Stanford cepa | Center for Education Policy Analysis

CEPA Working Paper No. 16-01

The Casual Effects of Cultural Relevance: Evidence from an Ethnic Studies Curriculum

AUTHORS

Thomas Dee
Stanford University

Emily Penner
Stanford University

ABSTRACT

An extensive theoretical and qualitative literature stresses the promise of instructional practices and content aligned with the cultural experiences of minority students. Ethnic studies courses provide a growing but controversial example of such "culturally relevant pedagogy." However, the empirical evidence on the effectiveness of these courses is limited. In this study, we estimate the causal effects of an ethnic studies curriculum piloted in several San Francisco high schools. We rely on a "fuzzy" regression discontinuity design based on the fact that several schools assigned students with eighth-grade GPAs below a threshold to take the course in ninth grade. Our results indicate that assignment to this course increased ninth-grade student attendance by 21 percentage points, GPA by 1.4 grade points, and credits earned by 23. These surprisingly large effects are consistent with the hypothesis that the course reduced dropout rates and suggest that culturally relevant teaching, when implemented in a supportive, high-fidelity context, can provide effective support to at-risk students.

Acknowledgements: The authors would like to thank Laura Wentworth, Bill W. Sanderson, and other members of the San Francisco Unified School District High School leadership team for supporting this research as well as Jim Shen and Mari Muraki for their invaluable assistance with data management. Financial support for this research came from the Stanford GSE Incentive Fund for Projects in SFUSD and from the Institute of Education Sciences Postdoctoral Training Fellowship under award number R305B130017. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

VERSION

January 2016

Suggested citation: Dee, T., & Penner, E. (2016). The Casual Effects of Cultural Relevance: Evidence from an Ethnic Studies Curriculum (CEPA Working Paper No.16-01). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp16-01

Introduction

The racial and ethnic gaps that exist across a variety of important student outcomes in the United States are both disturbingly large and stubbornly persistent. For example, data from the recently released 2015 National Assessment of Educational Progress (NAEP) indicate that, on average, the mathematics knowledge of eighth-grade black and Hispanic students in public schools lags behind their white peers by an amount equivalent to roughly two to three full years of learning (i.e., 0.84 and 0.59 standard deviations, respectively).[1] Black and Hispanic students are also substantially overrepresented among students diagnosed with specific learning disabilities relative to their white peers (Aud, Fox, & KewalRamani, 2010). Furthermore, while roughly 14 percent of white students in public high schools fail to graduate on time, the corresponding dropout rates for black and Hispanic students are roughly *twice* as large (Stetser & Stillwell, 2014). These striking patterns have motivated a broad array of aggressive federal, state, and local policies that have shaped the governance and operations of public schools over the last several decades. These contentious reforms have included different forms of school accountability and choice (e.g., No Child Left Behind, vouchers, and charters) as well as initiatives to promote effective teaching through performance-based compensation systems.

Over the same period, a fast-growing (and largely qualitative) research literature in education has instead focused on classroom pedagogy and stressed the importance of "culturally relevant pedagogy" (CRP) as a compelling way to unlock the educational potential of historically marginalized students (e.g., Ladson-Billings, 1992b, 1994, 1995; Ladson-Billings & Tate, 1995). The fundamental theoretical argument for CRP is that instructional practices are substantially more effective when differentiated to align with the distinctive cultural priors that individual students experience outside of school and when they also affirm both cultural identity and critical social engagement (e.g., Gay, 2010). The "ethnic studies" courses that expanded in the wake of the U.S. Civil Rights Movement

[1] See http://nces.ed.gov/nationsreportcard/naepdata/ for data on the main NAEP scale scores and standard deviations. Bloom et al. (2008) provide guidance on interpreting effect sizes as years of learning.

provide a particularly prominent example of culturally relevant pedagogy. In general, ethnic studies (hereafter, ES) refer to interdisciplinary programs of study that focus on the experiences of racial and ethnic minorities with a particular emphasis on historical struggles and social movements. Apart from the subject's relevance for students who are racial and ethnic minorities, ES courses often incorporate other elements of CRP through their emphasis on cultural identities and conscious engagement with social and political issues (Banks, 1997, 2012; Cammarota & Romero, 2009; Sleeter, 2014; Yosso, 2002, 2005).[2] While some school districts are currently experiencing sustained political controversy over their use of ES curricula (e.g., Tucson), other major urban school districts (e.g., Los Angeles and San Francisco) have begun implementing new ES courses in hopes of supporting the academic achievement of their diverse student populations.

However, the available quantitative evidence on the *causal* effects of ES courses (and, culturally relevant pedagogy, in general) on student outcomes is limited, particularly for larger-scale field settings. This study provides such evidence through examining the effects of a ninth-grade ES course piloted over several years in the San Francisco Unified School District (SFUSD). Specifically, using data on 1,405 students from five school-by-year cohorts, we examine the effects of ES participation on several proximate academic outcomes (i.e., attendance, grade point average, and credits earned) that are highly relevant for high school persistence. Our research design identifies the *causal* effects of taking the ES course by leveraging an institutional feature that was unique to SFUSD. High school students in our study cohorts were assigned to take the ES course if they were identified as at-risk of dropping out (i.e., an eighth-grade GPA below 2.0). We estimate the effects of ES participation through a "regression discontinuity" (RD) design that effectively compares outcomes among students whose eighth-grade GPA placed them just below versus just above this threshold condition. RD designs such as this can credibly support causal inferences because they are based on

[2] In fact, some partially attribute the development of CRP to the academic discipline of ethnic studies (Yosso, Parker, Solorzano, & Lynn, 2004).

the "as good as randomized" assignment to treatment that exists for students proximate to this threshold (D. S. Lee & Lemieux, 2010).

We find that ES participation had large, positive effects on each of our student outcomes. Specifically, ES participation increased student attendance (i.e., reduced unexcused absences) by 21 percentage points, cumulative ninth-grade GPA by 1.4 grade points, and credits earned by 23 credits.[3] These GPA gains were larger for boys than for girls as well as higher in math and science than in ELA. We find that these surprisingly large effects are robust to a variety of model specifications as well as checks for possible confounds related to the treatment contrast we study (e.g., unobserved teacher effects, the possibly independent effects of an at-risk designation, "heaping" of the assignment variable). We also argue that these large effects are consistent with the hypothesis that participation in the course reduced the probability of dropping out in addition to possibly improving the performance of enrolled students. Overall, our findings indicate that a culturally relevant curriculum implemented in a strongly supportive context can be highly effective at improving outcomes among a diverse group of academically at-risk students. However, we also note that the effectiveness of this ES course may reflect other theoretical mechanisms (e.g., buffering students against "stereotype threat") and that there are potentially serious challenges of successfully replicating and scaling up this curriculum.

Cultural Relevance and Ethnic Studies in Theory and Practice

Both academic and popular discussions have long emphasized the role that a community's culture may play in amplifying or ameliorating achievement gaps. For example, an older and largely discredited literature from the 1960s (e.g., Bereiter & Engelmann, 1966; Deutsch, 1967; Hess & Shipman, 1965) suggested that achievement gaps reproduce themselves, in part, because racial and ethnic minorities enter school with a deficit of "cultural capital" (e.g., skills and dispositions related to the dominant culture) that could otherwise support student success. A more contemporary literature

[3] In our main results, we define GPA and credits earned excluding the ES course and *all* other social studies courses (and physical education) to avoid possible confounds related to differences in assessment norms across different courses. We also show results specific to mathematics, science, and English/Language Arts courses.

based on an influential article by Fordham and Ogbu (1986) has advanced the related argument that, in response to discrimination, minority communities develop an "oppositional peer culture" that effectively devalues educational effort and success as "acting white." Several qualitative studies have strongly disputed this cultural characterization (e.g., Horvat & Lewis, 2003; O'Connor, 1997). Moreover, quantitative studies (e.g., Ainsworth-Darnell & Downey, 1998; Akerlof & Kranton, 2002; Cook & Ludwig, 1997; Downey & Ainsworth-Darnell, 2002; Tyson, Darity, & Castellino, 2005) have found little evidence to support the conjectured existence of an "oppositional" culture that contributes to achievement gaps.

Another body of qualitative studies has shifted the focus to evidence that school and classroom practices are frequently misaligned with the cultural priors and out-of-school experiences of minority students (Banks, 1991; Gay, 1988; Ladson-Billings, 1992a; Valenzuela, 1999). Specifically, several anthropological and sociolinguistic studies (e.g., Au & Jordan, 1981; Mohatt & Erickson, 1981) have provided evidence that teachers who are highly effective with minority students adopt culturally "appropriate" or "congruent" methods to engage their students (e.g., through their use of language and the design of classroom activities). In an influential body of work that drew, in part, on this earlier tradition, Ladson-Billings (1992b, 1994, 1995) examined and advocated for the practical and theoretical relevance of "culturally *relevant* pedagogy" (CRP).[4] One key element of CRP is the use of valid cultural referents in teacher practice. However, Ladson-Billings (1992b) argues that CRP does more than "fit" school culture to student culture; it also seeks to "use" student culture as a basis for classroom practice and to enhance both cultural competence and social and political awareness.

Interestingly, independent disciplinary traditions can provide alternative theoretical frames for situating how CRP might be effective in improving the academic performance of minority students. For example, the social-psychological literature on "stereotype threat" suggests that minority students underperform in highly evaluative settings such as classrooms because of the anxiety created by the

[4] This term is used interchangeably with the term "culturally *responsive* teaching" (CRT) (Gay, 2010; Ladson-Billings, 1992b; Sleeter, 2014).

expectation of being viewed through the lens of a negative stereotype (Steele & Aronson, 1995). Several field-based randomized trials of interventions that “buffer” students against stereotype threat have shown promise in reducing achievement gaps, though their efficacy appears to be context-dependent (Aronson & Dee, 2012; Dee, 2015; Yeager & Walton, 2011). Interestingly, the active ingredients in these stereotype-threat buffers (e.g., forewarning about stereotypes, values affirmation, external attribution for experiencing challenges, and growth mindsets) closely parallel the defining elements of CRP. The theoretical logic for CRP can also be understood in a microeconomics framework in which students have imperfect information about their own suitability for academic pursuits. Benabou and Tirole (2003) argue that, in these circumstances, individuals adopt a “looking-glass” perspective in which they come to understand their own place in the world based, in part, on the cues they receive about themselves from others (e.g., schools and teachers). In such a setting, CRP may be effective because both cultural congruence and an emphasis on cultural affirmation and integrity create positive signals about belongingness in school.

As commonly conceived and implemented, ES courses provide a prominent and controversial example of CRP. ES courses focus on the experiences, perspectives, and histories of traditionally underrepresented ethnic or racial groups and have several specific features. They are typically organized around the principal that CRP better engages underrepresented students and meets their needs by drawing on their cultural competencies to promote academic success. That is, ES courses are theorized to positively affect student outcomes through the creation of a relevant and meaningful curriculum that affirms students’ identities, draws from their funds of knowledge, and builds students’ critical intellectualism (Banks, 2012; Cammarota & Romero, 2009; Giroux & Simon, 1989; Sleeter, 2014; Tintiangco-Cubales et al., 2015). To support this type of curriculum, ES courses often adopt alternative organizational and pedagogical structures following central lessons from CRP. For example, many ES courses utilize a classroom structure in which teachers work to promote engagement by structuring collaborative, equitable, reciprocal relationships between themselves and students (Duncan-Andrade & Morrell, 2008; Sleeter, 2014; Tintiangco-Cubales et al., 2015). In addition to content that engages with students' cultural identities, and a student-focused classroom structure, ES courses also draw from critical pedagogies, using an educational praxis to provide students with tools for identifying, reflecting upon, critiquing, and acting against systemic racism and other forms of oppression (Freire, 2000; Giroux & Simon, 1989; Sleeter, 2014; Sleeter & Bernal, 2004). Recent examples of ES coursework guide students in exploring their own identities and engaging with their community, often incorporating assignments that require repeated engagement with community and family members and some type of social activism (Tintiangco-Cubales et al., 2015). Proponents of ES also stress the positive impact that these courses will have on standard educational outcomes such as students' grades, test scores, behavior, and school completion (Cabrera, Milem, Jaquette, & Marx, 2014; Matthews & Smith, 1994; Tintiangco-Cubales et al., 2015).

The first formal ES course was created at San Francisco State University in 1968, growing out of the civil-rights and anti-war movements. However, some argue that ES as an idea has a longer history tracing back to Freedom Schools, Black independent schools, and tribal schools, among others (Begay et al., 1995; C. D. Lee, 1992; Sleeter, 2014). Since their formalization at the post-secondary level, ES programs and curricula have spread to universities across the country, but are still relatively uncommon in secondary schools (Hurtado, Engberg, Ponjuan, & Landreman, 2002). Recently, several school districts have or are considering adopting ES courses as graduation requirements (Gilbertson, 2014; Tucker, 2014). However, the expansion and implementation of ES programs is often highly contentious. Critics often characterize ES programs as divisive, non-academic, and detrimental to students of color because they are substituting courses that promote the development of ethnic pride in place of the development of mainstream academic skills (Sleeter, 2014). When schools, colleges, and universities offer such courses or programs of study, they often become a contentious political flashpoint. For example, the school district in Tucson, Arizona, which had offered courses in Mexican-American studies, was recently found in violation of a new state law preventing the teaching of such courses as they "promote the overthrow of the United States government," "promote resentment toward a race or class of people" and "advocate ethnic solidarity instead of the treatment of pupils as individuals," (formerly *Arizona HB 2281*, 2010, *Arizona Revised Statute § 15-112*, 2010) and subsequently eliminated this programming under threat of losing state funding (Billeaud, 2011). Student protests of the school board meeting debating this policy and the ensuing controversy were covered by a diverse segment of the national media, including Fox News, *The Daily Show with Jon Stewart*, and the *New York Times* (Cabrera, Meza, Romero, & Rodríguez, 2013).

At the same time, other districts have expanded or are considering expanding their ES offerings. For example, the Los Angeles Unified School District and the El Rancho Unified School District recently included ES courses in their high school graduation requirements (Tucker, 2014). Recently introduced legislation in California would also require all high schools to offer ES courses (Clark, 2015; Gilbertson, 2014; Tucker, 2014).[5] The Texas State Board of Education also recently approved legislation allowing school districts to develop courses on Mexican-American studies (Isensee, 2014). In addition, the Berkeley Unified School District has offered a freshman ES course for over 20 years, requiring it for high school graduation during nearly all of this time (Artz, 2003; Levin, 2009; Noguera, 1994; Rubin et al., 2006; Veale, 2015). As we describe in the next section, the motivating context for this study is that the San Francisco Unified School District (SFUSD) was considering scaling-up access to a pilot ES curriculum and, possibly, requiring it as a graduation requirement.

While the expansion of ES courses illustrates both their appeal and concerns, the quantitative evidence on their effects is relatively limited. Furthermore, the evidence that is available relies on research designs that cannot necessarily support credible causal inference.[6] For example, a small-scale

[5] However, the governor of California, Jerry Brown, recently vetoed separate legislation that would have required the state to develop a model ES program for California's public schools and make it available to local districts (Ceasar, 2015).
[6] This appears to be true of CRP, more generally. One possible exception is a recent randomized trial by Kisker et al. (2012), which found that a culturally relevant math curriculum significantly improved the performance of second-grade Alaskan Natives. However, these gains may conflate the effects of general instructional quality as well as cultural

descriptive study by Cammarota (2007) focused on the "Social Justice Education Project" (SJEP), a "sub-curriculum" fielded among 17 at-risk Latina/o students in a Tucson high school over four semesters between 2003 and 2005. Cammarota (2007) reports that these students were successful both in completing high school and in engaging with advanced courses. A study by Lewis, Sullivan, and Bybee (2006) examined the effects of an "Emancipatory Education" course fielded over one semester among n=65 eighth-grade students in an urban, predominantly black school. They randomly assigned one of the two participating classes to receive this intervention and found positive effects on communal orientation, school connectedness, and achievement motivation. However, the availability of only two assignment units (and the lack of evidence on balance at baseline) makes it difficult to differentiate the true effects of the course from the effects of other unobserved traits that may have differed across these two classrooms.

Two other studies have relied on regression analyses of administrative data from the larger-scale implementation of ethnic studies in Tucson, Arizona. First, a brief report from the Arizona Department of Education (Francosi 2009) compared the test performance of Hispanic students in Tucson who took one or more ES course in the 2008-09 school year with Hispanic students statewide in regressions that controlled for other student traits (e.g., prior performance, mobility, and English learner status). This analysis found no evidence that course participation improved student performance. A more recent study by Cabrera et al. (2014) relied on administrative data from roughly 8,400 students over four cohorts (i.e., the graduating classes of 2008-2011) to examine the Mexican-American studies (MAS) program offered in four schools in Tucson.[7] In regression analyses that control for student demographic characteristics (race/ethnicity, gender, free/reduced price lunch eligibility, Census block median income, ELL, Special Ed, and GATE status, number of school transfers), prior academic achievement (ninth- and tenth-grade weighted GPA, tenth-grade

relevance. The intervention included teacher training and the intervention also improved the performance of students who were *not* Alaskan Natives.

[7] As noted by Cabrera and colleagues the development of this program was technically unrelated to AB 2281 and was instead a solution to a 40-year-old desegregation order for TUSD.

standardized test scores), and school-level context (school fixed effects), they find evidence that MAS participation improved student outcomes.[8] In particular, participation in MAS was associated with an increase in the probability of graduation of 9.5 percent across all cohorts. Among the subsample of students who initially failed the exit exam, MAS participation was associated with a 6.6 percent increase in the probability of passing the all three exit exams (the reading, writing, and math AIMS tests) on average across all cohorts.[9]

A central challenge to these empirical studies is that participation in the MAS program was voluntary. Thus, regression-adjusted comparisons among those who did and did not enroll may suffer from omitted variable biases of an uncertain direction. For example, if students who have a latent and unobserved capacity for school engagement are more likely to enroll in these courses, naïve regressions may overstate the program's benefits. In contrast, if at-risk students are more likely to be enrolled in MAS courses, their impact is likely to be understated. Cabrera et al. (2014, page 1094) discuss these methodological challenges and acknowledge the limitations of their study noting "our results may suffer from omitted variable bias and should not be considered true causal effects."

In sum, the theoretical arguments and public enthusiasm for ES curricula have not been matched by convincing quantitative evidence on their efficacy. Our study contributes to this gap in the literature by employing a research design that can credibly support a strong causal warrant. Specifically, we rely on an explicit student assignment rule to identify the causal effects of a year-long ES course in a regression discontinuity (RD) design. Our study is also unique in that it focuses on a mature, developed course situated within a novel setting (i.e., high schools in the San Francisco Unified School District). We describe our study context and research design in more detail below.

---

[8] This primary analytic sample is restricted to students who are in schools that offered the MAS curriculum. A secondary set of analyses includes nearly 17,000 students in all TUSD schools, including those without MAS programs.

[9] However, cohort-specific results suggest that this association was not significant for all tests in all years, particularly in the final 2011-2012 cohort. The authors speculate that the political turmoil surrounding the program in this year might have weakened its effectiveness, or that the expansion of MAS offerings to additional schools might also have contributed to the lack of significant results.

Ethnic Studies in the San Francisco Unified School District (SFUSD)

The genesis of the SFUSD ES curriculum was in 2007 when the District's Board of Education Curriculum Committee urged the district to create a high school ES curriculum. The District's Office of Learning Support and Equity, in collaboration with faculty from the College of Ethnic Studies at San Francisco State University (SFSU), subsequently initiated the curriculum design. Specifically, ten SFUSD social studies teachers formed the "Ethnic Studies Curriculum Collective" with SFSU faculty support. This group created a course framework drawing from ES curricula used in other districts and post-secondary programs across the country during the 2007-2008 school year. Over the next two years, the Collective created lesson plans, piloted the lessons in three high schools and met twice a month for lesson critique and development (SFUSD Ethnic Studies Curriculum Collective, 2012).

On February 23, 2010, the SFUSD school board unanimously approved a resolution to implement an ES pilot program in SFUSD high schools, explicitly referencing the promise of ES courses to contribute to closing achievement gaps. Five high schools participated in the pilot, offering a year-long, ninth-grade ES course from the 2010-2011 to 2012-2013 school years. The program continued into the 2013-2014 school year. In December of 2014 (i.e., after our study window), the school board voted to expand the program to be offered at all 19 of San Francisco's high schools. It is also being considered as a ninth-grade graduation requirement (Dudnick, 2014).

The design of SFUSD's ES course stressed the use of CRP as a way to engage with students that had previously felt marginalized by the traditional curriculum. Units focused on themes of social justice, discrimination, stereotypes, and social movements from U.S. history spanning the late 18th century until the 1970s. The course also encouraged students "to explore their individual identity, their family history, and their community history" and required students to design and implement service-learning projects based on their study of their local community. The designers of this curriculum hoped that these lessons and projects would increase students' commitment to social justice and improve self-esteem. In addition to the civic and psychological goals of the ES program, the program's stated intent was to close achievement gaps and reduce dropout rates (Office of Learning Support and Equity/Humanities, Academics and Professional Development, 2009; SFUSD Ethnic Studies Curriculum Collective, 2012).

While the ES curriculum was under development for several years and across several different high schools in San Francisco, the assignment of students varied. Some of the pilot schools chose to offer the ES course to all incoming ninth graders, while other schools used the program as an intervention for students identified as at-risk for academic failure through an early-warning system. The early-warning indicator (EWI) flagged students who, in eighth grade, had either an attendance rate below 87.5 percent or a GPA (excluding physical education) below 2.0. Prior research had shown that, in SFUSD, these binary variables were highly predictive of dropping out of school. In our data, very few students had an attendance rate below the 87.5-percent threshold so the relevant "assignment variable" in our RD design is the eighth-grade GPA.[10] Students whose eighth-grade GPA was below 2.0 were encouraged but not compelled to take the ES course. This partial compliance implies that our RD design is "fuzzy" and that there may be external-validity caveats to our inferences if the effect of taking the ES course is heterogeneous (Imbens and Angrist 1994). We take up this and other related issues after describing our data and methods below.

Data

We examine the impact of SFUSD's year-long ninth-grade ES course on student outcomes, primarily using data from three of the five high schools that piloted the curriculum. These three high schools assigned only some ninth-grade students, while two other schools chose to offer the ES course to all ninth-grade students. These schools typically offered two and four sections of the course in each year, although the course was not offered in all schools in every year. Our primary study sample draws from five unique school-year cohorts in these three high schools. In these five cohorts, enrollment in

[10] We exclude the few students with attendance rates below the threshold from our analysis. This implies that we are estimating a "frontier" RD (Reardon and Robinson 2012).

ES was encouraged, but not required, for students whose eighth-grade GPA was below 2.0. Students identified by the early-warning indicators as at risk of high school failure were automatically enrolled in the ES course when they received their course schedule at the start of their ninth-grade year. Students could opt out of the course after consulting with their academic counselor, but needed to actively select out of the course to do so.[11] One school used this rule over 3 years (i.e., AY 2011-12 though AY 2013-14) while two other schools used this in AY 2011-12 only. Critically, only 4 unique teachers taught the ES courses in these schools and years. We discuss, along with our other robustness checks, evidence indicating that our results are not simply due to effects unique to the effectiveness of these teachers.[12]

Our initial sample consists of ninth graders in these five school-year cohorts. However, we exclude those who are missing our assignment variable: a recorded eighth-grade GPA (n = 226). We also exclude a cluster of 128 students with eighth-grade GPAs that are distant from the threshold and clustered at a perfect 4.0 GPA. We also exclude a small number (n=27) of additional students with extremely low eighth-grade GPAs (i.e., less than 1.25).[13] These sample edits imply a final "intent-to-treat" (ITT) sample of 1,405 students. Our data on these students include several measures of baseline traits. These include binary indicators for gender and for whether the student was black, Hispanic, or Asian (with white serving as the reference category). We also have eighth-grade data on whether the student was in special education, ever suspended, or identified as an English Language Learner (ELL). We also have data on each student's attendance rate in eighth-grade, the value of their assignment variable (i.e., eighth-grade GPA exclusive of PE and centered on 2.0), and a binary indicator for our "intent-to-treat" (ITT) variable (i.e., an eighth-grade GPA less than 2.0).

[11] Similarly, students who had not been identified using the EWI system could opt into the course after consulting with their counselors if they desired to enroll, but were not automatically assigned to the course.
[12] Specifically, we examine the effectiveness of these teachers relative to their peers in other courses (i.e., other than ethnic studies).
[13] Based on the limited data available, we suspect some of the students with very low eighth grade GPA have unique special-education circumstances or missing data.

Table 1 presents descriptive statistics on these students. Interestingly, 60 percent of these students are of Asian descent and 23 percent are Hispanic. Only 6 percent of these students are black. Eighteen percent of these students are identified as ELLs and 12 percent have special education status. Among the cohorts in our sample, only 42 percent are female. This is due in part to the fact that there are fewer female students than male in the district overall (48 percent across all SFUSD schools), but particularly because female students are higher-achieving than our sample (recall that we exclude students who receive a perfect 4.0, which drives most of the difference in female representation between the full district and our sample). Thirteen percent of students enrolled in the ES course and 8 percent of the sample had an eighth-grade GPA below 2.0 (i.e., an intent-to-treat as taking the ES course).[14]

We examine three dependent variables in our analyses, ninth-grade attendance rates (which the district refers to as instructional time), ninth-grade GPA, and ninth-grade credits earned. The last two measures are defined exclusive of all social studies courses (i.e., like the ES course) and physical education. We also control for eighth-grade attendance and GPA in our models. While the average attendance rate increases slightly between eighth and ninth-grade (from 96.32 percent to 96.69 percent), GPA declines substantially during this important transition. The mean eighth-grade GPA is just above a 3.0 (a “B” on the four-point scale), by the end of ninth grade, the average GPA is 2.65 (a “C” on the four-point scale).

We measure these outcomes for *all* students observed at baseline in our intent-to-treat sample regardless of whether they completed ninth grade. So, we view the variation in these measures as reflecting both the academic progress of enrolled students and the probability a student has dropped out of school. For students to advance from ninth to tenth grade, they must complete at least 55 credits. Because we exclude physical education (which would account for 10 credits) and social studies (which would account for an additional 10), students should complete at least 35 credits by our measure in

[14] We define treatment uptake as being enrolled in the first-semester ES course, regardless of whether a student remained in the course.

order to advance to tenth grade. In our sample, we find that 7.3 percent of students have fewer than 35 credits at the end of ninth grade (i.e., suggesting they dropped out or were required to repeat ninth grade). Furthermore, the students at risk of dropping out tend to be concentrated among those encouraged to take ES. However, we also find evidence our results also reflect changes in the performance of enrolled students. In particular, we find virtually similar results to those we report below when we rely only on GPA from the *first* semester.

Regression Discontinuity (RD) Design

Our research design effectively compares those who were just eligible for assignment to the ES course (i.e., eighth-grade GPA below 2.0) to those who were just ineligible for this assignment (i.e., eighth-grade GPA at 2.0 or above). Specifically, we use a regression discontinuity (RD) design, which can provide causal inferences that are "as good as random assignment" (Lee and Lemieux 2010) in settings like this. An RD design asks whether, conditional on a students' eighth-grade GPA, student outcomes "jump" at the threshold that defined treatment eligibility (i.e., assignment to ES). The RD design is implemented by estimating reduced-form equations of the following general form:

$$Y_{ist} = \alpha + \beta I(G_{ist} < 0) + f(G_{ist}) + \lambda X_{ist} + \eta_{st} + \varepsilon_{ist}$$

where Y_{ist} is a student outcome (e.g., GPA) for ninth grader i in school s in year t. The variable, G_{ist}, is the "assignment variable" in this RD design: eighth-grade GPA centered on 2.0. The parameter of interest, β, identifies the jump in outcomes when eighth-grade GPA is below 2.0, conditional on $f(G_{ist})$, a smooth function of the assignment variable. We specify $f(G_{ist})$ as linear but allow for different slopes above and below the threshold.[15] We also explore flexibly non-parametric specifications (i.e., local

[15] We also examined models that added quadratic terms for the assignment variable. However, a comparison of Akaike information criterion (AIC) across these specifications privileged the linear specifications.

linear regressions).[16] The variable, X_{ist}, refers to student-level controls and η_{st} refers to fixed effects unique to each year at a particular school. We also rely on heteroscedastic-consistent standard errors.

In Table 2, we present the RD results from examining whether actually taking the ES course does indeed jump at the 2.0 threshold. We find robust evidence that the likelihood of taking the ES course jumps roughly 27 percentage points at the threshold. Figure 1 illustrates this finding graphically by showing the probability of taking ES as a function of eighth-grade GPA. This figure organizes the data in bins of width 0.1 defined by eighth-grade GPA. The top panel uses the full sample while the bottom panel uses data within a 0.7 GPA bandwidth of the threshold. These figures consistently illustrate the jump in treatment status at the threshold. However, they also underscore that, as is common in RD and experimental settings, we have partial compliance with the intent-to-treat implied by an eighth-grade GPA below 2.0. Roughly 20 percent of students with eighth-grade GPAs of 2.0 or slightly higher took ES while just over 50 percent of students below the threshold did so. This partial compliance does not confound the internal validity of the RD design because the identifying variation is based on eighth-grade GPA rather than the decision to take the course. In other words, our reduced-form estimates identify the effect of being assigned to take the ES course (i.e., the "intent-to-treat" effect) rather than the effect of taking the course. However, we can recover the estimated effect of actually *taking* the ES course (i.e., the "treatment-on-the-treated" effect) by dividing our reduced-form impact estimates by the corresponding treatment uptake at the threshold (i.e., roughly 0.25).

The fundamental treatment contrast leveraged in our study is among students eligible for assignment to the ES course and those who were not. To avoid any confounds related to different grading and attendance standards across the alternative courses students around this threshold took, we define our GPA and credits-earned measures excluding data from the ES course and all other social studies courses. A related concern is that taking ES may imply that a student takes different courses in *other* subject areas. However, we found that virtually all students were initially enrolled in math, ELA,

[16] More specifically, we show our key results using only the data in increasingly tight bandwidths around the threshold. We also found that estimates based on the optimal-bandwidth procedure introduced by Imbens and Kalyaranaman (2012) generated similar results.

and science courses and that course selection in these subject areas did not differ for students around the 2.0 GPA threshold.[17] We also present results using GPA measures specific to each of these three subjects.

The strong causal warrant of the RD design is based on the assumption that students' locations just above and below the 2.0 threshold are conditionally random. One compelling way to check this key assumption is by examining whether outcome-relevant student traits jump at the threshold. In Table 3, we present the key results from auxiliary regressions that examine this. Specifically, we present the results from RD regressions where the student observables are the dependent variables. The estimated jumps in these variables at the 2.0 threshold are consistently small and statistically insignificant.

A related concern in RD designs is whether students differentially manipulate their eighth-grade GPA to place themselves on one side of the 2.0 threshold. In general, efforts to raise the value of a forcing variable do not invalidate an RD design (Lee and Lemieux 2010). However, if individuals can systematically manipulate their position relative to the threshold, it can impugn an RD's internal validity. This is a unique concern in this context because eighth-grade GPA scores "heap" at a value of 2.0 and other integer and half-integer values (see Figure 5a). Students who earn an eighth-grade GPA of 2.0 may differ from those just below this value in unobserved ways that are relevant eighth-grade outcomes. The covariate balance at the threshold suggests that this is not an internal-validity threat. However, we also report results based on samples where we eliminated heaped observations. We also see (Figure 5b) that, when we eliminate these heaps, the distribution of observations is smooth at the threshold (McCrary, 2008).

Two other internal-validity concerns are unique to our study context. One is that our RD contrast may also identify any effects related to being flagged by an early-warning indicator. One way we examine this concern is to estimate our basic RD design using data from the other San Francisco

[17] Ninth-grade students in SFUSD typically take a ninth-grade English course, either Algebra 1 or Geometry, and either Biology or Physics.

high schools that did **not** offer ES over this period. If our RD design is valid, we expect to find null results at the GPA threshold in these schools. However, if the specification were misspecified, we might find effects unique to the threshold in these schools. Similarly, if early-warning status had independent effects, we would expect to find evidence in these schools. A second concern is that our RD framework may identify the effect of the four unique teachers in our study sample rather than the effect of the course per se. We investigate this issue by examining the comparative effectiveness of these teachers in the other courses they taught. We discuss these and other critical robustness checks as we outline our results below.

Main Results

Table 4 presents the main RD results examining the effects of ES eligibility on ninth-grade attendance, GPA, and credits earned. The baseline specification (i.e., the first column for each of the three outcomes) controls for the variable of interest (i.e., a binary indicator for whether the student had an eighth-grade GPA below 2.0), eighth-grade GPA, and a linear spline that allows this assignment variable to have distinct effects above and below the threshold. The subsequent specifications introduce controls for gender, race/ethnicity, and eighth-grade special education and ELL designations, eighth-grade attendance, and whether the student was ever suspended in eighth grade. These saturated specifications yield largely similar results, although the magnitude of the point estimates is reduced somewhat. Results from the most parsimonious to the most inclusive specifications consistently indicate that students with eighth-grade GPAs at the 2.0 threshold saw statistically significant improvements on all three ninth-grade academic outcomes. Drawing from the most unrestrictive model, we find robust evidence that attendance jumped by 5.6 percentage points for students at the 2.0 threshold, GPA increased by 0.39 points, and credits earned increased by 6.3 credits.

Figures 2, 3, and 4 provide graphical illustrations of these RD results. Figure 2 plots students' eighth-grade GPA scores by their ninth-grade attendance, with a line indicating the 2.0 GPA cutoff.

Figure 3 plots the relationship between eighth-grade GPA and ninth-grade GPA (excluding social studies and P.E.). Figure 4 plots the relationship between eighth-grade GPA and ninth-grade credits earned. Each of the figures shows a discontinuity at the 2.0 threshold, echoing the regression results shown in Table 4.

The instrumental-variable (IV) estimates implied by these results indicate that *taking* ES increased attendance by 21 percentage points, GPA by 1.4 grade points, and credits earned by 23 credits (or roughly four courses). We calculate these estimated effects of taking ES by inflating the effects of ES *eligibility* on academic outcomes (Table 4) by the effect of ES *eligibility* on ES take-up. This amounts to multiplying the reduced-form effects in Table 4 by roughly 3.7 (i.e., the inverse of the jump in ES uptake at the threshold in Table 2, Column 3). These effect sizes (i.e., roughly 1.5 to 2.0 of the corresponding standard deviations in Table 1) are quite large for interventions situated in field settings. However, several considerations should be noted. First and foremost, because we define these outcome measures for all students observed at baseline, some of these striking gains are likely to reflect reductions in dropping out as well as gains in the performance of enrolled students. Second, RD estimates like ours are effectively defined for students close to the 2.0 GPA threshold. These tend to be students who are at considerable academic risk so larger gains in academic performance are possible. We take up such issues of treatment heterogeneity after first exploring the robustness of our main findings.

Robustness Checks

Given the consistent, large findings across a variety of ninth-grade outcomes, we next turn to examining the robustness of the apparent effects associated with the eighth-grade 2.0 GPA discontinuity. One possible confounding explanation for these findings is that they reflect the effects of the early-warning indicator (EWI) rather than the ES course. In other words, students might be receiving other services and interventions as a result of the EWI identification and this designation or these services might be driving changes in student outcomes rather than ES. To examine this concern, we estimated the same RD specifications using similarly constructed data from SFSUD high schools that did **not** offer an ES course. We present these results in Table 5. The small and statistically insignificant coefficients for each specification and for each of the three outcomes (i.e., there are no jumps at the 2.0 threshold in these schools) indicate that EWI did not have an empirically meaningful effect on ninth-grade outcomes. These null results are consistent with the hypothesis that the Table 4 results reflect the effects of taking ES rather than the effects of an EWI designation.

An additional concern is related to the fact that student grades are reported in even grade points, leading to large clusters of students with GPAs at even-integer or half-integer GPA values (e.g., 3.0 and 3.5 rather than 2.99). As has been shown in other work using regression discontinuities to estimate causal effects, results can be biased by this heaping of the assignment variable (Barreca, Guldi, Lindo, & Waddell, 2011). We present several robustness specifications in Table 6 to examine whether our results are being driven by the preponderance of even and half-integer eighth-grade GPAs by excluding students with several specific values. In these "donut RDs" we first exclude students with eighth-grade GPAs of 2.0 exactly. In a second version, we exclude students with any whole- or half-integer value for their eighth-grade GPA. For each of the ninth-grade academic outcomes, the point estimates presented in Table 6 are from individual regressions for the variable eighth-grade GPA is less than 2.0, akin to the point estimates shown in Table 4 from models including student controls, with the first row replicating these estimates exactly.

The results in Table 6 show that our inferences are robust in specifications that exclude students whose eighth-grade GPA fell on the heaped values of 2.0 as well other integer and half-integer values. Each of the coefficients for all three of the ninth-grade academic outcomes is statistically significant at the 5 percent level and the magnitude of the coefficients is fairly consistent whether or not the students with GPAs of 2.0 or any integer or half-integer value are included in the sample.

Table 7 presents another important robustness check based on restricting the estimation sample to observations in increasingly tight bandwidths around the threshold for both the first-stage and reduced-form effects. These results provide evidence about whether the results are biased due to functional-form assumptions or are unduly influenced by observations that are far from the 2.0 GPA threshold. The results in Table 7 indicate that both the first-stage and reduced-form estimates are robust as the sample shrinks with each of the progressively tighter bandwidths, including a bandwidth that is within half of a grade point from the 2.0 threshold. If anything, the first-stage and reduced-form estimates are larger as the bandwidth tightens.

Table 8 presents another robustness check based on simultaneously estimating jumps at the GPA threshold that actually influenced assignment to the ES course, 2.0, and at other “placebo” thresholds that have no relevance. We examine six placebo thresholds at each quarter-integer interval between GPAs of 1.0 and 3.0. Across both the first-stage and reduced-form estimates, the only statistically significant effects are observed at the 2.0 threshold, with one exception. Students at the 2.25 GPA threshold, just below the cutoff 2.25 cutoff, earn significantly fewer ninth-grade credits than students on the other side of this cutoff. With this exception, the nearly universal lack of statistically significant effects at these false thresholds is consistent with the absence of specification error.

A final robustness check stems from the particular implementation of the ES curriculum in SFUSD. While ES was piloted at five high schools over several years, assignment to ES was based on the EWI in only five school-year cohorts at three schools. In each of these school-year cohorts, only one teacher taught ES, leaving us with a total of four unique teachers during our study window. This raises the possibility that the effects we observe are the result of effects unique to these teachers rather than the ES curriculum itself. To investigate this concern, we examined the effectiveness of ES teachers relative to their peers, when teaching courses *other than* ES. We began by identifying all of the non-ES courses taught by our four ES teachers in any of the study years and then identified all of the other teachers of those same courses. The majority of these courses were social studies courses, such as U.S. and world history, but the list also included some college counseling and homeroom-type courses, which we chose to exclude from the analysis. We focused on students in these social studies courses who had *not* taken ES. We then recovered teacher fixed effect estimates from regression models predicting each of our ninth-grade student outcomes (ninth-grade overall and subject-specific GPA, credits earned, and attendance), conditional on eighth-grade student controls and school by year fixed effects. For each of our outcomes, we examined the relative rankings of the teacher fixed effects to determine if the ES teachers were over-represented among teachers who had the largest fixed effects estimates. Across each of the outcome measures, we found the value-added estimates of the ES teachers to be quite uniformly disbursed throughout the distribution of teacher fixed effect estimates.[18] Wilcoxon rank-sum tests further suggest that the fixed effect estimates of ES teachers are not significantly different from those of non-ES teachers in the same subjects. Of the four ES teachers, on fairly consistently had the largest fixed-effect estimate. To ensure that this generally more effective teacher was not driving our results, we re-estimated our key ES results without this teacher. Doing so did not qualitatively alter the previously reported findings.

Treatment Heterogeneity

Our main impact estimates may obscure several forms of treatment heterogeneity that are worth noting and exploring. For example, one well-known caveat about external validity involves the "localness" of RD estimates. That is, because our research design leverages the targeting of ES courses to at-risk students, our resulting estimates may not speak to the effects these courses may have on students with high-performance in eighth grade.[19] Second, the impact of taking the ES course could conceivably vary across students with different demographic traits. In Table 9, we present evidence on this issue by showing the first-stage and reduced-form estimates in samples defined by race, gender,

---

[18] For example, for ninth-grade GPA, the ranks of the four ethnic-studies teachers were 9, 14, 30, and 35 among 37 total social studies teachers.
[19] The in-progress scale-up of this course across SFUSD high schools may provide opportunities to explore this heterogeneity.

and ethnicity. The point estimates show that there are consistently positive effects across male, female, Asian and Hispanic groups of students. However, the improved outcomes are particularly concentrated among boys and statistically insignificant for girls. For Hispanic students, the estimated effects are consistently large and statistically significant across all ninth-grade outcomes. For Asian students, while each of the point estimates is positive, they are only significant for the first-stage and ninth-grade instructional-time effects. This suggests that, while the ES course is not harmful for any of the enrolled students, it is particularly good for male students and Hispanic students.

In Table 10, we examine whether there are heterogeneous effects on student GPA by subject. Each cell in this table reports the key RD estimate (i.e., the estimated "jump" at the 2.0 threshold) from a unique regression. The first column presents point estimates conditional on linear splines of the assignment variable and on school-by-year fixed effects. The subsequent models introduce student and eighth-grade covariates. The point estimates show that there are consistently positive, statistically significant effects on GPA specific to math and to science, despite the distal nature of their respective content to that of ES. However, in ELA, while the point estimates remain positive, they are smaller and statistically insignificant.

The literature on causal inference has also recently emphasized another possible (and subtler) form of treatment heterogeneity based on the potential-outcomes framework and how individuals respond to their intent-to-treat (i.e., as "compliers", "always takers", and "never takers"). Specifically, Imbens and Angrist (1994) show that, when treatment effects are not homogenous across these groups, estimates like ours are "local average treatment effects" (LATE). Such LATE estimates identify the effect of the treatment for those who comply with their intent-to-treat but not necessarily for those who always (or never) take up the treatment regardless of the intent-to-treat. A recent study by Bertanha and Imbens (2014) provides straightforward guidance on assessing the empirical relevance of this possible treatment heterogeneity in "fuzzy" RD applications like ours. Specifically, they recommend estimating the reduced-form RD specifications for separate samples defined by whether the student took up the treatment (i.e., ES $\in$ 0,1). We report these results in Table 11 using our saturated model (i.e., column 3 in Table 4). In the first row, we repeat our full-sample results as a point of reference. In the second row, we show the estimated "jump" in outcomes using only data from students who did not take ES (i.e., ES = 0). In this sub-sample, the threshold separates never-takers (i.e., to the left of the threshold) from the population of never-takers and compliers who are to the right of the threshold. The fact that outcomes are higher to the left of the threshold (i.e., for at least two of the three outcomes) indicates that never-takers have unobserved traits that predispose them to better student outcomes relative to compliers. Intuitively, this finding suggests that students who insist on taking a health or college preparation/study skills course in lieu of ES have unobserved traits that imply better academic outcomes.

The next row identifies the jump at the threshold for each outcome measure using only data on students who took ES (i.e., ES = 1). The population to the left of the threshold consists of compliers and always-takers while the population to the right only contains always-takers. Our evidence that each student outcome jumps significantly at the threshold could indicate that taking the course is more effective for those who only take it when assigned relative to those who insist on taking it. This could occur, for example, if culturally relevant pedagogy is less novel and relevant for the types of students who insist on taking it. Overall, these findings are consistent with the type of heterogeneity implied by the LATE theorem. As a practical matter, this evidence of treatment heterogeneity has salience for the external validity we might expect when scaling up access to this course. In particular, these findings suggest that taking the course is less necessary for the type of student who refuses to take the course (i.e., never-takers) and less effective for students who insist on taking it when available (i.e., always-takers). We revisit issues of scalability in our concluding remarks.

Discussion

The results presented in this study indicate that the ninth-grade ES curriculum implemented in SFUSD led to large and statistically significant improvements in students' ninth-grade GPA, attendance, and credits earned. To our knowledge, this is the first study to examine the effect of any type of culturally relevant pedagogy (CRP) in a quantitative study that supports credible causal inferences. Specifically, our "regression discontinuity" (RD) design leveraged a class-assignment rule that encouraged academically at-risk students (i.e., those with eighth-grade GPA below 2.0) to take the course. We present several forms of evidence that affirm the validity of this discontinuous assignment rule as a quasi-experiment as well as evidence on the robustness of our main findings. We note evidence that these large effects appear to reflect both reductions in the probability of dropping out as well as improvements in the performance of enrolled students. We also find that the effects of this course were concentrated among males, Hispanics, and to a lesser degree, Asians.

Taken at face value, these findings provide a compelling confirmation of an extensive literature that has emphasized the capacity of CRP to unlock the educational potential of historically marginalized students. However, we also stress that our results are consistent with other theoretical frames as well. In particular, a field-experimental literature in social psychology has shown that quite modest interventions that buffer students against stereotype threat can, under the right circumstances, dramatically improve student outcomes. ES courses combine several of the active ingredients of these interventions (e.g., affirmation, external attribution for difficulties, forewarning about stereotypes) and expose students to them in an exceptionally intense and persistent manner (i.e., through a year-long course rather than a brief exercise). Furthermore, SFUSD's ES course was also targeted in a manner consistent with such "buffering" interventions (i.e., at the beginning of the school year and during a possibly difficult transition to high school). Further research that can measure alternative mediators can provide insight into the relevance of different theorized mechanisms.

As a matter of policy and practice, this study's findings should be interpreted in light of several important caveats related to external validity and scalability. First, as in all RD studies, our results focus on localized comparisons between students who are just above and below the eligibility threshold for ES enrollment. It is, thus, an open question whether the effects of this or any other ES curriculum would generalize to higher-performing students. Furthermore, we also find evidence that the benefits of taking such a course are larger among those who comply with the encouragement to take the course (i.e., relative to students who would always take it when available).

There are also several reasons to be cautious about the likely impact of scaling up or replicating this ES course. The implementation of ES in SFUSD was, arguably, conducted with a high degree of fidelity, forethought, and planning. In particular, it appeared to draw upon the work of a core group of dedicated teachers, engaging in a regular professional learning community, with outside support from experts in the subject to create and sustain the program. As scholars from a number of disciplines have noted that the effects of such smaller-scale interventions are often very different when the same policies are implemented at scale (Dodge, 2011; Welsh, Sullivan, & Olds, 2010). The broader school, district, and community contexts in which this course was situated may also be relevant. For example, the literature on stereotype threat stress that the success of buffering interventions depends critically on settings that can enhance and encourage positive "recursive" processes related to student engagement and success (Yeager and Walton 2011). Nonetheless, SFUSD's ES program appears to constitute an important proof of concept, indicating that culturally relevant pedagogy can be extraordinarily effective in supporting the academic progression of struggling students.

References

Ainsworth-Darnell, J. W., & Downey, D. B. (1998). Assessing the oppositional culture explanation for racial/ethnic differences in school performance. *American Sociological Review*, *63*(4), 536–553.

Akerlof, G. A., & Kranton, R. E. (2002). Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature*, *XL*, 1167–1201.

Arizona HB 2281 (2010).

Arizona Revised Statute § 15-112 (2010).

Aronson, J., & Dee, T. (2012). Stereotype threat in the real world. *Stereotype Threat: Theory, Processes, and Application*, 264–279.

Artz, M. (2003, December 12). BHS to Keep Ethnic Studies. Category: Page One from The Berkeley Daily Planet. *The Berkeley Daily Planet*, p. 1. Berkeley, CA.

Aud, S., Fox, M. A., & KewalRamani, A. (2010). Status and Trends in the Education of Racial and Ethnic Groups. NCES 2010-015. *National Center for Education Statistics*.

Au, K., & Jordan, C. (1981). Teaching reading to Hawaiian children: Finding a culturally appropriate solution. *Culture and the Bilingual Classroom: Studies in Classroom Ethnography*, 139–152.

Banks, J. A. (1991). A curriculum for empowerment, action, and change. In C. E. Sleeter (Ed.), *Empowerment through multicultural education* (pp. 125–141). Albany, NY: State University of New York Press.

Banks, J. A. (1997). *Teaching strategies for ethnic studies* (2nd ed.). Boston: Allyn and Bacon.

Banks, J. A. (2012). Ethnic studies, citizenship education, and the public good. *Intercultural Education*, *23*(6), 467–473.

Barreca, A. I., Guldi, M., Lindo, J. M., & Waddell, G. R. (2011). Saving Babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics*, *126*(4), 2117–2123.

Begay, S., Dick, G. S., Estell, D. W., Estell, J., McCarty, T. L., & Sells, A. (1995). Change from the inside out: A story of transformation in a Navajo community school. *Bilingual Research Journal*, *19*(1), 121–139.

Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, *70*(3), 489–520.

Bereiter, C., & Engelmann, S. (1966). *Teaching disadvantaged children in the preschool*. Prentice-Hall Englewood Cliffs, NJ.

Bertanha, M., & Imbens, G. W. (2014). "External Validity in Fuzzy Regression Discontinuity Designs" National Bureau of Economic Research Working Paper No. 20773.

Billeaud, J. (2011, December 27). Ariz schools' ethnic studies program ruled illegal. *Associated Press*. Retrieved from http://news.yahoo.com/ariz-schools-ethnic-studies-program-ruled-illegal-021635252.html

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, *1*(4), 289–328.

Cabrera, N. L., Meza, E. L., Romero, A. J., & Rodríguez, R. C. (2013). "If there is no struggle, there is no progress": Transformative youth activism and the school of Ethnic Studies. *The Urban Review*, *45*(1), 7–22.

Cabrera, N. L., Milem, J. F., Jaquette, O., & Marx, R. W. (2014). Missing the (Student Achievement) Forest for All the (Political) Trees Empiricism and the Mexican American Studies Controversy in Tucson. *American Educational Research Journal*, *51*(6), 1084–1118.

Cammarota, J., & Romero, A. (2009). The social justice project: A critically compassionate intellectualism for chicano/a students. In W. Ayers, T. Quinn, & D. Stovall (Eds.), *Handbook for Social Justice Education*. New York: Routledge.

Ceasar, S. (2015, October 12). California Gov. Jerry Brown vetoes ethnic studies bill - LA Times. *Los Angeles Times*. Los Angeles, CA. Retrieved from http://www.latimes.com/local/lanow/la-me-ln-brown-ethnic-studies-bill-20151011-story.html

Clark, S. (2015, January 13). New bill would require ethnic studies offering in all California high schools. *Santa Cruz Sentinel*. Santa Cruz, CA. Retrieved from http://www.mercurynews.com/education/ci_27311449/new-bill-would-require-ethnic-studies-offering-all

Cook, P. J., & Ludwig, J. (1997). Weighing the "burden of 'acting white'": Are there race differences in attitudes toward education? *Journal of Policy Analysis and Management*, *16*(2), 256–278.

Dee, T. S. (2015). Social Identity and Achievement Gaps: Evidence from an Affirmation Intervention. *Journal of Research on Educational Effectiveness*, *8*(2), 149–168.

Deutsch, M. (1967). *The disadvantaged child: Selected papers of Martin Deutsch and associates* (Vol. 1967). New York, NY: Basic Books.

Dodge, K. A. (2011). Context matters in child and family policy. *Child Development*, *82*(1), 433–442.

Downey, D. B., & Ainsworth-Darnell, J. W. (2002). The search for oppositional culture among black students. *American Sociological Review*, *67*(1), 156–164.

Dudnick, L. (2014, December 9). SFUSD board to vote on expanding ethnic studies curriculum to all high schools. *The San Francisco Examiner*. San Francisco, CA. Retrieved from http://www.sfexaminer.com/sanfrancisco/sfusd-board-to-vote-on-expanding-ethnic-studies-curriculum-to-all-high-schools/Content?oid=2913811

Duncan-Andrade, J., & Morrell, E. (2008). The art of critical pedagogy: The promise of moving from theory to practice in urban schools. *New York: Peter Lang*.

Fordham, S., & Ogbu, J. U. (1986). Black students' school success: Coping with the "burden of 'acting white.'" *The Urban Review*, *18*(3), 176–206.

Freire, P. (2000). *Pedagogy of the oppressed*. (M. Bergman Ramos, Trans.) (30th Anniversary). New York, NY: Continuum.

Gay, G. (1988). Redesigning Relevant Curricula for Diverse Learners. *Education and Urban Society*, *2*(4), 327–40.

Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.

Gilbertson, A. (2014, November 19). Ethnic studies requirement added for LAUSD graduation. Retrieved November 26, 2014, from http://www.scpr.org/blogs/education/2014/11/19/17582/lausd-requiring-ethnic-studies-for-graduation/

Giroux, H. A., & Simon, R. (1989). Popular culture and critical pedagogy: Everyday life as a basis for curriculum knowledge. *Critical Pedagogy, the State and Cultural Struggle*, 236–252.

Hess, R. D., & Shipman, V. C. (1965). Early experience and the socialization of cognitive modes in children. *Child Development*, *36*(4), 869–886.

Horvat, E. M., & Lewis, K. S. (2003). Reassessing the "burden of Acting White": The importance of peer groups in managing academic success. *Sociology of Education*, *76*(4), 265–280.

Hurtado, S., Engberg, M. E., Ponjuan, L., & Landreman, L. (2002). Students' precollege preparation for participation in a diverse democracy. *Research in Higher Education*, *43*(2), 163–186.

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467-475.

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

Isensee, L. (2014, May 7). Why Mexican American Studies is Coming to More Texas Schools. *Houston Public Media*. Houston, TX. Retrieved from http://www.houstonpublicmedia.org/news/why-mexican-american-studies-is-coming-to-more-texas-schools/

Kisker, E. E., Lipka, J., Adams, B. L., Rickard, A., Andrew-Ihrke, D., Yanez, E. E., & Millard, A. (2012). The potential of a culturally based supplemental mathematics curriculum to improve the mathematics performance of Alaska Native and other students. *Journal for Research in Mathematics Education*, *43*(1), 75–113.

Ladson-Billings, G. (1992a). Liberatory consequences of literacy: A case of culturally relevant instruction for African American students. *Journal of Negro Education*, 378–391.

Ladson-Billings, G. (1992b). Reading between the lines and beyond the pages: A culturally relevant approach to literacy teaching. *Theory into Practice*, *31*(4), 312–320.

Ladson-Billings, G. (1994). *The dreamkeepers: Successful teachers of African American students*. San Francisco, CA: Jossey-Bass.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, *32*(3), 465–491.

Ladson-Billings, G., & Tate, W. (1995). Toward a critical race theory of education. *The Teachers College Record*, *97*(1), 47–68.

Lee, C. D. (1992). Profile of an independent Black institution: African-centered education at work. *Journal of Negro Education*, *61*(2), 160–177.

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*, 281–355.

Levin, S. (2009, June 17). Racial Tensions Boil Over at Berkeley High. *East Bay Express*. Retrieved from http://www.eastbayexpress.com/eastbay/racial-tensions-boil-over-at-berkeley-high/Content?oid=1370149

Matthews, C. E., & Smith, W. S. (1994). Native American related materials in elementary science instruction. *Journal of Research in Science Teaching*, *31*(4), 363–380.

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

Mohatt, G., & Erickson, F. (1981). Cultural differences in teaching styles in an Odawa school: A sociolinguistic approach. In H. Trueba, G. Guthrie, & K. Au (Eds.), *Culture and the bilingual classroom: Studies in classroom ethnography* (pp. 105–119). Rowley, MA: Newbery.

Noguera, P. A. (1994). Ties that bind, forces that divide: Berkeley High School and the challenge of integration. *University of San Francisco Law Review*, *29*, 719.

O'Connor, C. (1997). Dispositions toward (collective) struggle and educational resilience in the inner city: A case analysis of six African-American high school students. *American Educational Research Journal*, *34*(4), 593–629.

Office of Learning Support and Equity/Humanities, Academics and Professional Development. (2009). *Proposal to pilot a 9th grade Ethnic Studies course in the San Francisco Unified School District during the 2010-11 school year*. San Francisco, CA: San Francisco Unified School District.

Rubin, B. C., Wing, J. Y., Noguera, P. A., Fuentes, E., Liou, D., Rodriguez, A. P., & McCready, L. T. (2006). Structuring inequality at Berkeley High. *Unfinished Business: Closing the Racial Achievement Gap in Our Schools*, 29–86.

SFUSD Ethnic Studies Curriculum Collective. (2012). *Ethnic Studies Strategic Plan 2012-2013*. San Francisco Unified School District.

Sleeter, C. E. (2014). *The Academic and Social Value of Ethnic Studies: A Research Review*. Washington, D.C.: National Education Association Research Department.

Sleeter, C. E., & Bernal, D. D. (2004). Critical pedagogy, critical race theory, and antiracist education: Implications for multicultural education. *Handbook of Research on Multicultural Education*, *2*, 240–258.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797.

Stetser, M. C., & Stillwell, R. (2014). *Public High School Four-Year On-Time Graduation Rates and Event Dropout Rates: School Years 2010-11 and 2011-12.* (No. NCES 2014-391.). National Center for Education Statistics.

Tintiangco-Cubales, A., Kohli, R., Sacramento, J., Henning, N., Agarwal-Rangnath, R., & Sleeter, C. (2015). Toward an Ethnic Studies Pedagogy: Implications for K-12 Schools from the Research. *The Urban Review*, *47*(1), 104–125.

Tucker, J. (2014, November 22). Ethnic studies requirement proposed for S.F. high schools. Retrieved November 24, 2014, from http://www.sfchronicle.com/news/article/Ethnic-studies-requirement-proposed-for-S-F-high-5911492.php

Tyson, K., Darity, W., & Castellino, D. R. (2005). It's not "a black thing": Understanding the burden of acting white and other dilemmas of high achievement. *American Sociological Review*, *70*(4), 582–605.

Valenzuela, A. (1999). *Subtractive Schooling: US-Mexican Youth and the Politics of Caring.* Ithaca, NY: State University of New York Press.

Veale, L. (2015, February 26). Bay Area schools put race on the syllabus. San Francisco, CA: kalw. Retrieved from http://kalw.org/post/bay-area-schools-put-race-syllabus

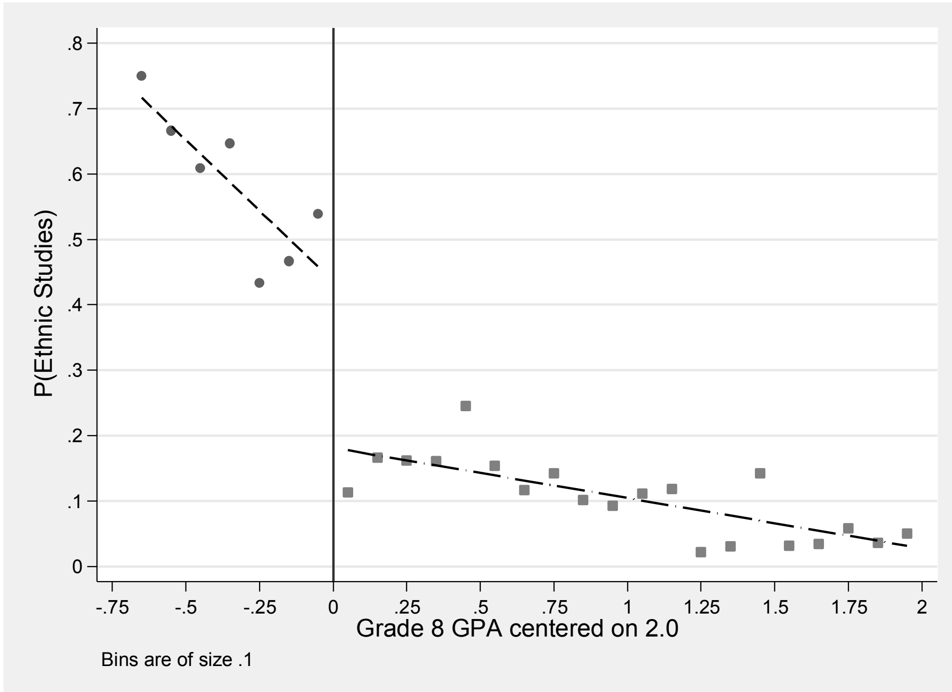
Welsh, B. C., Sullivan, C. J., & Olds, D. L. (2010). When early crime prevention goes to scale: A new look at the evidence. *Prevention Science*, *11*(2), 115–125.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*(2), 267–301.

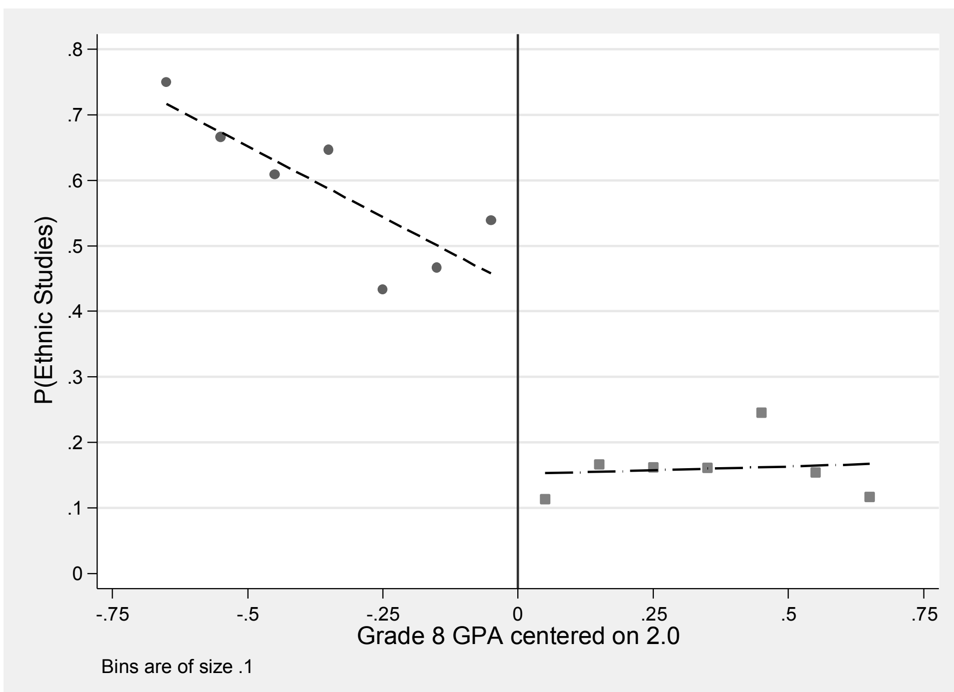
Yosso, T. J. (2002). Toward a critical race curriculum. *Equity &Excellence in Education*, *35*(2), 93–107.

Yosso, T. J. (2005). Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethnicity and Education*, *8*(1), 69–91.

Yosso, T. J., Parker, L., Solorzano, D. G., & Lynn, M. (2004). From Jim Crow to affirmative action and back again: A critical race discussion of racialized rationales and access to higher education. *Review of Research in Education*, *28*, 1–25.

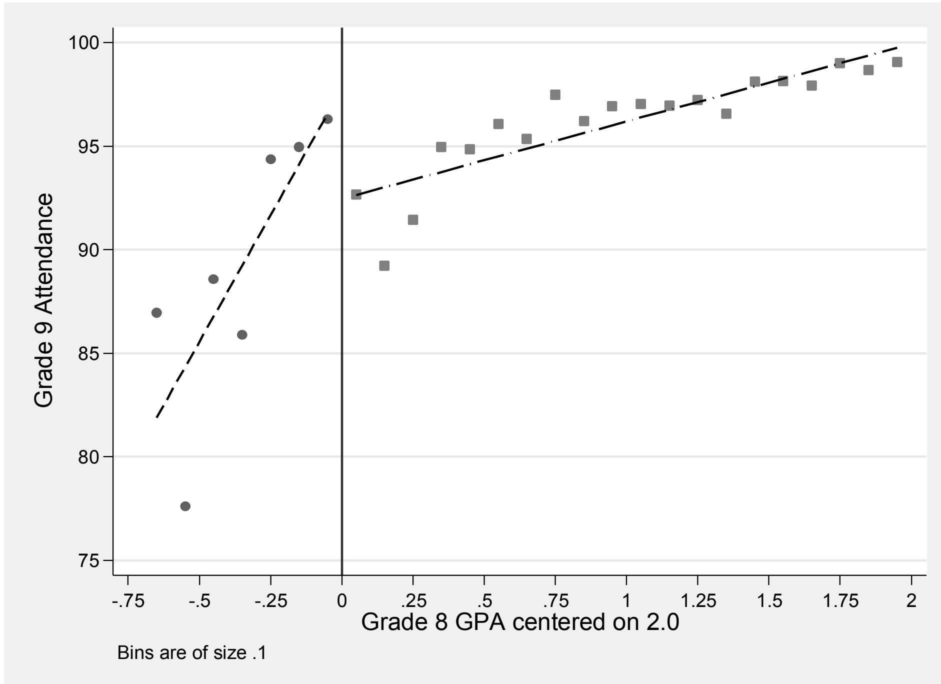


(a) Full Sample

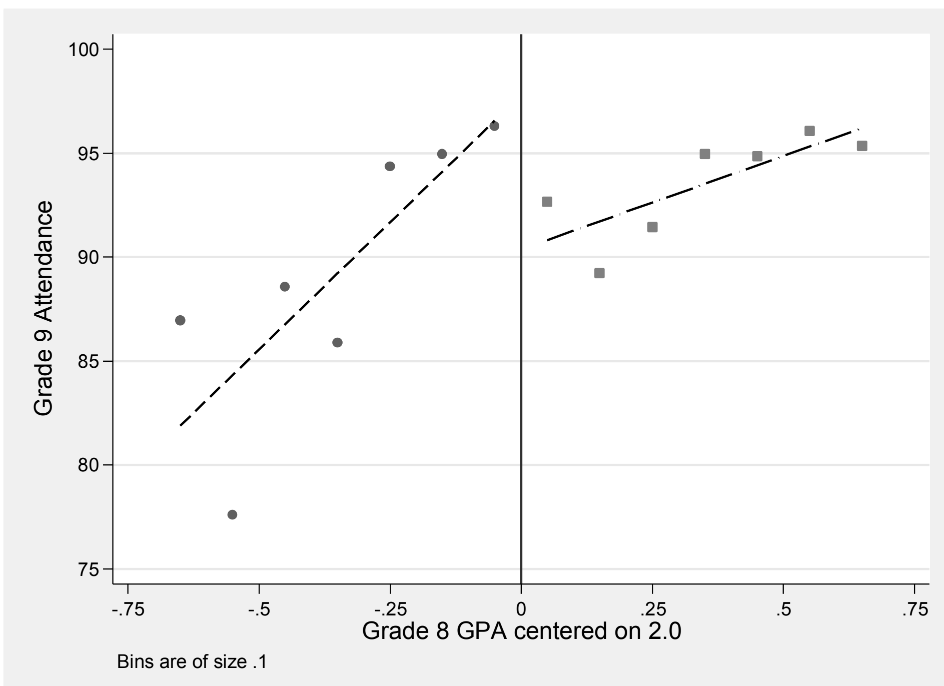


(b) ±0.7 Bandwidth

Figure 1 - Ninth-Grade Ethnic-Studies Participation by Eighth-Grade GPA

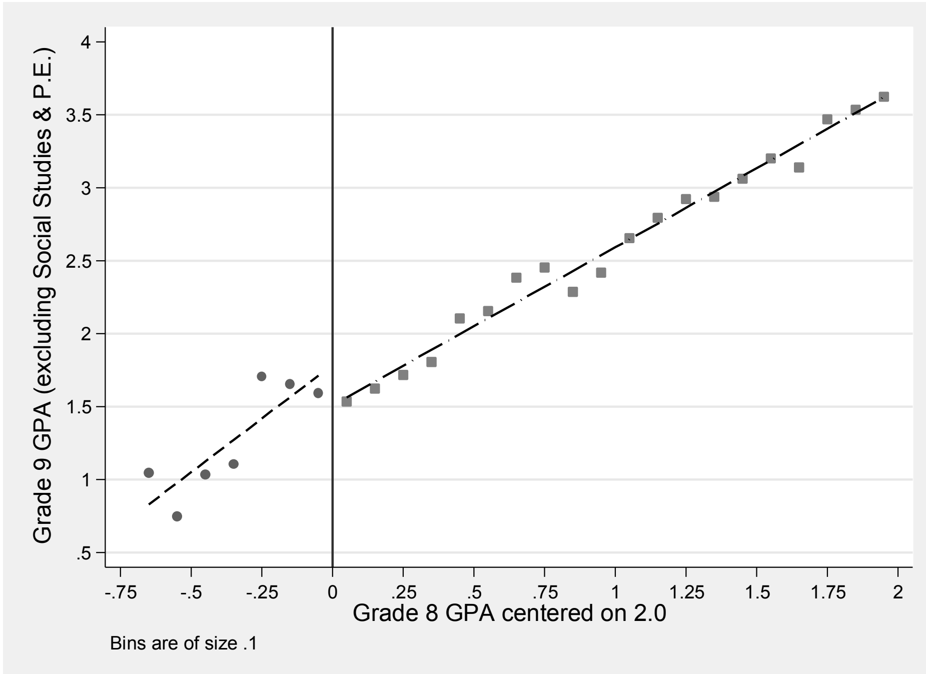


(a) Full Sample

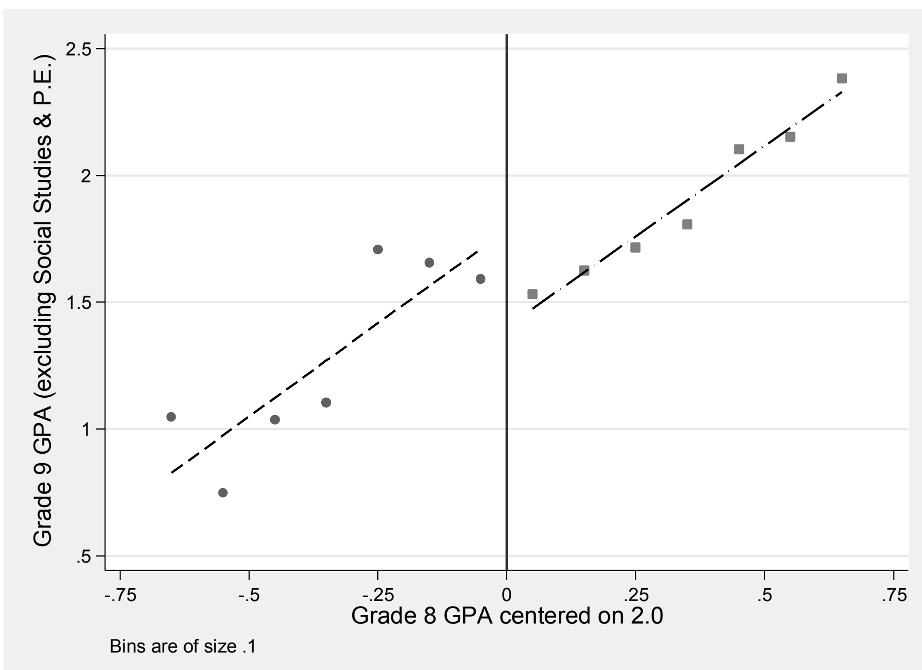


(b) ±0.7 Bandwidth

Figure 2 - Ninth-Grade Attendance by Eighth-Grade GPA

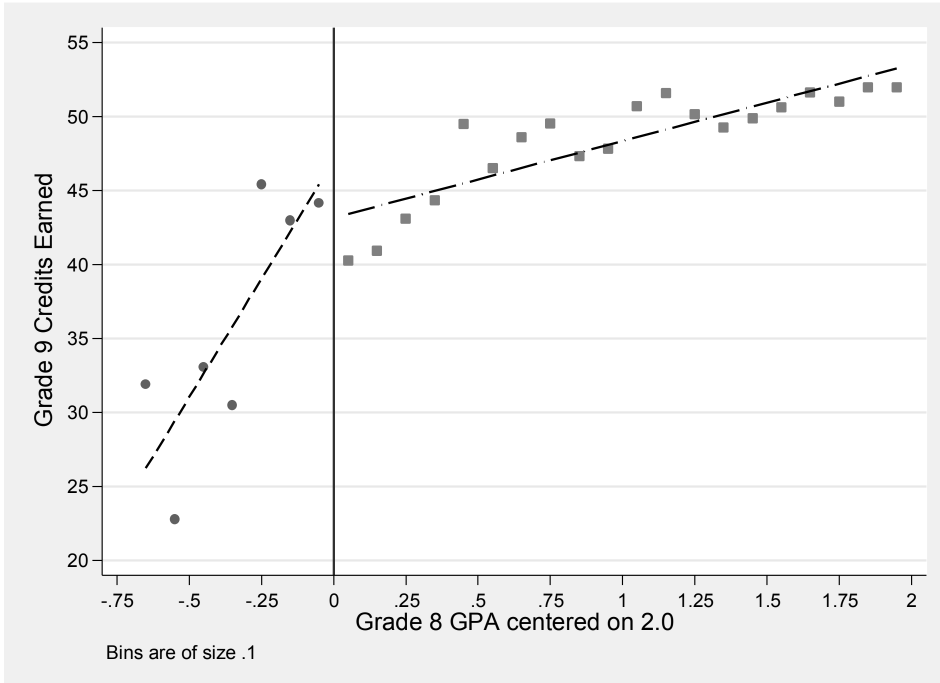


(a) Full Sample

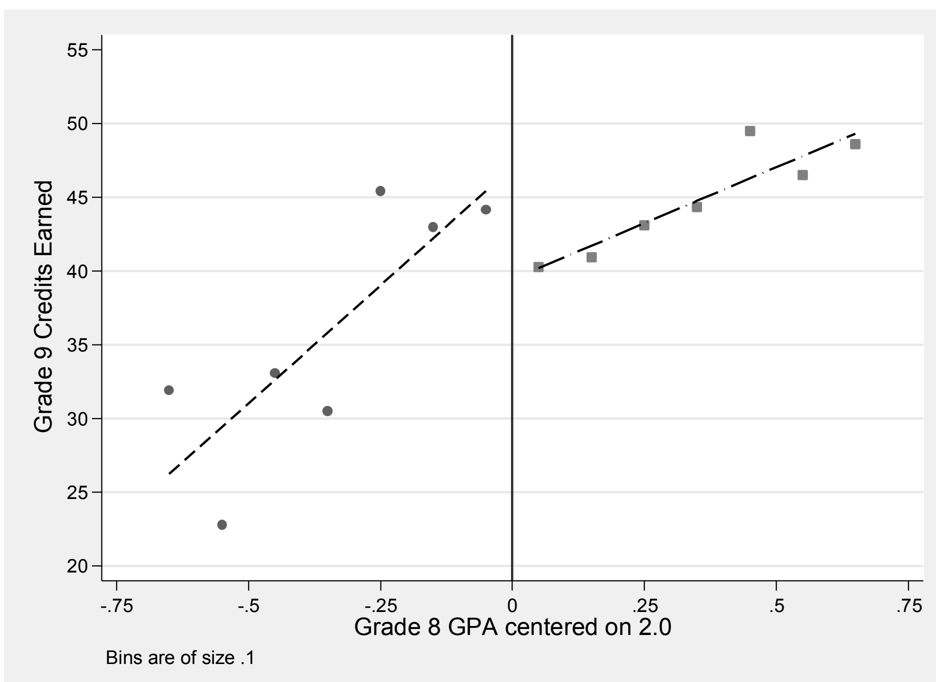


(b) ±0.7 Bandwidth

Figure 3 - Ninth-Grade GPA by Eighth-Grade GPA

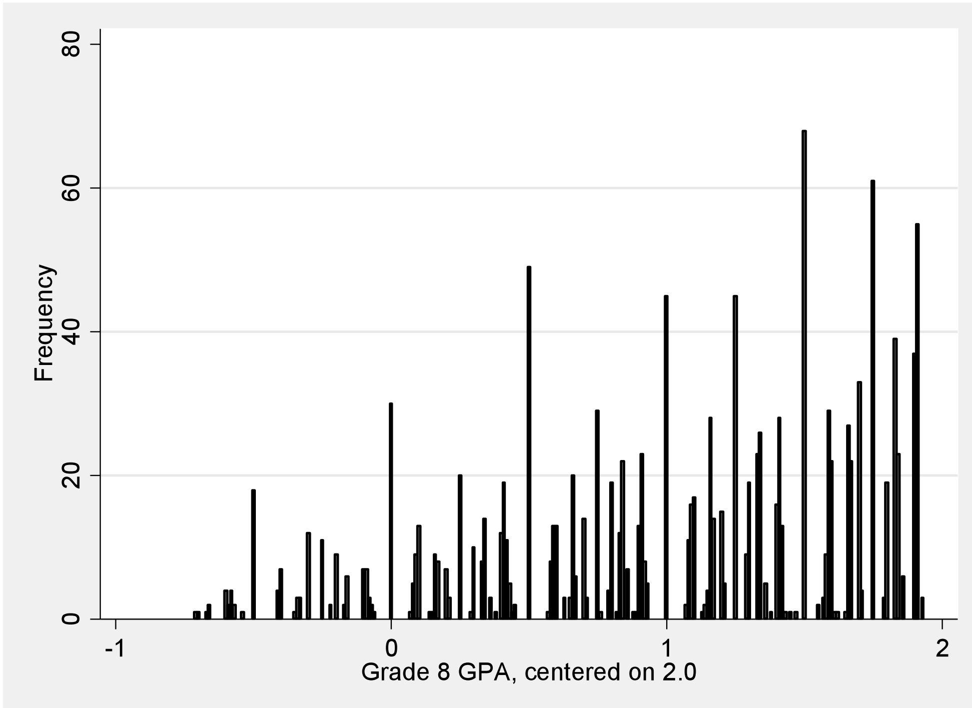


(a) Full Sample

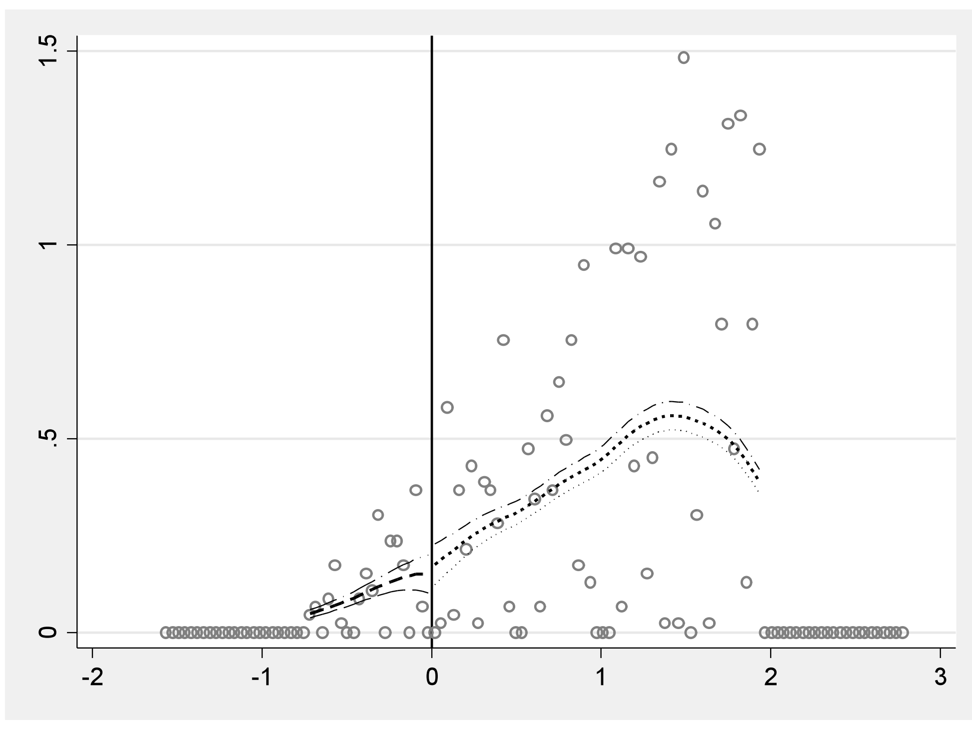


(b) ±0.7 Bandwidth

Figure 4 - Ninth-Grade Credits Earned by Eighth-Grade GPA



(a) Frequency Histogram, Full Sample



(b) Density Test (McCrary 2008), Excluding Integer and Half-Integer Heaps

Figure 5 - Distribution of Eighth-Grade GPA centered on 2.0

Table 1 - Summary Statistics

Variable	Mean	Std. Dev	Min	Max
Grade-9 Attendance	96.32	6.41	20.11	100
Grade-9 GPA (excluding P.E. & social studies)	2.65	0.97	0	4
Grade-9 Credits Earned	48.21	9.81	0	80
Grade-9 Ethnic Studies	0.13	0.33	0	1
Grade-8 GPA (excluding P.E.)	3.03	0.67	1.29	3.93
I(Grade-8 GPA < 2.0)	0.08	0.28	0	1
Female	0.42	0.49	0	1
Black	0.06	0.24	0	1
Hispanic	0.23	0.42	0	1
Asian	0.60	0.49	0	1
Grade-8 Special Education	0.12	0.33	0	1
Grade-8 Attendance	96.68	3.14	87.50	100
Grade-8 Ever Suspended	0.02	0.13	0	1
Grade-8 English Language Learner	0.18	0.39	0	1

Notes: N = 1,405 9th graders from five analysis cohorts, in three SFUSD high schools in fall 2011, 2012, 2013. Grade-8 GPA, while not centered here for comparability with Grade-9 GPA, is centered in all analyses.

Table 2 - Regression Discontinuity Estimates, Determinants of Grade 9 Ethnic-Studies Participation

Independent Variable	(1)	(2)	(3)	(4)	(5)
I(Grade-8 GPA < 2.0)	0.274***	0.268***	0.273***	0.255***	0.247**
	(0.098)	(0.100)	(0.097)	(0.097)	(0.099)
Female		0.023	0.012	0.012	0.019
		(0.017)	(0.017)	(0.017)	(0.019)
Black		0.106**	0.144***	0.144***	0.146**
		(0.051)	(0.050)	(0.051)	(0.057)
Hispanic		0.077**	0.073**	0.076**	0.092***
		(0.031)	(0.030)	(0.031)	(0.033)
Asian		-0.017	-0.021	-0.018	-0.014
		(0.023)	(0.024)	(0.024)	(0.026)
Grade-8 Special Ed			-0.170***	-0.182***	-0.183***
			(0.023)	(0.023)	(0.026)
Grade-8 Attendance			-0.007*	-0.007*	-0.006
			(0.004)	(0.004)	(0.004)
Grade-8 Ever Suspended			-0.056	-0.053	-0.053
			(0.059)	(0.063)	(0.063)
Grade-8 ELL			0.026	0.027	0.027
			(0.024)	(0.024)	(0.027)
Excluding Grade-8 GPA = 2.0	no	no	no	yes	no
Excluding Grade-8 GPA = any integer or half-integer	no	no	no	no	yes
R^2	0.178	0.195	0.222	0.228	0.218
Sample Size	1405	1405	1405	1375	1195

Notes: Student data are from five school-by-year cohorts of SFUSD 9th graders. All models condition on school-by-year fixed effects and grade-8 GPA with separate splines above and below the threshold. Grade-8 GPA is centered at 2.0. Robust standard errors are reported in parentheses.

* p<0.10; ** p<0.05; *** p<0.01

Table 3 - Regression Discontinuity Estimates, Covariate Balance

Independent Variable	Estimate
Female	0.014
	(0.090)
Black	0.090
	(0.071)
Hispanic	-0.040
	(0.098)
Asian	0.064
	(0.098)
Grade-8 Special Ed	-0.009
	(0.078)
Grade-8 Attendance	1.198*
	(0.658)
Grade-8 Ever Suspended	0.009
	(0.036)
Grade-8 ELL	-0.042
	(0.092)

Notes: Each point estimate is from a separate RD regression where the baseline covariate is the dependent variable. All models condition on school-by-year fixed effects and grade-8 GPA with separate splines above and below the threshold. N = 1,405 in all models. Robust standard errors are reported in parentheses.
* p<0.10; ** p<0.05; *** p<0.01

Table 4 - Regression Discontinuity Estimates, Grade-9 Outcomes

Independent Variable	Grade-9 Attendance			Grade-9 GPA			Grade-9 Credits Earned		
I(Grade-8 GPA < 2.0)	6.164***	6.397***	5.638***	0.413***	0.442***	0.387***	6.482***	6.723***	6.328***
	(1.607)	(1.577)	(1.449)	(0.138)	(0.136)	(0.132)	(2.272)	(2.273)	(2.201)
Female		-0.393	-0.094		0.084**	0.111***		-0.321	0.026
		(0.310)	(0.289)		(0.035)	(0.035)		(0.466)	(0.462)
Black		-3.638***	-3.346***		-0.456***	-0.479***		-2.980**	-3.414**
		(0.982)	(0.860)		(0.093)	(0.093)		(1.401)	(1.384)
Hispanic		-1.184**	-1.197**		-0.363***	-0.350***		-2.165**	-2.134**
		(0.549)	(0.496)		(0.065)	(0.064)		(0.915)	(0.912)
Asian		0.810*	-0.495		0.039	-0.023		-0.308	-0.646
		(0.422)	(0.447)		(0.055)	(0.056)		(0.749)	(0.774)
Grade-8 Special Ed			0.616			0.179***			4.239***
			(0.495)			(0.054)			(0.761)
Grade-8 Attendance			0.725***			0.046***			0.384***
			(0.089)			(0.007)			(0.106)
Grade-8 Ever Suspended			-8.434***			-0.377***			-7.727***
			(2.225)			(0.143)			(2.475)
Grade-8 ELL			-0.605			-0.101**			-0.889
			(0.456)			(0.047)			(0.654)
R^2	0.213	0.248	0.379	0.564	0.597	0.618	0.236	0.245	0.281
Sample Size	1405	1405	1405	1404	1404	1404	1404	1404	1404

Notes: Student data are from five school-by-year cohorts of SFUSD 9th graders. All models condition on school-by-year fixed effects and grade-8 GPA with separate splines above and below the threshold. Grade 8 average GPA is centered at 2.0. Robust standard errors are reported in parentheses.

* p<0.10; ** p<0.05; *** p<0.01

Table 5 - Regression Discontinuity Estimates, Grade-9 Outcomes in High Schools without Ethnic Studies

Independent Variable	Grade-9 Attendance			Grade-9 GPA			Grade-9 Credits Earned		
I(Grade-8 GPA < 2.0)	1.603	2.093	1.799	-0.111	-0.053	-0.063	-2.875	-2.403	-2.460
	(1.786)	(1.711)	(1.580)	(0.128)	(0.125)	(0.120)	(2.597)	(2.562)	(2.451)
Female		-1.245***	-0.717***		0.056**	0.099***		-0.691*	-0.132
		(0.265)	(0.241)		(0.025)	(0.024)		(0.356)	(0.347)
Black		-4.078***	-3.720***		-0.537***	-0.533***		-4.054***	-3.688***
		(0.916)	(0.829)		(0.066)	(0.064)		(1.061)	(1.042)
Hispanic		-1.784***	-1.954***		-0.437***	-0.451***		-2.698***	-2.987***
		(0.504)	(0.492)		(0.049)	(0.049)		(0.669)	(0.669)
Asian		0.827**	-0.982**		0.020	-0.094**		0.004	-1.646***
		(0.384)	(0.430)		(0.037)	(0.039)		(0.461)	(0.514)
Grade-8 Special Ed			-0.777			0.097*			0.849
			(0.592)			(0.053)			(0.671)
Grade-8 Attendance			0.800***			0.054***			0.703***
			(0.066)			(0.005)			(0.084)
Grade-8 Ever Suspended			-6.293***			-0.612***			-8.740***
			(1.559)			(0.116)			(2.656)
Grade-8 ELL			-1.053**			-0.123***			0.172
			(0.445)			(0.037)			(0.563)
R^2	0.140	0.182	0.277	0.454	0.500	0.534	0.301	0.316	0.353
Sample Size	2860	2860	2860	2851	2851	2851	2851	2851	2851

Notes: Student data are from 9th graders in SFUSD high schools that did not offer Ethnic Studies. All models condition on school-by-year fixed effects and grade-8 GPA with separate splines above and below the threshold. Grade-8 GPA is centered at 2.0. Robust standard errors are reported in parentheses.

* p<0.10; ** p<0.05; *** p<0.01

Table 6 - Regression Discontinuity Estimates, Grade-9 Outcomes in High Schools with and without Heaping

Sample	Grade-9 Attendance	Grade-9 GPA	Grade-9 Credits Earned	Sample Size
Full Sample	5.638***	0.387***	6.328***	1405
	(1.449)	(0.132)	(2.201)	
Excluding Grade-8 GPA = 2.0	5.452***	0.371***	5.490**	1374
	(1.438)	(0.133)	(2.198)	
Excluding Grade-8 GPA = any integer or half-integer	5.831***	0.344**	5.127**	1194
	(1.609)	(0.136)	(2.279)	

Notes: Student data are from five school-by-year cohorts of SFUSD 9th graders. All models condition on school-by-year fixed effects, grade-8 GPA with separate splines above and below the threshold, and other student controls. Grade-8 GPA is centered at 2.0. The sample size for GPA and credits earned is 1,404. One student attended part of the fall semester, but left before earning final grades in their courses. Robust standard errors are reported in parentheses.

* p<0.10; ** p<0.05; *** p<0.01

Table 7 - Regression Discontinuity Estimates, Grade-9 Outcomes by Bandwidth Restrictions

Bandwidth Sample	Ethnic Studies	Grade-9 Attendance	Grade-9 GPA	Grade-9 Credits Earned	Sample Size
Full Sample	0.273***	5.638***	0.387***	6.328***	1405
	(0.097)	(1.449)	(0.132)	(2.201)	
± 1.0	0.261***	6.802***	0.450***	8.884***	633
	(0.097)	(1.536)	(0.142)	(2.378)	
± 0.9	0.278***	7.152***	0.468***	9.522***	545
	(0.097)	(1.575)	(0.143)	(2.435)	
± 0.8	0.300***	7.081***	0.481***	9.726***	486
	(0.098)	(1.600)	(0.144)	(2.483)	
± 0.7	0.298***	6.777***	0.509***	9.677***	429
	(0.097)	(1.670)	(0.148)	(2.547)	
± 0.6	0.335***	7.693***	0.557***	11.079***	378
	(0.104)	(1.937)	(0.156)	(2.732)	
± 0.5	0.363***	6.567***	0.524***	10.405***	340
	(0.111)	(1.641)	(0.164)	(2.830)	

Notes: Student data are from five school-by-year cohorts of SFUSD 9th graders. All models condition on school-by-year fixed effects, grade-8 GPA with separate splines above and below the threshold, and other student controls. Grade-8 GPA is centered at 2.0. Robust standard errors are reported in parentheses.
* $p<0.10$; ** $p<0.05$; *** $p<0.01$

Table 8 - Regression Discontinuity Estimates, Placebo and Actual Thresholds

Threshold	Ethnic Studies	Grade-9 Attendance	Grade-9 GPA	Grade-9 Credits Earned
Grade-8 GPA <1.5	0.085	-4.975	0.034	-1.538
	(0.181)	(6.373)	(0.274)	(5.979)
Grade-8 GPA <1.75	0.018	-4.173	-0.255	-2.540
	(0.191)	(3.167)	(0.271)	(5.037)
Grade-8 GPA <2	**0.347*** **	**5.339** **	**0.441** **	**8.493*** **
	(0.130)	(2.270)	(0.182)	(3.274)
Grade-8 GPA <2.25	-0.086	-1.202	-0.090	-4.579**
	(0.053)	(1.399)	(0.108)	(1.788)
Grade-8 GPA <2.5	0.046	-1.502	-0.065	-0.519
	(0.050)	(0.921)	(0.099)	(1.313)
Grade-8 GPA <2.75	0.044	-0.269	0.106	-0.815
	(0.040)	(0.513)	(0.089)	(1.057)
Grade-8 GPA <3	0.018	0.729*	-0.049	-0.986
	(0.034)	(0.423)	(0.082)	(0.888)

Notes: Student data are from five school-by-year cohorts of SFUSD 9th graders. All models condition on school-by-year fixed effects,grade-8 GPA with separate splines above and below the threshold, and other student controls. Grade-8 average GPA is centered at 2.0. Robust standard errors are reported in parentheses.

* p<0.10; ** p<0.05; *** p<0.01

Table 9 - Regression Discontinuity Estimates by Student Traits

	Ethnic Studies	Grade-9 Attendance	Grade-9 GPA	Grade-9 Credits Earned	Sample Size
Full Sample	0.273***	5.638***	0.387***	6.328***	1404
	(0.097)	(1.449)	(0.132)	(2.201)	
Male	0.300***	6.432***	0.395**	8.021***	818
	(0.113)	(1.655)	(0.154)	(2.603)	
Female	0.189	2.662	0.319	2.449	586
	(0.193)	(2.809)	(0.273)	(3.632)	
Hispanic	0.307**	5.430*	0.406*	7.945**	324
	(0.146)	(2.789)	(0.211)	(3.757)	
Asian	0.445***	4.831***	0.311	3.943	844
	(0.155)	(1.579)	(0.239)	(3.795)	

Notes: Student data are from five school-by-year cohorts of SFUSD 9th graders. All models condition on school-by-year fixed effects, grade-8 GPA with separate splines above and below the threshold, and other student controls. Grade-8 GPA is centered at 2.0. Robust standard errors are reported in parentheses.

* p<0.10; ** p<0.05; *** p<0.01

Table 10 - Regression Discontinuity Estimates, Effect of Ethnic-Studies Eligibility on Subject-Specific GPA

Variable	(1)	(2)	(3)
Grade-9 GPA	0.413***	0.442***	0.387***
	(0.138)	(0.136)	(0.132)
Grade-9 GPA - Math	0.505**	0.521**	0.462**
	(0.205)	(0.211)	(0.210)
Grade-9 GPA - Science	0.459**	0.481***	0.430**
	(0.192)	(0.185)	(0.180)
Grade-9 GPA - ELA	0.217	0.253	0.195
	(0.189)	(0.178)	(0.175)
Basic RD controls	yes	yes	yes
Student Demographics	no	yes	yes
Grade-8 Traits	no	no	yes

Notes: Student data are from five school-by-year cohorts of SFUSD 9th graders. All models condition on school-by-year fixed effects, grade-8 GPA with separate splines above and below the threshold, and other student controls. Grade-8 GPA is centered at 2.0. Robust standard errors are reported in parentheses.

* p<0.10; ** p<0.05; *** p<0.01

Table 11 - Regression Discontinuity Estimates by Complier Status

Sample	Grade-9 Attendance	Grade-9 GPA	Grade-9 Credits Earned
Full Sample	5.638***	0.387***	6.328***
	(1.449)	(0.132)	(2.201)
ES = 0	4.477**	0.330*	3.991
	(2.025)	(0.181)	(3.312)
ES = 1	5.074**	0.536***	6.414*
	(2.091)	(0.202)	(3.414)

Notes: Student data are from five school-by-year cohorts of SFUSD 9th graders. All models condition on school-by-year fixed effects, grade-8 GPA with separate splines above and below the threshold, and other student controls. Grade-8 GPA is centered at 2.0. The last two rows show RD estimates for separate samples of students who did (ES=1) and did not (ES=0) take the ES course (Bertanha and Imbens 2014). Robust standard errors are reported in parentheses.

* $p<0.10$; ** $p<0.05$; *** $p<0.01$