordingly, the results generally support the current literature (e.g., Calcagno & Long, 2008; Martorell & McFarlin, 2011) that developmental course work has little impact on academic outcomes for students on the margin of needing developmental course work.

Impacts of Lower-Level Developmental Education (vs. Higher Level)

Table 3 reports results of the impact of lower-level reading compared to higher-level reading (upper panel) and lower-level writing compared to higher-level writing (bottom panel) on educational outcomes. In contrast to the generally nonsignificant patterns across estimates for higher-level developmental course work versus no developmental education, lower-level developmental sequence shows substantial and significantly negative impacts compared to higher-level developmental sequence on several academic outcomes, particularly for lower-level developmental reading, where all the estimates are significant except for passing the English gatekeeper course.

Specifically, lower-level developmental course work in both subjects shows a significant negative impact on 1st-year retention rate, where taking lower-level reading versus higher-level reading increases the probability of dropping out of college within the 1st year by 13 percentage points; similarly, taking lower-level writing developmental course work versus higher level increases the 1st-year dropout rate by 19 percentage points. It is worth mentioning, though, that both estimates are significant only at the 0.1 level.

Associated with this negative impact of developmental course work on 1st-year retention is a reduced probability of ever attempting the English gatekeeper course. Students who barely passed lower-level reading were more likely to enroll in the English gatekeeper course by 16 percentage points, which is significant at the 0.05 level. I then examined whether taking developmental course work can improve the probability of passing English gatekeeper course among those who enrolled in the course. The estimates are consistently small and insignificant for both reading and writing.

In terms of distal outcomes, enrollment in longer reading developmental sequence versus shorter sequence reduces both total credits and college-level credits in 5 years by about 10 points, both of which are significant at the 0.05 level. Taking

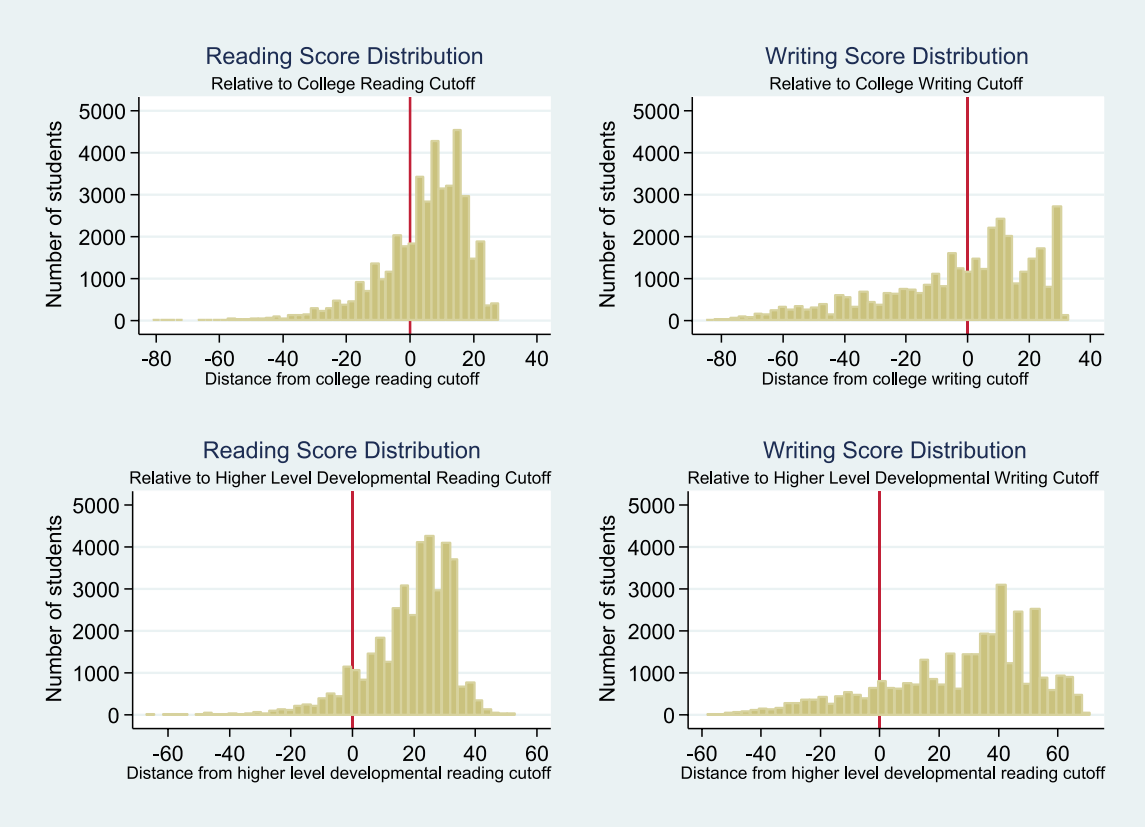

FIGURE 3. *Regression discontinuity validity check: Density of observations around cutoff*

longer developmental reading also reduces the likelihood of earning a degree or transferring to a 4-year university in 5 years by 14 percentage points ($p < .05$). As for writing developmental course work, the estimates for all the three long-term outcomes were also negative, although only the impact on the total college-level credits is significant at the 0.1 level.

One concern, however, is that the negative impacts of lower-level developmental course work on long-term outcomes are mainly due to its negative impact on early college dropout. I explored this possibility by estimating the long-term impact on a reduced sample where students persisted into the 2nd year. However, the negative impacts persisted for this subgroup of students, suggesting that assignment to lower-level developmental course work has a long-lasting negative effect on individual academic progress even beyond early stages of one's college career. The negative impacts on degree completion and transfer are particularly alarming, given that research has generally shown that completion of a certificate, an associate degree, or transfer to a higher-level college has positive effects on earnings (Jaeger & Page, 1996; Kane & Rouse, 1999).

Heterogeneous Impacts of Developmental Course Work

To examine whether the aggregate estimates may mask benefits for certain groups of students, I examined the effect of developmental course work by gender, age, and race. Specifically, I first included interaction terms between each of the key demographic variables with a specific level and subject of the developmental course work explored. The interaction terms were generally nonsignificant for higher-level developmental course work. In contrast, the interaction terms were all significant for lower-level reading and writing. I therefore further conducted analysis separately for each subgroup for lower-level reading and lower-level writing. Detailed results on lower-level reading developmental course work are presented as an illustration in Table 4, and the results on lower-level writing show similar patterns.

The results indicate that the negative impacts of enrolling in longer developmental sequence in reading were mainly driven by the influence on females, younger students, and Black students. A possible explanation for such variation is that students along different demographic lines might react in different ways toward the unintended negative impacts of developmental course work academically, economically, and psychologically. For example, the economic burden might be less strong for older students (students who entered college after 25) who are more likely to be economically independent than younger students. These findings lend support to the argument (e.g., Angrist, Lang, & Oreopoulos, 2009) that college-level interventions typically vary substantially among demographic lines, and neglecting the heterogeneity of developmental course work among different types of students can overlook important policy implications.

Discussion and Conclusion

The effects of developmental course work on helping underprepared students succeed in college are of great interest to policymakers, school administrators, and taxpayers. Using an RD design, this study extends current understanding of college

Table 2
Impacts of Higher-Level Reading and Writing Developmental Course Work on Academic Outcomes (Bandwidth: ±5 Points)

	Short-Term Impact			Long-Term Impact		
Variable	Dropped Out Within the 1st Year	Took English Gatekeeper	Complete English Gatekeeper	Total Credits Earned in 5 Years	Total College-Level Credits in 5 Years	Earned Degree or Transfer to 4-Year University in 5 Years
Enrollment in higher-level reading developmental course work						
Intent-to-treat estimates	–0.000 (0.011)	–0.010 (0.017)	0.004 (0.022)	0.937 (1.201)	–0.636 (1.119)	–0.037* (0.020)
Instrumental variable estimates	–0.001 (0.050)	–0.026 (0.047)	–0.010 (0.063)	2.495 (3.221)	–1.632 (3.007)	–0.098* (0.053)
Observations	9,039	9,039	6,054	9,039	9,039	9,039
Enrollment in higher-level writing developmental course work						
Intent-to-treat estimates	0.028** (0.012)	–0.056*** (0.012)	0.013 (0.015)	0.035 (0.792)	–0.744 (0.749)	–0.012 (0.014)
Instrumental variable estimates	0.020 (0.096)	–0.080 (0.093)	–0.017 (0.087)	–6.181 (6.283)	–8.105 (5.948)	–0.098 (0.107)
Observations	5,146	5,146	3,540	5,146	5,146	5,146

Note. Each cell represents a separate regression within a 10-point band. Each analysis includes college fixed effects and the following covariates: placement test score in either reading or writing, squared term of the corresponding placement test score, an interaction term between placement test score and whether the student was directly assigned to college course work in that subject area (as opposed to higher-level developmental course work), an interaction term between the squared term of the placement test score and college course assignment, a gender dummy variable, race dummy variables, cohort dummy variables, a dummy variable for receiving need-based federal financial aid, a dummy variable for whether in a transfer-track program (vs. career-tech track), and a dummy variable for dual enrollment prior to college. Standard error is clustered at the college level.
*$p < .10$. **$p < .05$. ***$p < .01$.

developmental course work by exploring the impacts of different levels of developmental English course work on students' short-term and long-term academic outcomes in an entire community college system. The results suggest that developmental courses do differ in their impacts by the level of assignment. Specifically, although the estimated effects are generally small in magnitude and statistically insignificant for students on the margin of needing developmental course work, lower-level and therefore longer developmental sequence versus higher level (or shorter sequence) imposes negative impacts on various academic outcomes. Heterogeneity analysis suggests that such negative impacts are mainly driven by influences on females, younger students, and Black students.

These results have several important policy implications. First, results regarding the robust negative impacts of the longer developmental sequence versus shorter sequence lend support to the increasing public concerns of whether the economic, psychological, and academic burdens imposed by these lengthy developmental sequences might in fact outweigh their intended benefits. Considering that optimizing students' college retention is an imperative when it comes to economic opportunity for disadvantaged students, of greater concern is the finding that taking the longer developmental sequence may increase the chance that the students drop out completely from college during their early academic career. Because community colleges assume a critical role in addressing the national equity agenda by disproportionately enrolling students from historically underrepresented groups, the negative impacts of the longer developmental sequence on student college persistence would imply that the economic, academic, and psychological burdens imposed by the traditional developmental sequence may harm not only these students' educational outcomes but also their labor market opportunity.

Current English acceleration reforms that have been in place across all Virginia community colleges since spring 2013 may help ameliorate the heavy burdens of developmental requirement on students. Specifically, the developmental redesign involved combining reading and writing sequences into a single, shorter developmental English sequence; aligning the developmental English curriculum to college-level English; reducing the amount of time needed to complete developmental English requirements; and allowing students who place into the highest developmental course to take college English and developmental English concurrently. Similar models in other states have had a positive impact on students' likelihood of completing college English and college credit accumulation (Jaggars, Hodara, Cho, & Xu, 2015). Moreover, there have also been statewide efforts to reduce the proportion of students testing into remediation, which entails improving placement accuracy and tailoring developmental education requirements to students' academic goals, thus eliminating unnecessary prerequisites (Kalamkarian, Raufman, & Edgecombe, 2015). Future studies may wish to directly compare different developmental assignment policies and curriculum on students' academic performance.

In addition, besides the current study, Boatman and Long (2010) provides the only other causal estimate in the literature that explores the impacts of different levels of developmental English on all students. However, the results for their study and the current study are substantially different from each other: They identified large negative effects on the margin of needing

Table 3
Impacts of Lower-Level Reading and Writing Developmental Course Work on Academic Outcomes (Bandwidth: ±5 Points)

	Short-Term Impact			Long-Term Impact		
Variable	Dropped Out Within the 1st Year	Took English Gatekeeper	Complete English Gatekeeper	Total Credits Earned in 5 Years	Total College-Level Credits in 5 Years	Earned Degree or Transfer to 4-Year University in 5 Years
Enrollment in lower-level reading developmental course work						
Intent-to-treat estimates	0.061* (0.033)	–0.069** (0.033)	–0.015 (0.044)	–4.882** (2.021)	–4.378** (1.794)	–0.065** (0.032)
Instrumental variable estimates	0.133* (0.074)	–0.162** (0.072)	–0.047 (0.110)	–10.720** (4.492)	–9.612** (3.954)	–0.144** (0.070)
Observations	3,122	3,122	1,620	3,122	3,122	3,122
Enrollment in lower-level writing developmental course work						
Intent-to-treat estimates	0.036* (0.021)	–0.012 (0.215)	0.020 (0.027)	–1.027 (1.309)	–2.053* (1.173)	–0.022 (0.021)
Instrumental variable estimates	0.188* (0.110)	0.082 (0.113)	–0.044 (0.146)	–8.509 (6.832)	–10.261* (6.130)	–0.144 (0.108)
Observations	2,027	2,027	1,056	2,027	2,027	2,027

Note. Each cell represents a separate regression within a 10-point band. Each analysis includes college fixed effects and the following covariates: placement test score in either reading or writing, squared term of the corresponding placement test score, an interaction term between placement test score and whether the student was assigned to higher-level developmental course work in that subject area (as opposed to lower-level developmental course work), an interaction term between the squared term of the placement test score and developmental course work assignment, a gender dummy variable, race dummy variables, cohort dummy variables, a dummy variable for receiving need-based federal financial aid, a dummy variable for whether in a transfer-track program (vs. career-tech track), and a dummy variable for dual enrollment prior to college. Standard error is clustered at the college level.
*$p < .10$. **$p < .05$.

Table 4
Heterogeneous Effects of Lower-Level Reading Developmental Course Work by Subgroup (Bandwidth: ±5 Points)

	Short-Term Impact			Long-Term Impact		
Variable	Dropped Out After the 1st Year	Took English Gatekeeper	Complete English Gatekeeper	Total Credits Earned in 5 Years	Total College-Level Credits in 5 Years	Earned Degree or Transfer to 4-Year University in 5 Years
Gender						
Female (n = 1,933)	0.112	–0.265***	–0.181	–15.214***	–14.152***	–0.244***
	(0.091)	(0.088)	(0.134)	(5.782)	(5.073)	(0.090)
Male (n = 1,189)	0.181	0.011	0.254	–4.308	–2.612	0.018
	(0.127)	(0.129)	(0.220)	(7.324)	(6.503)	(0.113)
Age when started college						
Above or equal to 25 (n = 391)	–0.133	–0.154	0.311	–10.943	–7.051	–0.239
	(0.282)	(0.293)	(0.856)	(15.016)	(12.971)	(0.244)
Below 25 (n = 2,731)	0.186**	–0.138*	–0.091	–12.622***	–11.495***	–0.145**
	(0.076)	(0.074)	(0.116)	(4.715)	(4.147)	(0.073)
Race						
White (n = 1,418)	0.171	–0.135	–0.042	–8.780	–7.997	0.015
	(0.126)	(0.121)	(0.204)	(8.057)	(7.217)	(0.119)
Black (n = 1,247)	0.076	–0.220*	–0.019	–10.732*	–9.358*	–0.250**
	(0.118)	(0.116)	(0.212)	(6.502)	(5.657)	(0.109)

Note. Each cell represents a separate regression using instrumental variable approach within a 10-point band (±5 points). Each analysis includes college fixed effects and the following covariates: placement test score in either reading or writing, squared term of the corresponding placement test score, an interaction term between placement test score and whether the student was directly assigned to college course work in that subject area (as opposed to higher-level developmental course work), an interaction term between the squared term of the placement test score and developmental course work assignment, a gender dummy variable, a gender dummy variable, race dummy variables, cohort dummy variables, a dummy variable for receiving need-based federal financial aid, a dummy variable for whether in a transfer-track program (vs. career-tech track), and a dummy variable for dual enrollment prior to college. Standard error is clustered at the college level.
*$p < .10$. **$p < .05$. ***$p < .01$.

developmental course work but smaller and sometimes positive effects on students placed in lower-level developmental sequence compared to higher-level developmental sequence. Yet the specific mechanisms driving such between-state differences are largely unknown. The two systems may differ in multiple ways, including student characteristics, developmental education assignment and enforcement policies, the specific developmental education curriculum, and student academic support, all of which may lead to differences in the estimated impacts of developmental education. These sharp between-state differences in the impacts of developmental course work on students underscore the importance of evaluating developmental education programs relative to the institutional context rather than solely relying on evidence from other states. Therefore, these between-state differences highlight the necessity for future studies to uncover how specific developmental program design features may influence student outcomes and draw attention to the importance for colleges and systems to conduct their own analyses to assess the effectiveness of developmental course work when contemplating program elimination and reform.

Finally, the heterogeneous impacts of developmental education across subgroups of students deserve policy attention. In particular, the negative impacts of being assigned to the lowest level of English developmental sequence are especially strong on Black students. This is troubling from an equity perspective: If this pattern holds true across other state community systems, it would imply that the educational achievement gaps between key demographic groups, such as between White and ethnic-minority students, are exacerbated with the implementation of the developmental sequence. Yet, there is no definite answer as to the mechanism by which these heterogeneous effects may operate. One possible explanation is that the economic and psychological burden of developmental education might be stronger to students who are under greater economic constraint and have limited access to academic resources. Therefore, as colleges move forward with the reform movement, it would be imperative that stakeholders involved in the planning, teaching, or supervision of developmental education recognize the additional challenges of completing developmental course work and develop effective strategies to better support struggling students in these programs.

NOTE

The research reported here was supported by the Bill & Melinda Gates Foundation to Teachers College, Columbia University. The opinions expressed are those of the author and do not represent views of the Bill & Melinda Gates Foundation. I would like to thank all the staff at the Community College Research Center for their support during this research project. I am grateful to Thomas Bailey, Clive Belfield, Jonah Rockoff, Aaron Pallas, Shanna Smith Jaggars, Michelle Hodara, and Judith Scott Clayton for their valuable comments and suggestions about this research. I also thank the 23 colleges for providing high-quality data and expert guidance on the college contexts for this research.

REFERENCES

Angrist, J., Lang, D., & Oreopoulos, P. (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, *1*(1), 136–163.

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

Bettinger, E., & Long, B. T. (2005). *Addressing the needs of under-prepared students in higher education: Does college developmental coursework work?* (NBER Working Paper No. 11325). Cambridge, MA: National Bureau of Economic Research.

Boatman, A., & Long, B. T. (2010). *Does developmental coursework work for all students? How the effects of postsecondary developmental and developmental courses vary by level of academic preparation* (NCPR working paper). New York, NY: National Center for Postsecondary Research.

Calcagno, J. C., & Long, B. T. (2008). *The impact of postsecondary developmental coursework using a regression discontinuity approach: Addressing endogenous sorting and noncompliance*. Mimeo.

Dadgar, M. (2012). *Essays on the economics of community college students' academic and labor market success* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3506175)

Greene, J., & Foster, G. (2003). *Public high school graduation and college readiness rates in the United States* (Education Working Paper No. 3). New York, NY: Manhattan Institute, Center for Civic Information.

Hodara, M. (2012). *Language minority students at community college: How do developmental education and English as a Second Language affect their educational outcomes?* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3505981)

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

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrics*, *61*(2), 467–476.

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

Jaeger, D., & Page, M. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *Review of Economics and Statistics*, *78*(4), 733–740.

Jaggars, S. S., Hodara, M., Cho, S. W., & Xu, D. (2015). Three accelerated developmental education programs: Features, student outcomes, and implications. *Community College Review*, *43*, 3–26.

Kalamkarian, J. S., Raufman, J., & Edgecombe, N. (2015). *Statewide developmental education reform: Early implementation in Virginia and North Carolina*. New York, NY: Community College Research Center, Columbia University. Retrieved from http://ccrc.tc.columbia.edu/media/k2/attachments/statewide-developmental-education-reform-early-implementation.pdf

Kane, T. J. (2003). *A quasi-experimental estimate of the impact of financial aid on collegegoing* (NBER Working Paper No. 9703). Retrieved from http://www.nber.org/papers/w9703

Kane, T. J., & Rouse, C. (1999). The community college: Education students at the margin between college and work. *Journal of Economic Perspectives*, *13*(1), 63–84.

Lee, D. (2008). Randomized experiments from non–random selection in US house elections. *Journal of Econometrics*, *142*, 675–697.

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

Lesik, S. (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*, *48*(5), 566–591.

Levin, H., & Calcagno, J. C. (2008). Developmental coursework in the community college: An evaluator's perspective. *Community College Review*, *35*, 181–207.

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

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

Reardon, S. F., & Robinson, J. P. (2010). *Regression discontinuity designs with multiple rating-score variables*. Stanford, CA: Center for Education Policy Analysis. Retrieved from https://cepa.stanford.edu/sites/default/files/reardon%20robinson%20multiple%20rating%20score%20RD%20paper%20july%202010.pdf

Roksa, J., Jenkins, D., Jaggars, S. S., Zeidenberg, M., & Cho, S. (2009). *Strategies for promoting gatekeeper success among students needing remediation: Research report for the Virginia Community College System*. New York, NY: Columbia University, Teachers College, Community College Research Center.

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

AUTHOR

DI XU, PhD, is an assistant professor in the School of Education at the University of California–Irvine, 3200 Education Building, Irvine, CA 92697; *dix3@uci.edu*. Her current research focuses on using an economic lens to understand factors and policies that improve students' access to and success in college, especially among historically underrepresented groups.

Manuscript received October 17, 2015
Revisions received June 9, 2016, and October 28, 2016
Accepted November 18, 2016

Appendix
Characteristics of Study Sample Compared to Nationally Representative Sample of Community College Students

Characteristic	National Sample[a]	Virginia Community Colleges
Sample size	7,095	46,362
College entry term	Fall 2003	Fall 2006-2008
Gender		
Female	57%	56%
Male	43%	44%
Race-ethnicity		
Black	14%	24%
Hispanic	16%	6%
White	61%	64%
Other	9%	6%
Age upon entry		
Average age	24	20
18 or younger	34%	57%
19	22%	14%
20–23	17%	12%
24–29	9%	8%
30 or older	18%	9%
Developmental education enrollment		
Took any developmental education	68%	60%
Took developmental math	60%	49%
Took developmental reading/writing	17%	38%

[a]Author derived data from the U.S. Department of Education, National Center for Education Statistics (NCES), BPS: 2009 Beginning Postsecondary Students study using the NCES QuickStats tool. BPS: 2009 contains student-level data on a nationally representative sample of students who enrolled in college for the first time in 2003–2004, tracked to 2009. I report data on students who started in a public, 2-year college only. Sample size is approximate because BPS: 2009 reports approximate sample sizes.