A Meta-Analysis of Regression Discontinuity Studies Investigating the Effects of Placement into Developmental Education: A Working Paper

Jeffrey C. Valentine
University of Louisville

Spyros Konstantopoulos
Michigan State University

Sara Goldrick-Rab
Temple University

November 2016

This is a working paper. Please let us know if you spot any errors or have suggestions for improvement (send to jeff.valentine@louisville.edu).

The work reported herein was supported in part by the U.S. Department of Education's Institute of Education Sciences (contract ED–IES–12–C–0084). However, the views expressed are the authors' and do not necessarily represent the positions or policies of the Institute of Education Sciences or the U.S. Department of Education. Attendees at the *Society for Research Synthesis Methods* annual meeting (July, 2015) provided helpful feedback on portions of this manuscript.

Abstract

This working paper reports a systematic review and meta-analysis of studies that use regression discontinuity to examine the effects of placement into developmental education. Results suggest that placement into developmental education is associated with effects that are negative, statistically significant, and substantively large for three outcomes: (a) the probability of passing the college level course in which remediation was needed, (b) college credits earned, and (c) attainment. Several sensitivity analyses suggest these results are not a function of particular stylized studies, or the choices made in assembling the meta-analytic database. Two exploratory moderator analyses suggest that the negative effects of placement into developmental education are stronger for university students than for community college students, and worse for students placed in reading or writing than in math. This work can inform debate and research on postsecondary policies and on alternative mechanisms for ensuring that college students have the skills needed to meet their goals.

What Happens to Students Placed into Developmental Education?: A Meta-Analysis of Regression Discontinuity Studies

Almost two in five beginning college students are placed in developmental education (National Center for Educational Statistics, 2016). Broadly speaking, the term "developmental education" connotes a set of policies and practices designed for students who are underprepared to do college level work in a given area. The goal of this experience is to give students the knowledge, skills, and habits that will help them be successful in the college level version of the course. The growing use of developmental education reflects an increasingly normative transition from high school to college, which while predicated on completion of secondary schooling does not necessarily imply adequate preparation for what is deemed "post-secondary" work.

The specific mechanisms for deciding which students should be placed into developmental education vary. These policies are sometimes set at the state level (as is true in Florida), sometimes set at the system level (as in the California State University system), and sometimes set by individual institutions. Community colleges and other open access institutions generally require all students to take placement exams. Institutions that require an entrance exam like the SAT or ACT often use a tiered system. For example, in Tennessee (Boatman & Long, 2010) students scoring below 26 on the ACT's math subtest (approximately two-thirds of all test takers score below this threshold) are required to take the COMPASS Algebra test, a placement exam developed by ACT. Students scoring 50 and above are placed into college algebra, while students scoring below 50 are placed into intermediate algebra. Depending on the specific policy in place, students may or may not be able to retake the placement exam.

Nationally, about 60% of students taking a placement exam are recommended for placement into developmental education, but not all students recommended for placement actually end up in the courses (see Bailey, 2009). Rates of developmental course taking are somewhat higher in community colleges (about 42%) than in public and private doctoral degree granting institutions (about 25 and 22%, respectively), but even in these latter institution types developmental course taking is common. Math is the most common subject in which remediation is needed, with participation rates (about 15%) that are two to two and a half times the participation rates in English, reading, and writing (which range from 6 to 7%).

Placement into developmental education adds costs and, critically, time to a student's journey to a degree or certificate. With respect to student costs, Barry and Dannenberg (2016) estimate that each developmental course costs students $3,000 and adds $1,000 in student loan debt (and this analysis did not include the opportunity costs that students experienced). In addition, states are increasingly concerned about "paying twice" for courses taken both in high school and in college. Nationwide, Breneman and Haarlow (1998) estimate cost of development education to be $1 billion at public postsecondary institutions in 1996 dollars, while Pretlow and Wathington (2012), using similar methodologies, arrived at about $1.13 billion (again in 1996 dollars) for the 2004-2005 year. More recently, Barry and Dannenberg (2016) put the estimate at $1.5 billion.

What are the effects of utilizing developmental education when it comes to college academic outcomes? Given the high personal and societal costs of developmental education, the effectiveness of developmental education has become an important public policy question that has spurred both research and reform efforts (e.g., Complete College America, 2012). Most simple comparisons of students assigned to developmental education relative to those not assigned suggest that assignment to developmental education is associated with several negative outcomes, not least of which is a much lower likelihood of postsecondary attainment (i.e. graduation or certification). For example, using data from the National Educational Longitudinal Survey (NELS) from the 1992 high school class, Attewell, Lavin, Domina, and Levey (2006) found that for students attending 2-year colleges graduation rates were about 30% lower for students who enrolled in at least one developmental education course than students who did not (28% versus 36%). For students attending 4-year institutions, the picture is even bleaker, with students enrolling in at least one developmental course graduating at a much lower rate (52%) than students not enrolling in developmental courses (77%). But it is far from clear whether these lower completion rates are *caused* by development education. We wrote this paper, which reports the results of a state-of-the-art systematic review and meta-analysis on the effects of placement into developmental education, in an attempt to address this question. We examine the effects of placement on four indicators of college attainment, including credit accumulation and degree or certificate completion.

Studying the Effects of Placement into Developmental Education

Some of the observed differences in outcomes between students placed into developmental education in at least one subject and students not placed into developmental education are real in the sense that they reflect different levels of academic opportunities, preparation, and motivation. However, the raw statistics do little to untangle the causal effects of being placed into developmental education. There are two aspects to this problem. One is the distinction between enrollment and assignment. Attewell et al.'s (2006) data point to the negative association between enrollment in developmental education and attainment, but some students assigned to developmental education never take a developmental education course, either because they somehow avoid the placement decision and go directly into the college level course, or because they take assignment to developmental education as a signal that they are unlikely to succeed in college and drop out (see Bailey, Jeong, & Cho, 2010; Scott-Clayton & Rodriguez, 2012); if true, this suggests that Attewell et al.'s analysis understates the negative impact of assignment to developmental education.

The second part of the problem is untangling the causal relationships. To test the effect of assignment to developmental education, researchers could identify a group of students for whom an institution's policy suggests developmental education is needed, and randomly recommend students for placement into either the developmental course or into the college level course in the subject in which remediation is needed. For example, Aiken, West, Schwalm, Carroll, and Hsiung (1998) randomly assigned students to either placement into Freshman Composition or into a developmental writing course followed by Freshman Composition. However, their analyses were conditional on either passing the first assigned course (either developmental writing or Freshman Composition, depending on assignment) or on passing Freshman Composition, depending on the specific analysis[1] (see also Moss, Yeaton, & Lloyd, 2014). Much more

[1] Of the approximately 100 students who passed their assigned course in the fall semester, students in the college class scored statistically significantly higher on a writing sample than did students placed into developmental education. There was not a statistically significant difference between groups on the Test of Standard Written English. Among the approximately 90 students who passed Freshman Composition, this pattern was reversed: the remediation group scored statistically significantly higher on the Test of Standard Written English, while there was no statistically significant difference between the two groups on the writing sample.

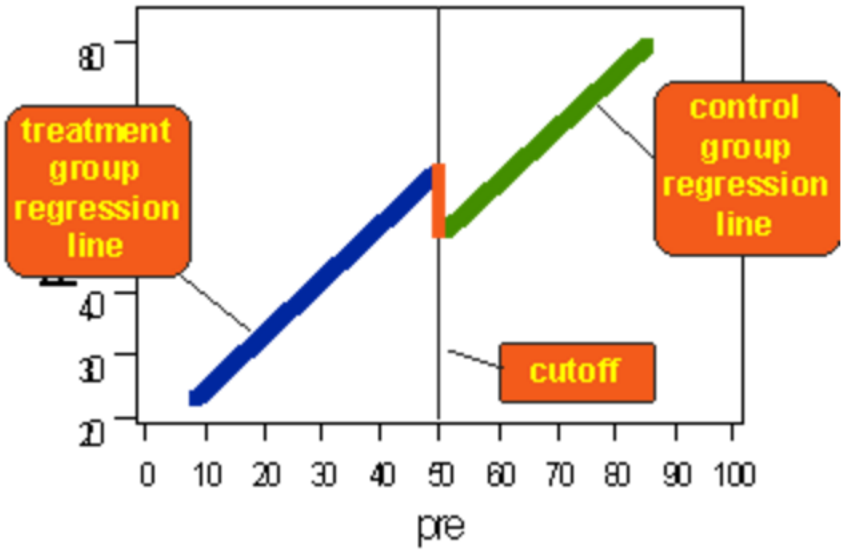
common are non-randomized experiments that adopt a similar approach of conditioning, in one way or another, on success in the developmental course. A randomized experiment that followed students regardless of whether they actually enrolled in the developmental course (or even enrolled in college), and assessed an outcome that is not dependent on course participation (e.g., whether or not students ultimately passed the college level course in which remediation was needed) would provide a fair test of whether placement into developmental education is helpful to students. Given the scarcity of randomized trials in this area, it seems likely that institutions are reluctant to randomly assign students to developmental education or not. But because students are typically assigned to developmental education on the basis of a test score, regression discontinuity is a viable option for studying the effects of assignment to developmental education.

The Regression Discontinuity Design

The basic requirement of a regression discontinuity (RD) design study is that assignment to conditions is done using a score on a continuous variable. An example is the Tennessee process described by Boatman and Long (2010) above. Students are assigned to college algebra if they score 50 and above on the COMPASS Algebra placement test, and to developmental algebra if they score below 50. Thus in a RD designed study (a) groups are formed by design, (b) the assignment mechanism is completely known if (c) the cut score is adhered to (all of these features are shared with randomized experiments). The fact that the assignment mechanism is known allows for unbiased inferences if the assumptions of RD a met and if the data are analyzed properly.

Relative to randomized experiments, RD studies have lower statistical power (Schochet, 2009) and are dependent on more assumptions, some of which are untestable (Valentine & Thompson, 2013). Despite these drawbacks, RD is growing in popularity as researchers become more familiar with its strengths and the conditions under which it is particularly useful. Google Scholar (as of May 16, 2016) lists about 1,500 hits for "regression discontinuity" in 1995, 4,080 in 2005, 7,040 in 2010, and 11,400 in 2015. Good primers on RD are available (e.g., Jacob, Zhu, Somers, & Bloom, 2012; Shadish, Cook, & Campbell, 2002; What Works Clearinghouse, 2015), but the basic logic underlying RD is easy to visualize. Figure 1(from the Social Science Research Methods website; Trochim, 2006) shows the scores of students on a pretest (the x-axis) plotted against their scores on a post-test (the y-axis). The solid line down the middle of the figure represents the assignment cut score. The Figure shows what happens when there is a positive treatment effect, which can be seen visually as a "discontinuity" – a break – *at the cut point* between the scores of students who do and do not receive the intervention. Therefore RD is similar to an interrupted time series approach, except that assignment is based on a score instead of time.

Figure 1. Illustration of an Effect Observed in an RD Study



From Trochim, 2006 (http://www.socialresearchmethods.net/kb/quasird.php).

All studies should be evaluated for the rigor with which they were designed and analyzed (Valentine & Cooper, 2008), and this statement is especially true for RD studies. While still a developing field of study, the What Works Clearinghouse's (2015) regression discontinuity standards provide a good example of how a quality assessment of an RD study might be carried out. The WWC articulates five quality markers for RD studies. These are (a) the variable used to create groups cannot be manipulated, (b) data loss due to attrition should be minimal, (c) there must be no evidence of a discontinuity anywhere other than at the cutoff variable, (d) the functional form of the relationship between the variable used to create groups and the outcome is properly specified (i.e. if the relationship is quadratic it should be modeled as such), and (e) the analyses are constrained to a proper "bandwidth" around the cutoff variable.

Systematic Review and Meta-Analysis

Systematic reviewing and meta-analysis are now the standard set of tools that researchers use to investigate the effectiveness of policies, procedures, and practices when multiple studies pertaining to the specific research question exist. As described below, we found 11 studies using RD that examine the effects of placement into developmental education. Though meta-analysis of RD studies is rare, it is not unknown (Deke, Dragoset, Bogan, & Gill, 2012; Quinn, Lynch, & Kim, 2014), and we anticipate that it will become more common in the future. We begin with a description of how we located, assessed for inclusion, and coded the studies in our analyses. We present our findings on the effects of placement into developmental education on four outcomes: (a) college level credits earned, (b) whether or not students eventually passed the college level course in which remediation was needed, (c) student grades in the course in which remediation was needed, and (d) whether or not students earned a degree or certification. As will be seen, the data mostly suggest statistically significant and potentially important negative impacts on these outcomes. We conclude with suggestions about how placement into developmental education might be improved, and a discussion of the cautions and limitations that go along with our work.

Methods

Literature Search

This review is part of a larger project examining interventions for developmental education students. The electronic literature search was initially conducted in ERIC and PsycInfo, from 1993 – March 2013. Search terms were divided into three groups: (a) terms that identified as study as involving developmental education (e.g., developmental or non-credit or remedia*); (b) terms that identified the context as postsecondary education (e.g., universit* or community college); and (c) a term that identified the document a study that used regression discontinuity (discontin*). Documents with at least one search term from each these categories were screened for relevance by at least two trained individuals who worked independently. Disagreements were resolved by a third screener. We included only studies that examined the effects of placement into developmental education relative to placement directly into the college level course (and not, for example, studies that examined the effects of placement into different levels of developmental education; see Melguizo, Bos, Ngo, Mills, & Prather, 2016).

We also engaged in ancillary searches to find studies of the effects of placement in developmental education. First, because the *Journal of Higher Education* does not publish abstracts, we hand searched that journal from 1993 forward. In addition, we conducted Google Scholar searches for relevant studies, and forward citation searches on the researchers who authored relevant papers. The last literature searches were run in November 2015.

Coding

Once potentially relevant studies were identified, studies were coded by two reviewers who worked independently. We coded characteristics related to study context, the developmental education placement process, the sample, and the study's outcomes. These characteristics included institution type (community college, university), the number of institutions in the study, whether or not the study was published, the process used to place students into developmental education (e.g., the specific placement test used and the cutoff for placement), and information about the students in the sample (e.g., whether the study included only first time, full time students).

Analytic Model

To synthesize the results of the eligible studies, we used standard meta-analytic techniques including inverse variance weighting, which gives proportionally more weight to larger studies. Many studies presented the results of multiple models (e.g., models with more or fewer covariates). Rather than adopting a robust variance estimation approach, when we chose models in a deliberate attempt to maximize the conceptual similarity of the studies in the analysis. Therefore, when we had a choice we always selected the model with (a) the largest number of control variables in it, (b) the narrowest bandwidth, and (c) results that were as close to three years from the time of assignment as possible (except for attainment, for which we selected the longest follow up point).

Researchers undertaking a meta-analysis need to consider whether to employ a fixed effect or a random effects analytic model. Using the fixed effect model, study effect sizes can be thought of as estimating a single population value, and therefore any differences in effect sizes across studies are treated as solely due to random sampling and identifiable covariates. Using the random effects model, reviewers assume that studies do not in fact share a single population value but instead come from a distribution of effect sizes. Therefore any differences in effect sizes across studies are due to random sampling error, any identifiable covariates, and other random factors that cannot be identified.

The choice between fixed effect and random effects models can be an important one, because the confidence intervals arising from a random effects analysis will never be smaller and are often larger than their fixed effect counterparts; this has implications for both the statistical significance tests and interpreting the likely range of population effects. Often, the random effects model is thought to be the most defensible choice, in part due to its somewhat better generalization properties (Hedges & Vevea, 1998). However, one issue with the random effects model is that if the number of studies is small, the estimate of the between-studies variance component (i.e. the extent to which population effect sizes differ from one another) is both highly uncertain and highly unstable. That is, the between-studies variance component is estimated with a great deal of error, and it can be very sensitive to the inclusion of new information (e.g., a new study in an updated review). Due to these considerations we report both the fixed effect and the random effects models in this review. In addition, we report several sensitivity analyses as robustness checks.

Finally, we should note that three studies examined the effects of placement into developmental education in multiple subjects (Boatman & Long, 2010; Calcagno & Long, 2008; Scott-Clayton & Rodriguez, 2012). Within each study, we treat these effects as independent. However, it is possible that some students could have been placed into developmental education in multiple subjects, and therefore be in our analysis more than once. For example, a student in Calcagno and Long's study could have been placed into developmental math and developmental reading, and might have appeared in both of their bandwidth-constrained analyses; this would violate the statistical assumption of independence. We do not know the extent to which this combination of events happened. However, in Boatman and Long (2010), 17% of students were recommended for placement into developmental education in two subjects, and 5% were recommended for placement into three subjects. Therefore in that study the maximum overlap is 22%, but the overlap within the optimal bandwidth (i.e. students who scored between 47 and 52 on the math placement test and between 65 and 70 on the reading placement test) is likely much smaller (though probably not zero).

Results

The literature search uncovered 11 reports, with a total of 21 independent samples, that use RD to investigate the effects of placement into developmental education (henceforth, we refer to independent samples as “studies”). However, Harmon (2011) does not appear in our analyses, as that study did not examine one of our four primary outcomes[2]. The studies varied widely in size. The within study sample sizes (based on the analyses we used) ranged from 185 to 59,334, with a median sample size of about 1,000 students (in all, well over 100,000 students are represented in the meta-analytic database).

Credits Earned

Sixteen analyses examined the effect of placement into developmental education on college credits earned. As can be seen in Table 1, credits earned were typically examined about three years after assignment. The mean effect size under fixed effect assumptions was -1.86 credits, $p < .001$. The homogeneity test was statistically

[2] This study focused on the overall grade point average of students who successfully exited developmental education relative to students not assigned to developmental education.

significant, Q (15) = 43.18, p < .001 (I^2 = 68%). The mean effect size under random effects assumptions was -3.00 credits, p = .002. Below we report two sensitivity analyses and two exploratory moderator analyses on this dataset.

Table 1. Study Characteristics and Outcomes: Credits Earned

Author(s) (Year)	Institution Type	Placement Test	Levels of Dev. Education	Retake Policy	Subject	Analytic Sample Size	Timing (from Placement Semester)	Covariates Included in the Model	Effect Size (SE)
Boatman (2012) (APSU)	4	ACT	1	Allowed	Math	928	2 years	Gender Race/Ethnicity HS Achievement	0.662 (7.098)
Boatman (2012) (CSCC)	2	ACT	1	Allowed	Math	489	2 years	Gender Race/Ethnicity HS Achievement	1.724 (2.305)
Boatman (2012) (JSCC)	2	ACT	1	Allowed	Math	624	2 years	Gender Race/Ethnicity HS Achievement	1.969 (0.242)
Boatman & Long (2010)	4	COMPASS	1	Rare	Math	263	3 years	Gender Race/Ethnicity SES	-4.8965 (3.4042)
Boatman & Long (2010)	2	COMPASS	1	Rare	Math	227	3 years	Gender Race/Ethnicity SES	-8.3323 (4.178)
Boatman & Long (2010)	4	COMPASS	1	Rare	Reading	559	3 years	Gender Race/Ethnicity SES	-4.1206 (3.0274)
Boatman & Long (2010)	2	COMPASS	1	Rare	Reading	938	3 years	Gender Race/Ethnicity SES	-9.5066 (1.8544)
Boatman & Long (2010)	4	COMPASS	1	Rare	Writing	336	3 years	Gender Race/Ethnicity SES	-2.0349 (3.1653)
Boatman & Long (2010)	2	COMPASS	1	Rare	Writing	622	3 years	Gender Race/Ethnicity SES	-7.3279 (2.4366)

Calcagno & Long (2008)	2	CPT	Unknown	Used “no retake” sample	Math	9,593	6 years	Gender Race/Ethnicity	-0.244 (3.641)
Calcagno & Long (2008)	2	CPT	Unknown	Used “no retake” sample	Reading	8,755	6 years	Gender Race/Ethnicity	-1.59 (2.124)
Hodara (2012)	2	CUNY	Unknown	Unclear	Writing/ESL	12,773	3 years	Gender HS Achievement Race/Ethnicity SES	-1.146 (0.925)
Martorell & McFarlin (2011)	2	TASP	Unknown	Allowed (but used first attempt score)	Reading or math	59,344	6 years	Race/Ethnicity SES	-3.96 (2.15)
Martorell & McFarlin (2011)	4	TASP	Unknown	Allowed (but used first attempt score)	Reading or math	33,910	6 years	Race/Ethnicity SES	-7.59 (2.71)
Scott-Clayton & Rodriguez (2015)	2	COMPASS	Unknown	Allowedbut strict	Math	17,641	3 years	Gender Race/Ethnicity	0.007 (0.796)
Scott-Clayton & Rodriguez (2015)	2	COMPASS	Unknown	Allowed but strict	Reading	1,374	3 years	Gender Race/Ethnicity	-3.183 (3.023)

Notes: Effect sizes are unstandardized regression coefficients, so represent the observed effect in terms of credit hours earned. Because all sample sizes are not small, a *z* test for each effect size can be given by the effect size ÷ standard error. For Martorell and McFarlin (2011), the regression coefficient represents the total number of college level credits attempted over a six-year period.

For institution type, 2 = community college, 4 = university.

For placement test, CPT = Florida College Entry Level Placement Test, CUNY is the City University of New York’s placement test, and TASP is the Texas Academic Skills Program.

Ever Pass College Level Course for Subject?

Table 2 houses the effect size estimates for the six analyses involving whether students ever passed the college level course in which remediation was needed. For both fixed and random effects models, the mean effect size was a 7.9 percentage point reduction in the proportion of students eventually passing the college level course in which remediation was needed (e.g., from 75% to 68%), $p < .001$ for the fixed effect model and $p = .004$ for the random effects model. The homogeneity test was statistically significant, $Q(5) = 27.31$, $p < .001$ ($I^2 = 81\%$).

Table 2. Study Characteristics and Outcomes: Ever Pass College Level Course

Author(s) (Year)	Institution Type	Placement Test	Levels of Dev. Education	Retake Policy	Subject	Analytic Sample Size	Timing (from Placement Semester)	Covariates Included in the Model?	Effect Size (SE)
Calcagno & Long (2008)	2	CPT	Unknown	Used "no retake" sample	Reading	8,755	6 years	Gender Race/Ethnicity	-0.036 (0.017)
Calcagno & Long (2008)	2	CPT	Unknown	Used "no retake" sample	Math	9,593	6 years	Gender Race/Ethnicity	-0.002 (.064)
Hodara (2012)	2	CUNY	1	Unknown	Writing/ESL	14,733	3 years	Gender HS Achievement Race/Ethnicity SES	-0.147 (0.017)
Lesik (2006)	4	Unknown	1	Allowed but none in sample did	Math	212	4 years	Placement score only	+0.307 (0.668)
Scott-Clayton & Rodriguez (2015)	2	COMPASS	Unknown	Allowed but strict	Math	17,641	3 years	Gender Race/Ethnicity	-0.059 (0.015)
Scott-Clayton & Rodriguez (2015)	2	COMPASS	Unknown	Allowed but strict	Reading	1,374	3 years	Gender Race/Ethnicity	-0.143 (0.055)

Notes: Effect sizes are ordinary least squares regression estimates with a binary dependent variable, so represent the observed effect in terms of percentages passing the college level course (e.g., -0.147 means that students assigned to developmental education passed the first college level course in which remediation was needed at a rate that was 14.7 percentage points less than the rate at which students assigned directly to the college level course passed it). Because all sample sizes are not small, a *z* test for each effect size can be given by the effect size ÷ standard error.

For Lesik (2006), we assumed that the base rate of passing was 50%, which resulted in the most optimistic effect size possible. The .307 effect size represents a translation of the logged odds ratio reported in Table 4 of 1.43.

For institution type, 2 = community college, 4 = university.

For placement test, CPT = Florida College Entry Level Placement Test, and CUNY is the City University of New York's placement test.

Achievement in College Level Course if Taken and Completed

Table 3 contains effect size estimates for nine analyses addressing the academic performance in the college level course in which remediation was needed, conditional on students taking and completing that course. Of our four main outcomes, this is the one that is most likely to be biased by treatment-induced attrition, though the direction of this bias is difficult to predict. For the fixed effect analysis, the estimated effect size is 0.00 (p = .98). For the random effects model, the estimated effect size is +.01 grade points (p = 94). The homogeneity test was not statistically significant, Q (8) = 15.16, p = .06 (I^2 = 46%).

Table 3. Study Characteristics and Outcomes: Achievement in College Level Course

Author(s) (Year)	Institution Type	Placement Test	Levels of Dev. Education	Retake Policy	Subject	Analytic Sample Size	Timing (from Placement semester)	Covariates Included in the Model?	Effect Size (SE)
Boatman & Long (2010)	4	COMPASS	1	Rare	Math	185	3 years	Gender Race/Ethnicity SES	0.2169 (0.2025)
Boatman & Long (2010)	2	COMPASS	1	Rare	Math	227	3 years	Gender Race/Ethnicity SES	0.1451 (0.2422)
Boatman & Long (2010)	4	COMPASS	1	Rare	Reading	460	3 years	Gender Race/Ethnicity SES	-0.1058 (0.1382)
Boatman & Long (2010)	2	COMPASS	1	Rare	Reading	596	3 years	Gender Race/Ethnicity SES	0.0473 (0.1171)
Boatman & Long (2010)	4	COMPASS	1	Rare	Writing	315	3 years	Gender Race/Ethnicity SES	0.2971 (0.1464)
Boatman & Long (2010)	2	COMPASS	1	Rare	Writing	467	3 years	Gender Race/Ethnicity SES	-0.1438 (0.1356)
Horn et al. (2009)	2	COMPASS	1	Unknown	Reading	328	Unknown	Gender Race/Ethnicity	-0.552 (0.216)
Moss & Yeaton (2006)	2	ASSET	Unknown	Unknown	Reading	1,473	6 years	Placement score only	-0.02 (0.09)
Moss et al. (2014)	2	COMPASS	1		Math		1 semester	Placement score only	+0.34 (.33)

Notes: Effect sizes are unstandardized regression coefficients, so represent the observed effect in terms of student grades (eg., +.22 means that students assigned to developmental education scored .22 grade points higher on average than students assigned directly to the college level course). Because all sample sizes are not small, a *z* test for each effect size can be given by the effect size ÷ standard error.

Horn et al. (2009) did not report the standard error for the regression coefficient, but did report that the coefficient's *p*-value was less than .05 (but presumably larger than .01). 0.216 is the standard error that yields p = .011. The model without covariates has a standard error of 0.245 so 0.216 does seem reasonable.

For institution type, 2 = community college, 4 = university.

We should note here that while Scott-Clayton and Rodriguez (2015) also measured academic achievement in the college level course in which remediation was needed, they did so by dummy coding achievement as whether students earned a B in the college level course. In and of itself this does not create a problem for our analysis, but Scott-Clayton and Rodriguez coded as “0” any student who either (a) earned less than a B or (b) never took the college level course. Because this analysis conflates two aspects of the educational experience that we think should be kept separate, we did not use the two effect sizes from this study in our meta-analysis. Both were negative and statistically significant.

Degree Attainment

Thirteen studies examined the effect of placement into developmental education on degree or certificate attainment (see Table 4). For both fixed and random effects models, the mean effect size was a 1.5 percentage point reduction in the proportion of students eventually earning a degree (e.g., from 30% to 28.5%; $p = .03$ for both models). The homogeneity test was not statistically significant, Q (12) = 13.39, $p = .34$ ($I^2 = 7\%$).

Table 4. Study Characteristics and Outcomes: Attainment

Author(s) (Year)	Institution Type	Placement Test	Levels of Dev. Education	Retake Policy	Subject	Analytic Sample Size	Timing (from Placement Semester)	Covariates Included in the Model	Effect Size (SE)
Boatman & Long (2010)	4	COMPASS	1	Rare	Math	263	6 years	Gender Race/Ethnicity SES	-.1176 (0.1687)
Boatman & Long (2010)	2	COMPASS	1	Rare	Math	227	6 years	Gender Race/Ethnicity SES	-.4397 (0.2080)
Boatman & Long (2010)	4	COMPASS	1	Rare	Reading	559	6 years	Gender Race/Ethnicity SES	-.1424 (0.1506)
Boatman & Long (2010)	2	COMPASS	1	Rare	Reading	938	6 years	Gender Race/Ethnicity SES	-.1526 (0.1157)
Boatman & Long (2010)	4	COMPASS	1	Rare	Writing	366	6 years	Gender Race/Ethnicity SES	.1982 (0.1626)
Boatman & Long (2010)	2	COMPASS	1	Rare	Writing	652	6 years	Gender Race/Ethnicity SES	.0035 (0.1488)
Calcagno & Long (2008)	2	CPT	Unknown	Used "no retake" sample	Math	9,593	6 years	Gender Race/Ethnicity	-.027 (.015)
Calcagno & Long (2008)	2	CPT	Unknown	Used "no retake" sample	Reading	8,755	6 years	Gender Race/Ethnicity	-.031 (.026)
Hodara (2012)	2	CUNY	Unknown	Unclear	Reading and writing	12,773	3 years	Gender HS Achievement Race/Ethnicity SES	-0.001 (0.016)

Martorell & McFarlin (2011)	2	TASP	Unknown	Allowed (but used first attempt score)	Reading or math	59,344	6 years	Race/Ethnicity SES	-.023 (0.016)
Martorell & McFarlin (2011)	4	TASP	Unknown	Allowed (but used first attempt score)	Reading or math	33,910	6 years	Race/Ethnicity SES	-.040 (.028)
Scott-Clayton & Rodriguez (2015)	2	COMPASS	Unknown	Allowed but strict	Math	17,641	3 years	Gender Race/Ethnicity	-.001 (0.010)
Scott-Clayton & Rodriguez (2015)	2	COMPASS	Unknown	Allowed but strict	Reading	1,374	3 years	Gender Race/Ethnicity	-.029 (.039)

Notes: Effect sizes are unstandardized regression coefficients, so represent the observed effect in terms of percentage of the sample earning a certificate or degree. Because all sample sizes are not small, a *z* test for each effect size can be given by the effect size ÷ standard error.

For institution type, 2 = community college, 4 = university.

For placement test, CPT = Florida College Entry Level Placement Test, CUNY is the City University of New York's placement test, and TASP is the Texas Academic Skills Program.

The raw magnitude of this effect depends on (a) the size of the incoming class and (b) the proportion of these students assigned to developmental education. At the institution level, in small institutions and in institutions with low developmental education placement rates, the negative effect of placement into developmental education will not matter that much. But in larger institutions, and in institutions with higher placement rates, this effect might be large enough to matter. For example, imagine a typical mid-sized university with 6,000 incoming students, 25% of whom are assigned to developmental education. This institution could be expected to award 22 or 23 fewer degrees in that class than it would have if placement into developmental education had no effect on attainment (i.e. if the graduation rate among non-developmental students is 60%, then 58.5% of the 1,500 developmental students are expected to earn a degree, and the difference between the two attainment rates is 22.5 degrees).

Of course, at the policy level the consequences are staggering. Assume that in a given year 2.5 million students start their college careers in either a university or a community college setting, that one third of these students are placed into developmental education, and that the overall six year graduation rate is 34%. The 1.5 percentage point reduction can be thought of as suggesting that 35% of students not placed into developmental education and 33.5% of students placed into developmental education will graduate in 6 years. This works out to a loss of about 12,500 certificates or degrees for that year's cohort of students.

Exploratory Moderator Analyses

Our dataset of studies examining the effects of placement into developmental education on credits earned is the only one large enough to support even tentative moderator analyses; we report two of these analyses below. The first examines the effects observed in community colleges relative to universities, and the second examines effects observed separately for reading, writing, and math. Even though we approached these hypothesis tests with specific predictions in mind, we believe that they are best conceptualized as exploratory analyses because, as Lipsey and Wilson (2001) observed, studies have personalities in the sense that their traits tend to cluster together. For a meta-analysis, this means that study characteristics tend to correlate with one another, confounding univariate analyses of the relationship between study characteristics and outcomes. As a result moderator analyses in meta-analysis should generally be multivariate so that study characteristics can be examined net of other characteristics in the model. However, meta-regression (the meta-analytic analog to multiple regression) generally requires a large number of studies for both reasonable statistical power and stable estimates. The analyses below are univariate, and as such warrant an extra level of caution when interpreting them.

Effects for Two- Versus Four-Year Institutions (Credit Accumulation Only)

In our meta-analytic dataset we have five estimates of the effects of placement into developmental education on college credit accumulation that are based on four-year institutions and 11 estimates that are based on two-year institutions. For universities, the fixed effect and random effects mean effect size is -4.64 credits, p = .002. The homogeneity test within these five estimates was not statistically significant, Q (4) = 2.46, p = .65. For community colleges, a somewhat different picture emerges. The mean effect size under fixed effect assumptions was -1.56 credits, p = .001. The homogeneity test was statistically significant, Q (10) = 36.84, p < .001. The mean effect size under random effects assumptions was -2.62 credits, $p = .03$.

Effects for Different Subjects

Our meta-analytic dataset includes four analyses of developmental education for reading, three analyses of developmental education for writing, and seven analyses of developmental education for math. For math, the fixed effect and random effects mean effect size is -0.08 credits, $p = .90$. The homogeneity test within these seven studies was not statistically significant, $Q(6) = 7.38$, $p = .29$. For reading, the mean effect size under fixed effect assumptions was -5.45 credits, $p < .001$. The homogeneity test was statistically significant, $Q(3) = 8.58$, $p = .04$. The mean effect size under random effects assumptions was -4.87 credits, $p = .01$. For writing, the mean effect size under fixed effect assumptions was -1.93 credits, $p = .02$. The homogeneity test was not statistically significant, $Q(2) = 5.63$, $p = .06$. The mean effect size under random effects assumptions was -3.18 credits, $p = .11$.

Sensitivity Analyses

Because we have the most information on credits earned, we used this dataset to conduct several sensitivity analyses. First, we Winsorize the meta-analytic weights and next, we drop studies one at a time from the analysis. Both of these strategies are intended to ensure that our results are not being driven by a single study. Finally, five studies allow us to tentatively test the extent to which study results are sensitive to the bandwidth that was used.

Influence Analyses

Under fixed effect assumptions two studies (Hodara, 2012 and Scott-Clayton and

Rodriguez's, 2012 math analysis) have relative weights of 25% and 33%, suggesting that these studies are large relative to the other studies in the dataset. Perhaps more important, Boatman's (2012) community college reading analysis is very influential. By this we mean that the analysis' weight (which is above the mean) and effect size (the absolute value of which is the largest in the database) combine to exert a large influence on the fixed effect analysis of college credits earned. We addressed these potentially influential studies in two ways. First, we Winsorized the outlier weights. We defined an outlier using Tukey's (1977) rule (i.e. an outlier is an observation that is more than two standard deviations beyond the 75th percentile). As we suspected, Hodara (2012) and Scott-Clayton and Rodriguez's (2015) math analysis were identified as outliers. We then trimmed the weights iteratively (recoding the weights so that they were no longer outliers, then rechecking for outliers) until no outliers were identified. This process had the effect of inflating the standard errors for these two studies (from .925 to 1.462 for Hodara, and from .796 to 1.462 for Scott-Clayton and Rodriguez). The patterns of statistical significance were unchanged across the mean effect size under fixed effect assumptions, the mean effect size under random effects assumptions, and the homogeneity analysis. Winsorizing resulted in a much larger point estimate for the fixed effect analysis, and had virtually no effect on the random effects analysis. The estimate of between study heterogeneity dropped somewhat with Winsorized weights (I^2 values were 68% for the main specification versus 59% for the Winsorized analysis).

Next, we addressed potentially influential studies by dropping one study at a time from the main analysis of the effects of placement into developmental education on college credits earned. Again, most of the changes are minor, but dropping Boatman's

(2010) community college reading effect results in a large change to the fixed estimate (from -1.86 to -1.44 credits) and to the random effects estimate (-3.04 to -2.30 credits). Dropping both Hodara (2012) and Scott-Clayton and Rodriguez's (2015) math analysis resulted in less dramatic increases to both fixed effect and random effects estimates. Across these "drop one study" analyses the statistical conclusions did not change (i.e. the mean effect was negative and statistically significant under fixed and random effects assumptions, and the homogeneity test was statistically significant), and the substantive interpretation of the effects was highly similar.

RD Assumptions

Two studies (Calcagno & Long, 2008; Martorell & McFarlin, 2011) provide effect sizes both for all students in the analysis and for a specific bandwidth. Similarly, three studies in our meta-analytic database (Hodara, 2012; Moss et al., 2014; Scott-Clayton & Rodriguez, 2015) used at least two bandwidths as a sensitivity check. We tested whether results were sensitive to bandwidth choice using *z* tests (procedures described in Borenstein, Hedges, Higgins, & Rothstein, 2009, pp. 229-230) for computing the variance of two correlated variables. This procedure requires that researchers know or estimate the extent to which the standard errors are based on independent information. Though not realistic, we chose zero for this value because doing so yields the smallest possible standard error. This means that the statistical tests in are more likely to result in a rejection of the null hypothesis of no difference, even when there is no actual difference between the estimates and as such represent a "worst case" scenario.

Only one of the 13 tests resulted in a rejection of the null hypothesis (for that study, p = .048). Correcting for multiple comparisons using any common family wise or false discovery rate procedure (e.g., a Bonferroni correction or the Benjamini-Hochberg correction) yields nonsignificant results for all tests. Further, there was no consistency in the direction of the differences, and the median *p*-value across these 13 analysis is .41. As such we cannot find evidence in these studies that the observed effect sizes were unduly influenced by our decision to use the most narrow bandwidth given in the studies.

Discussion

This paper reviewed evidence on the effects of placement into developmental education as evaluated with regression discontinuity designs. If the causal inferences are correct and our effect sizes are reasonably accurately estimated, the meta-analyses of studies using regression discontinuity to investigate the effects of placement into developmental education suggest that placement into developmental education results in statistically significant and substantively sizable negative impacts. Relative to their peers who are also on the margins of college readiness but who were placed into college level courses, students placed into developmental education earned fewer college credits after about 3 years (our estimates ranged from about 2 to 3 credit hours, depending on model specification), were about 8 percentage points less likely to eventually pass the college level course in which remediation was needed, and were about 1.5 percentage points less likely to earn a certificate or degree. We cannot reject the null hypothesis that marginal students placed into developmental education perform similarly (i.e. earn similar grades) in the college level course in which remediation was needed relative to marginal students placed into the college level course. The results for college credits earned were not sensitive to either outlier effect sizes (there were none) or outlier weights. Influential studies similarly did not affect the statistical significance of the results, though in the fixed effect model there was some variation in the effect sizes observed depending on which studies were in the analysis (effect sizes in the random effects model were very similar regardless of which studies were in the analysis). There is no evidence that the observed effect sizes were influenced by the decision to focus on the narrowest bandwidth presented in the studies in the review.

The exploratory moderator analyses using the studies that assessed college credits earned suggest that the negative effects of placement into developmental education are stronger for university students (but still statistically significant and negative for community college students), and for students placed into developmental education in reading and writing (recall that for writing, the fixed effect estimate was statistically significant but the random effects estimate was not, $p = .11$), but not math (the fixed and random effects estimates were close to zero and were not statistically significant). This latter point merits additional research attention, as Roksa, Jenkins, Jaggars, Zeidenberg, and Cho (2009) found that among students attending community college in Virginia, among those who enrolled in college level English the probability of passing that course was unrelated to placement test scores. Though just one study, this finding raises questions about the adequacy of placement test scores as a basis for assigning students to developmental education.

How Can Educational Systems and Institutions Improve the Situation?

This study was designed to assess, across multiple studies in many contexts, if placement into developmental education helps students be successful in college. It was not designed to address how or why any positive or negative effects might have occurred. That said, much of the national conversation on developmental education has focused on misplacement rates. As mentioned earlier, placement is generally based on a single test. No one believes – or at least, no one should believe – that these tests are perfect indicators of college readiness (see Armstrong, 2000). A general principle of psychological measurement is when a construct (like college readiness) is measured imperfectly, one way to improve measurement is to measure the construct in multiple ways. Incorporating information that many institutions already have - like high school grade point average and scores on standardized entrance tests - into placement decisions is only relatively easy way to modify the placement rubric that has the potential to reduce misplacement rates. Title 5 §55502 of the California Code of Regulations explicitly recognizes this by requiring institutions to use multiple measures for placement into developmental education, and even placement test developers recommend that institutions use multiple measures for placement (Westrick & Allen, 2014). If we were responsible for running an institution, attempting to reduce misplacement rates by using multiple measures would be where we would start reform efforts (see Bahr, Hayward, Hetts, Lamoree, Newell, Pellegrin, Sorey, & Willett, 2014).

Furthermore, it is not clear that all students need a semester long course to achieve college readiness, and researchers have been experimenting with other ways to accomplish this goal. For example, Logue, Watanabe-Rose, and Douglas (2016) conducted an RCT in which algebra instruction was embedded into a college level statistics course supplemented with weekly workshops that focused on algebra. Compared to students who took the usual developmental algebra course, or that course supplemented with weekly workshops, students taking the college statistics course earned more college level credits over three semesters (21 to about 15 in the other two groups), and were more likely to pass the course to which they were assigned. Other possible ways of remediating deficits include summer bridge programs, targeted one credit pre-semester tune up courses, and by providing additional supports (e.g., mandatory tutoring sessions during the semester). The important point is that educational leaders should think carefully about who gets placed into developmental education and develop flexible systems to help students develop the skills that they need to be successful in college (see Bailey, 2009).

Limitations and Conclusion

An important conceptual limitation is that this study did not address the effect of placement at different levels of developmental education (e.g., elementary versus intermediate algebra). Due to the relatively small numbers of students placed at the lowest levels of developmental education, and the fact that all else being equal statistical power in RD is much lower than in an RCT, it is likely that a series of randomized experiments will be needed to address this question.

With respect to the questions that we were able to address, perhaps the greatest threat to the conclusions we draw in this paper is our analyses are based on studies with characteristics that differ in fundamental and probably important ways. Most obviously, we included studies that examined the effects of placement into developmental reading, writing, and math, and studies that occurred in both community colleges and universities. We were only able to test these two potential modifiers of the effects of placement in developmental education for one outcome (credits earned) because we had too few studies to support parallel analyses for the other outcomes. Those analyses did suggest that there is reason to suspect heterogeneous effects (e.g., placement into developmental education appears to have more negative effects on university students than on community college students). However, these analyses were not multivariate, and therefore could confound the effects of other study characteristics with the ones we were examining. Readers can draw some reassurance from our extensive sensitivity analyses, which suggest that our results are not unduly influenced by exceptional studies or by some of the important decisions we made when assembling our meta-analytic dataset.

Even exercising appropriate caution in drawing causal conclusions from our research, based on the studies we review it is very difficult to walk away with the conclusion that placement into developmental education *helps* students. More than 75% of the estimates in our meta-analytic database are negative, and the meta-analytic estimates for the probability of passing the college level course in which remediation was needed, college credits earned, and attainment are all negative, statistically significant, and large enough to be meaningful. Our hope is that this work spurs thoughtful debate and research on placement policies and on alternative mechanisms for ensuring that college students have the skills needed to meet their goals.

References

Aiken, L. S., West, S. G., Schwalm, D. E., Carroll, J. L., & Hsiung, S. (1998). Comparison of a randomized and two quasi-experimental designs in a single outcome evaluation efficacy of a university-level remedial writing program. *Evaluation Review, 22*, 207-244.

Armstrong, W. B. (2000). The association among student success in courses, placement test scores, student background data, and instructor grading practices. *Community College Journal of Research and Practice*, *28*, 681-695. DOI:10.1080/10668920050140837

Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *Journal of Higher Education, 77,* 886-924.

Bahr, P., Hayward, P., Hetts, J., Lamoree, D., Newell, M., Pellegrin, N., Sorey, K., & Willett, T. (2014). *Multiple measures for placement and assessment: White paper*. Educational Results Partnership. Available online at http://rpgroup.org/system/files/MMAP_WhitePaper_Final_September2014.pdf

Bailey, T. (2009). Challenge and opportunity: Rethinking the role and function of developmental education in community college. *New Directions for Community Colleges, 145*, 11-30. DOI: 10.1002/cc.352

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.

Barry, M. N. and Dannenberg, M. (2016). *Out of pocket: The high costs of inadequate high schools and high school student achievement on college affordability*. Washington, DC: Education Reform Now. Available online at https://edreformnow.org/app/uploads/2016/04/EdReformNow-O-O-P-Embargoed-Final.pdf

* Boatman, A. (2012). *Evaluating institutional efforts to streamline postsecondary remediation: The causal effects of the Tennessee developmental course redesign initiative on early student academic success*. New York, NY: National Center for Postsecondary Research.

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

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley and Sons.

Breneman, D. W., & Haarlow, W. N. (1998). *Remediation in higher education: A symposium featuring developmental education: Costs and consequences. Fordham Report, 2*. Washington, DC: Thomas B. Fordham Foundation

* Calcagno, J. C. & Long, B. T. (2008). *The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance*. New York, NY: National Center for Postsecondary Research.

Complete College America (2012). *Remediation: Higher education's bridge to nowhere*. Available online at http://www.completecollege.org/docs/CCA-Remediation-final.pdf

Deke, J., Dragoset, L., Bogan, K., & Gill, B. (2012). *Impacts of Title I supplemental educational services on student achievement* (NCEE 2012-4053). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Harmon, T. B. (2011). *Remedial policy in the California State University system: An analysis*. Retrieved from Proquest Digital Dissertations (UMI Number 3468568)

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods, 3*, 486-504.

* Hodara, M. (2012). *Language minority students at community college: How do developmental education and English as a second language affect their educational outcomes*? Retrieved from Proquest Digital Dissertations (UMI Number 3505981)

* Horn, C., McCoy, Z., Campbell, L., & Brock, C. (2009). Remedial testing and placement in community colleges. *Community College Journal of Research and Practice, 33*, 510-526.

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

Jacob, R. T., Zhu, P., Somers, M. A., and Bloom, H. S. (2012). *A practical guide to regression discontinuity*. New York, NY: MDRC. Available online at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.671.4723&rep=rep1&type=pdf

* Lesik, S. A. (2006). Applying the regression-discontinuity design to infer causality with non-random assignment. *The Review of Higher Education, 30*, 1-19.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage publications.

Logue, A. W., Watanabe-Rose, M., & Douglas, D. (2016). Should students assessed as needing remedial mathematics take college-level quantitative courses instead? A randomized controlled trial. *Educational Evaluation and Policy Analysis* DOI:10.3102/0162373716649056

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

Melguizo, T., Bos, J. M., Ngo, F., Mills, N., & Prather, G. (2016). Using a regression discontinuity design to estimate the impact of placement decisions in developmental math. R*esearch in Higher Education, 57*, 123-151.

* Moss, B. G., & Yeaton, W. H. (2006). Shaping policies related to developmental education: An evaluation using the regression-discontinuity design. *Educational Evaluation and Policy Analysis, 28*, 215-229.

* Moss, B. G., Yeaton, W. H., & Lloyd, J. E. (2014). Evaluating the effectiveness of developmental mathematics by embedding a randomized experiment within a regression discontinuity design. *Educational Evaluation and Policy Analysis, 36*, 170-185.

National Center for Educational Statistics (2016).

Pretlow III, J. & Wathington, H. D. (2012). Cost of developmental education: An update of Breneman and Haarlow. *Journal of Developmental Education, 36*, 4-14.

Quinn, D. M., Lynch, K., & Kim, J. S. (2014). *Replicating the moderating role of income status on summer school effects across subject areas: A meta-analysis*. Paper presented at the Spring Conference of the Society for Research on Educational Effectiveness.

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: Columbia University, Teachers College, Community College Research Center. Available online at http://67.205.94.182/publications/gatekeeper-course-success-virginia.html.

Schochet, P. Z. (2009). Statistical power for regression discontinuity designs in education evaluations. *Journal of Educational and Behavioral Statistics, 34*, 238-266.

* 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. doi:10.1162/EDFP_a_00150

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage Learning.

Trochim, W. M. (2006). The research methods knowledge base, 2nd edition. Available online at: http://www.socialresearchmethods.net/kb/index.php

Valentine, J. C., & Cooper, H. (2008). A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: the Study Design and Implementation Assessment Device (Study DIAD). *Psychological Methods, 13*, 130-149.

Valentine, J. C., & Thompson, S. G. (2013). Issues relating to confounding and meta-analysis when including non-randomized studies in systematic reviews on the effects of interventions. *Research Synthesis Methods, 4*, 26-35.

Westrick, P. A., & Allen, J. (2014). *Validity evidence for ACT COMPASS tests*. Iowa City, IA: ACT. Available online at http://files.eric.ed.gov/fulltext/ED546849.pdf

What Works Clearinghouse (2015). *Preview of regression discontinuity design standards*. Available online at http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_rdd_standards_122315.pdf

* Indicates studies that were used in the meta-analyses.