PROJECT MUSE

Shifting College Majors in Response to Advanced Placement Exam Scores

Christopher Avery, Oded Gurantz, Michael Hurwitz, Jonathan Smith

Journal of Human Resources, Volume 53, Number 4, Fall 2018, pp. 918-956 (Article)

Published by University of Wisconsin Press

For additional information about this article
https://muse.jhu.edu/article/706372

Access provided at 18 Feb 2020 20:17 GMT from University of Missouri - Columbia

Shifting College Majors in Response to Advanced Placement Exam Scores

Christopher Avery
Oded Gurantz
Michael Hurwitz
Jonathan Smith

ABSTRACT

Do signals of high aptitude shape the course of collegiate study? We apply a regression discontinuity design to understand how college major choice is impacted by receiving a higher Advanced Placement (AP) integer score, despite similar exam performance, compared to students who received a lower integer score. Attaining higher scores increases the probability that a student majors in that exam subject by approximately 5 percent (0.64 percentage points), with some individual exams demonstrating increases as high as 30 percent. A substantial portion of the overall effect is driven by behavioral responses to the positive signal of receiving a higher score.

Christopher Avery is Roy E. Larsen Professor of Public Policy and Management at the Harvard Kennedy School (christopher_avery@hks.harvard.edu). Oded Gurantz is an assistant professor at the Truman School of Public Affairs at the University of Missouri (gurantzo@missouri.edu). Michael Hurwitz is senior director at the College Board (mhurwitz@collegeboard.org). Jonathan Smith is an assistant professor of economics at Georgia State University (jsmith500@gsu.edu). *The authors thank Jessica Howell, Jack Buckley, Trevor Packer, Maureen Reyes, and seminar/conference participants at the Society of Labor Economists (SOLE), the Association for Education Finance and Policy (AEFP), the Society for Research on Educational Effectiveness, (SREE), the Association for Public Policy Analysis and Management (APPAM), and the Western Economic Association International (WEAI) conferences, Michigan State University, University of California, Santa Cruz, University of Washington, Stanford University, and the University of Pittsburgh for useful feedback. This research reflects the views of the authors and not their corresponding institutions. The data used for the project are proprietary and cannot be freely posted online. Researchers can gain access to these data for replication purposes by contacting the College Board (The College Board, 250 Vesey Street, New York, NY 10281; 212-713-8088) or through the direct online portal at http://research.collegeboard.org/data/request. The authors are willing to assist (mhurwitz@collegeboard.org).*
[Submitted October 2016; accepted April 2017]; doi:10.3368/jhr.53.4.1016-8293R
JEL Classification: I2, I23, and I26
ISSN 0022-166X E-ISSN 1548-8004 © 2018 by the Board of Regents of the University of Wisconsin System
Supplementary materials are freely available online at: http://uwpress.wisc.edu/journals/journals/jhr-supplementary.html

I. Introduction

A student's choice of college major may have long-lasting implications, including on future earnings. The average difference in lifetime earnings between the top-paying and lowest-paying majors is estimated to be several million dollars (Carnevale, Cheah, and Hanson 2015), and there is growing evidence that at least some portion of the connection between college major and wages is causal (Hastings, Neilson, and Zimmerman 2013; Kirkebøen, Leuven, and Mogstad, 2016). Despite these seemingly strong labor market incentives, there remains a mismatch between demand and supply of workers in some relatively lucrative fields. For example, a 2012 Federal Government report estimates a ten-year shortfall of 1 million college graduates with STEM (science, technology, engineering, and mathematics) majors. As these findings suggest, some students may not be choosing their college majors optimally, either because they lack adequate information on the relative benefits and challenges of certain majors, or because they enter college with inadequate academic preparation for a preferred major (Betts 1996; Oreopoulos and Dunn 2013; Stinebrickner and Stinebrickner 2012; Wiswall and Zafar 2015a, 2015b).

In addition to the earnings potential associated with each field of study,[1] the previous literature emphasizes the importance of heterogeneous tastes and predilections on a student's choice of major. Morgan, Gelbgiser, and Weeden (2013) find in their analysis of Educational Longitudinal Survey (ELS) data that specific occupational plans reported by students prior to entering college yield much sharper predictions of their college majors than test scores and other observable performance data. Similarly, Altonji, Arcidiacono, and Maurel (2015) suggest that a combination of major-specific abilities and individual preferences drive the choice of major for most students, while Wiswall and Zafar (2015a) estimate that 80% of the variation in major-specific tastes remains unexplained by observable characteristics.

Although this literature suggests that each student's chosen field of study can be highly personal and driven by factors in place before entering college, there is also a small body of evidence that the choice of major is subject to external factors that can be shaped by policy. For instance, peers (Ost 2010), early exposure to a subject via required coursework (Fricke, Grogger, and Steinmayr 2015), and the final grade achieved by a student in an introductory course (Goldin 2015) can also have strong influence on a student's subsequent course of study. Several recent papers assess the effects of explicit financial incentives in directing students to particular fields of study. Denning and Turley (2017) find that the "SMART" Program, which provides U.S. Department of Defense scholarships to college juniors and seniors pursuing STEM majors, significantly increased the probability of completing college with a major in those fields, although Evans (2017) finds no significant effects in Ohio. Similarly, Castleman, Long, and Mabel (2018) find that the Florida State Access Grant (FSAG) program significantly increased the probability of completing college with a STEM major even though FSAG funding was not tied in any way to the choice of major.[2]

1. See for example, (Arcidiacono, Hotz, and Kang 2012; Beffy, Fougère, and Maurel 2011; Long, Goldhaber, and Huntington-Klein 2015; Shu 2013; Stinebrickner and Stinebrickner 2014; Wiswall and Zafar 2015a).
2. In related work, Stange (2015) finds that differential tuition policy can alter the demand for a particular degree.

In this paper we focus on the role of Advanced Placement (AP) exam scores and their signals, which reflect a nationally recognized college-level curriculum taken by hundreds of thousands of high school students each year, in encouraging students to choose a college major in a subject of interest. In particular, to isolate the causal impact of different AP exam scores among students with similar mastery of the content and skills of an AP course, we compare students with very similar performance on the AP exam, but who receive different AP scores by falling on either side of the cutoff score that separates an AP integer score of 5 from an AP integer score of 4, the cutoff score that separates an AP 4 from an AP 3, and so on. We investigate two channels by which a higher reported AP exam score, among students with otherwise comparable mastery of the course content and skills, can increase the probability that the student completes a college major in a field of study connected to that AP course. First, a higher AP score can coincide with an increase in college credits (and/or preferential course placement), both towards graduation requirements and towards completion of a particular major at a given college. Second, students may have a behavioral response to a higher AP score, such that they perceive themselves to have more ability in the field, or use the high score as a guidepost for choosing initial courses or major.

As in our previous and related study, which finds a causal effect of AP exam scores on degree completion (Smith, Hurwitz, and Avery 2017), we apply a regression discontinuity design to AP exam scores from millions of students who graduated high school between 2004 and 2009. Students and colleges only observe an integer exam score between 1 and 5, but we rely on the underlying continuous scores that map to the integer score. These data allow us to compare the majors of students who just barely attain a 3, for example, relative to those just shy of the threshold who instead attain a 2. Isolating the impact of attaining a higher score by comparing identical students distinguishes our paper from previous work, which establishes the strong predictive component of AP scores and major (Mattern, Shaw, and Ewing 2011). To be clear, our analyses compare two essentially identical students who have both elected to take an AP, but we are unable to measure the effect of exposure to or quality of the AP curriculum on major choice. Participating in AP courses may have strong and independent causal impacts on student major (and other outcomes), but, in these analyses, we are not able to separate this effect from other unobserved factors that might impact both AP exam performance and student major.

Similar to previous work in the area, we show a strong positive relationship between AP integer scores and choice of college major. For example, students in our sample who attain a 5 on an AP exam—the highest possible score—are 5.7 percentage points (64 percent) more likely to major in the same subject as the AP exam than students who attain a 4 on the exam. However, when comparing students whose raw scores barely placed them into the 5 category compared to those who just missed a score of 5, we find a 0.64 percentage point (5 percent) increase in majoring in the same subject as the AP exam. This implies that approximately 11 percent of the increase in the probability of majoring in the same subject as the AP exam can be explained not by differences in students but rather, the direct impact of receiving a higher integer score. We also see causal effects that are smaller in magnitude by attaining a 3 over a 2 and 4 over a 3. AP and its scoring impacts millions of students each year across the entire nation and is delivered prior to the beginning of college, which is unique among causal studies on major choice. Further, for students with nearly identical performance on the AP Exam (adjacent scores on the continuous scale), we find evidence that the effect of an increase in AP score on the choice of college major is primarily driven by the behavioral effect of the positive signal communicated by the higher integer score, as shifts occur even when the higher integer score does not coincide with a jump in potential college credits.

Finally, we document several other results about how AP scores influence a student's choice of college major. First, our estimates do not detect any strong heterogeneous effects, suggesting that the causal impacts on major choice hold across students differing on gender and underrepresented minority status. Second, the strong impact of a score of 5 is attentuated when students also receive additional high AP scores, implying that the power of an additional signal depends on how many other positive signals the student has received. Finally, although students who attain higher AP scores on STEM exams are, on average, considerably more likely to major in STEM (for example, students scoring a 5 on a STEM AP exam are 42 percent more likely than students scoring a 4 on a STEM AP exam), the impact of a higher AP integer score among students with otherwise comparable AP exam performance shifts students across STEM disciplines, which we will discuss later in the paper. In other words, factors, many of which are unobservable, such as quality of AP instruction, students' mastery of the required content and skills, or students' interest and motivation in a subject, likely explain the strong positive relationship between AP integer scores and the student's likelihood of majoring in that AP discipline, rather than the unique signaling effect of a higher integer score to a student who has content and skill mastery similar to a student who received a lower integer score.

This paper is organized as follows. Section II describes the Advanced Placement program, scoring, and literature. Sections III and IV describe our data and methodology, respectively. Section V presents our main findings on the response to relatively higher AP scores, along with the exploration of underlying mechanisms, including credit policies and behavioral responses to positive signals. Section VI investigates some of the broader impacts of our findings, including heterogeneous effects, the impact of multiple signals, and changes in STEM degree production. Section VII concludes.

II. AP Background and Literature Review

A. AP Background

The history of the Advanced Placement Program is rooted in philosophies that college-level academic opportunities should be extended to high-achieving high school students and that demonstration of proficiency in such coursework should exempt college students from retaking courses (see Smith, Hurwitz, and Avery 2017 for more details). Collaborating with high school teachers and college professors, the AP program develops curricula that are reflective of the content typically taught in introductory-level college courses, and exams are constructed to certify whether students have mastered the content and skills required for course exemption. Since its introduction in the 1950s, the AP program has extended its reach beyond college preparatory schools and well-funded public schools, and currently, more than 9 out of 10 public school students in the

United States have access to at least one AP exam at their schools (Theokas and Saaris 2013).[3] In 2015, high school students took nearly 4.5 million AP exams in 36 subjects. Exams take place over a two-week period in May, with only one administration per subject per year, and scores are released several months later.[4] The exact number of AP exams has varied over time, as some exams were retired due to low participation rates, and new exams were introduced as a result of high student demand. This paper only considers the 19 most popular subject exams, with at least 100,000 exam takers between 2004 and 2009 (see Online Appendix Table 1 for details on all 34 exams).

AP scores are reported to students and colleges on a 1 through 5 scale, where 1 translates into "no recommendation" and 5 translates into "extremely well-qualified." The integer scores are based on students' raw scores, which reflect performance on multiple choice and free-response sections. Because the AP exams are criterion based, cutoff scores are established based on earning a predetermined number of points that predict college performance at varying levels and not on relative performance. The exams are designed so students earning a score of 3 on one test administration have an identical mastery of material as students earning a 3 on a separate administration.[5]

In order to receive credit, course exemption, and placement, students must submit AP scores to the institutions at which they enroll. Variation exists in how AP exam scores are treated, both across postsecondary institutions and across exams within postsecondary institutions. Most students enrolling at four-year institutions attend colleges that award credits toward graduation if students meet certain threshold minima—generally a 3 or 4 on the standard 1–5 scale. Along with receipt of college credit, the student is generally eligible to enroll immediately in the sequent course. Colleges independently decide how many credits students receive for meeting AP thresholds, the sequent courses for which they are eligible, and whether scores exceeding the credit-granting thresholds are appropriate for the awarding of additional credits and course exemptions.

B. AP Literature Review

Our paper contributes to a small, but expanding body of literature that separates out the predictive effects of AP participation and performance from the causal effects of receiving higher AP integer scores. A substantial prior literature documents a positive relationship between early college credit and choice of major (Dodd et al. 2002; Keng and Dodd 2008; Murphy and Dodd 2009; Tai et al. 2010).[6] More recently, Mattern,

3. The College Board official statistics are slightly lower at around 60 percent (see http://apcentral.collegeboard.com/apc/public/program/index.html, accessed February 26, 2018).
4. Approximately 0.3 percent of students retake an AP exam.
5. Continuous raw scores range from 0 to 180 points, though there is considerable variation in the scoring range and maximum across exams.
6. There also exists a series of studies that demonstrate a strong positive correlation between AP participation, AP exam scores, and subsequent academic performance across a range of measures, including college attendance (Chajewski, Mattern, and Shaw 2011) and success in subject performance (Patterson and Ewing 2013), overall performance (Mattern, Marini, and Shaw 2013), and college completion (Dougherty, Mellor, and Jian 2006; Hargrove, Godin, and Dodd 2008; Mattern, Marini, and Shaw 2013). There is a similar line of research on dual enrollment. For example, see Karp et al. (2007).

Shaw, and Ewing (2011) find that students who take a particular AP exam are much more likely to major in that subject: students who take AP Computer Science are 4.5 times more likely to major in computer science than students who did not take the AP course. These large estimates rest on a selection on observables identification strategy, which relies on logistic regressions that include demographic and academic (for example, SAT scores, self-reported GPA) characteristics in an attempt to control for observed differences between students with higher and lower AP exam scores.

Sources of randomization in the context of AP research are hard to come by, and many of the most compelling studies examining the long- and short-term consequences of AP course and exam taking have relied on models that control for observed covariates to deal with potential confounders between high- and low-performing students (Evans 2017; Long, Conger, and Iatarola 2012; Murphy and Dodd 2009).[7] Two notable exceptions are our own study linking AP scores to college graduation outcomes (Smith, Hurwitz, and Avery 2017) and Jackson (2010), who finds that the introduction of a program that paid teachers and students for success on AP examinations increased SAT/ACT scores and college matriculation. Despite the convincing case for causality, Jackson (2010) is unable to generalize about the relative contributions of improved teaching, increased exposure to rigor and the direct effects of the fact that some students may have earned higher AP scores as a result of this incentive program. In what follows, we isolate the effect of higher AP scores and demonstrate its effects on choice of major.

III. Data and Descriptive Statistics

A. College Board Data

This paper uses student-level data from the 2004–2009 graduating high school cohorts collected from two main sources, College Board (CB) data on AP examinees and National Student Clearinghouse (NSC) data. College Board maintains a database of all students who take at least one AP exam. This database contains not only the students' AP exam scores on the 1–5 integer scale, but their underlying continuous scores on most exams taken between 2004 and 2009. From these two pieces of information, we identify the exact continuous scores that sharply form the boundaries of the scaled scores.[8] In addition to student performance on each AP exam, the CB data also contain a host of student demographic information, such as a student's gender, race/ethnicity, and parental income.[9] We also observe student SAT scores, if they take the exam. We frequently divide our analyses into separate results for STEM and non-STEM AP exams, which are listed in Online Appendix Table 1. The AP exams used in this paper that

7. There are currently some randomized AP evaluations underway, which will be very informative, but they are limited in their scope of exams and populations (Long, Conger, and McGhee 2014).

8. Data on raw scores are available only for exams taken during the 2003–2004 school year or later. Therefore, some AP test takers, particularly in the 2004 and 2005 cohorts, will not have raw scores that can be mapped to their scaled scores taken in the sophomore or junior year of high school. The few exams without an accompanying raw score are removed from our analyses.

9. Parental income is collected on the SAT registration forms, and so some AP test takers who did not participate in the SAT will have missing demographic information. Even among SAT participants, some students fail to respond to these questions.

are considered STEM include Biology, Calculus, Chemistry, Environmental Sciences, Physics, and Statistics.

B. National Student Clearinghouse, CIP Codes, and IPEDS

College Board data are then merged with the NSC data. As of 2015, over 3,600 postsecondary institutions participate in NSC, which collects postsecondary enrollment information on more than 98 percent of students enrolled in public and private colleges within the United States.[10] In this study, we track a student's postsecondary trajectory including enrollment and degree completion. We observe students college trajectories for six years after they graduate high school for the 2004–2007 cohorts, five years for the 2008 cohort, and four years for the 2009 cohort.

The majors in the NSC data are provided only for graduating students, and we focus exclusively on majors associated with a bachelor's degree. The NSC provides full six-digit Classification of Instructional Program (CIP) code information,[11] which we simplify by focusing on the first two digits.[12] Two-digit CIP codes translates into general fields such as biology, history, or English.

In order to assess whether college majors are impacted by different AP scores for similar exam performance, we match each AP subject to the closest two-digit CIP code (see Online Appendix Table 2). In some cases the match is fairly exact; for example, students taking AP Biology are linked to the CIP code denoting Biological Sciences. In other cases we are required to group AP exams, as both Chemistry and Physics are most closely linked to the two digit CIP code for Physical Sciences.[13] In addition, we consider whether AP exams alter students' major in the broader field of STEM majors. We select all CIP codes where the first two digits correspond to our STEM AP exams, namely 11, 14, 15, 26, 27, and 40. Although we do not capture all STEM majors with this approach, we do capture most STEM degrees at four-year universities.[14]

10. Due to data privacy laws and potential complications with student matching, the actual NSC coverage may be a bit lower than the 98 percent rate (Dynarski, Hemelt, and Hyman 2015).
11. The CIP codes are a taxonomic scheme created by the U.S. Department of Education to ensure a uniform system of tracking across colleges.
12. CIP codes are not provided for the 2004 cohort and approximately one-third of institutions in other cohorts but are instead in text form that we unify into CIP codes. Since there is the chance for classification error, we test the sensitivity of the results by only using the students with a CIP code. Results hold and are presented in the Online Appendix.
13. The only deviation from this approach that we adopt is a grouping of AP Calculus and Statistics with majors in either math/statistics or engineering, primarily because relatively few students major in math, and engineering is far more prevalent among test takers in these subjects. Also, 4 percent of students double major. If one of the two majors is related to the AP exam, students are counted as majoring in that subject. Results are not sensitive to excluding double majors (see Online Appendix Table 5).
14. This is most problematic for students without CIP codes (but a textual description of major), of which we exclude in robustness tests. Other commonly used STEM classification systems typically include a relatively small number of CIP codes in the two-digit fields of 1 (Animal and Plant Sciences), 3 (Natural Resource Conservation), 29 (Military Technologies), 30 (Multi-disciplinary Studies), 41 (Science Technologies), and 51 (Pharmaceutical Sciences), along with a small number of other specific majors. As most majors in these broad two-digit disciplines are not STEM-related, their inclusion was deemed incorrect. In alternate analyses not presented here, we show that our STEM results in Table 9 are robust to using only schools that report six-digit STEM codes and using alternate STEM classifications, such as U.S. Immigration and Enforcement lists of STEM programs that qualify foreigners for expedited work visas.

Finally, we append to our data several variables from the Integrated Postsecondary Education Data System (IPEDS). These include the average standardized test scores (ACT and SAT) of incoming students and whether the college is public or private.[15]

C. AP Credit Policies

We use AP credit policies from two sources: the Annual Survey of Colleges (ASC) and data collected by the authors from college websites. Administered annually by the College Board to nearly 4,000 colleges, the 2004 survey included information on the minimum credit-granting scores by AP subject (after 2005, College Board did not include information on the minimum credit granting scores on the survey). We supplement these data by constructing an enhanced "policy sample." To accomplish this, we collected the more nuanced AP credit and placement data directly from the websites of the 500 largest four-year institutions in the country, as measured by full-time-equivalent students. This inclusion rule captures a wide swath of postsecondary institutions—both selective and nonselective colleges, along with a representative mix of public and private colleges—and represents approximately 82% of students who take an AP exam. We create a binary "AP Credit" variable for each combination of AP exam and threshold at each college. We code the AP Credit variable as a "1" for each exam–college threshold combination if a college provides any beneficial advantage at that threshold, including credit towards graduation, credit towards major, or placement into any advanced course.[16] For example, some colleges provide four units of credit for a scaled score of "at least a 3." In this example, the AP credit variable would be coded as "1" for an AP scaled score of 3 and "0" for any other AP scaled score (2, 4, or 5). As another example, a college may provide four units of credit for a score of 3 and eight units of credit for a score of 4 on a given AP exam. In this example, we would code the AP Credit variable as "1" for a scaled score of 3, "1" for a scaled score of 4, and a "0" for a scaled score of 5. Online Appendix Table 1 provides summary statistics of the credit policies across these 500 colleges.

We highlight several limitations in the use of the AP credit data that we collected for these 500 colleges in the summer and fall of 2015. First, these policies reflect current practices at these colleges, whereas our data apply to students who graduated from high school between 2004 and 2009. Even so, we find that at least 70% of colleges have identical minimum credit-granting policies from 2004 (derived from ASC data) and in 2015 (from our manual data collection), so we conduct sensitivity analyses on the subset of colleges and thresholds with identical minimum credit-granting thresholds for AP credit in 2004 and 2015. (See Section V.C for results.) Second, the coding of our binary AP credit variables does not account for a variety of nuances in policies across colleges. For example, some colleges may place caps on AP credits used towards college graduation, and/or provide conditional credit for scores on certain AP exams based on a student's choice of major. For this reason, we intentionally adopt a conservative approach through expansive coding rules in the creation of the AP Credit variables, ensuring

15. To estimate average composite SAT scores, we add the 25th and 75th percentiles of the Math and Critical Reading sections, as reported by IPEDS, and divide by two. For colleges that only report ACT scores to IPEDS, we use an SAT conversion table found at http://research.collegeboard.org/sites/default/files/publications/2012/7/researchnote-2009-40-act-sat-concordance-tables.pdf (accessed Jan. 31, 2018).

16. Note that we use the word "credit," but in some instances it is only placement with no credit.

that imprecision in the coding of these variables will induce downward bias since some fraction of students will not be receiving credit despite being coded as having done so in a relevant AP credit variable. In the case of shifts in major that are driven by a behavioral response to higher AP scores, these caveats about the coding of the AP credit variables should have no impact on our estimates.

D. Descriptive Statistics

We present summary statistics of students in our analytic sample in Table 1. We find that approximately 69 percent of the sample is white, 43 percent are male, and 50 percent reported having a parent who attended at least some college. On average, students earned a 1207 on the SAT, took almost three AP exams, and scored an average of 2.7 on the exams. We further disaggregate our summary statistics to describe students close to each of the four integer score thresholds (that is, from the 1/2 threshold up to the 4/5 threshold). As regression discontinuity designs assess causal impacts that are local to students near the threshold, understanding characteristics of marginal students helps us assess to whom we might be able to extrapolate these results. Students who perform better on AP exams are more likely to be white, Asian, male, have more educated parents, and come from families with higher income. As might be expected, AP-taking students have high SAT scores; the average composite scores of 1,147 and 1,324 at the 1/2 and 4/5 thresholds correspond to roughly the 65th and 90th percentile of the national SAT distribution, respectively. Roughly one-third of students at the 4/5 threshold attend a postsecondary institution that belongs to the Barron's Most Competitive ranking, and close to 90% will earn a bachelor's degree within six years.

Table 2 lists the probability that a student with a given AP score on an AP exam chooses the college major most closely associated with that subject, and then more generally in any STEM field.[17] Consistent with previous research, there is a systematic increase in the probability of choosing the most related college major for every field. Using AP Biology as one example, the probability of majoring in biology monotonically increases with each integer score, such that students who receive a 5 are nearly five times more likely to major in the subject as students who receive a 1. Similar patterns exist across all the exams, although the exact magnitude varies, demonstrating the strong predictive power of AP scores in major choice. The second set of columns show similar patterns for the likelihood of majoring in any STEM field, regardless of whether it is directly tied to the particular AP subject. As the interests, abilities, and supportive structures of students with a higher scaled score on a given AP exam are (presumably) systematically different than those of students with a lower scaled score on that same test, the values in Table 2 can be viewed as unrealistically large upper bounds on the causal effect of an increase in scaled score on the choice of college major. In general, the relationships between AP integer scores and the probability of majoring in STEM are stronger for STEM than for the non-STEM exams. We still observe a strong correlation between integer scores in English Language or World History and majoring in a STEM field, though we would not assume that the curricular content in these courses has any particular impact on scientific knowledge.

17. These probabilities look similar when conditioning on graduates. Our primary analyses do not condition on graduating, so we only present these statistics on the entire sample of graduates and nongraduates.

Table 1
Summary Statistics

		Threshold			
	Overall	1/2	2/3	3/4	4/5
Student demographics					
White	69.3%	60.9%	69.5%	72.9%	72.4%
Asian	10.7%	12.4%	12.3%	13.2%	15.2%
Black	6.8%	9.8%	5.5%	3.3%	2.1%
Latino/Hispanic	9.2%	12.8%	8.7%	6.8%	6.4%
Male	43.2%	40.4%	42.4%	45.6%	49.2%
Parental education: Some college	49.6%	45.2%	52.2%	58.8%	63.8%
Income <$50,000	13.0%	17.2%	12.9%	10.3%	8.9%
Income $50,000–$100,000	19.1%	19.9%	20.4%	19.7%	18.5%
Income >$100,000	18.0%	16.4%	19.4%	21.9%	23.8%
Exam scores					
SAT	1207	1147	1203	1265	1324
Number of AP exams taken	2.8	3.3	3.6	4.1	4.7
Number of STEM AP exams taken	1.0	1.1	1.2	1.4	1.7
Average AP exam score	2.7	1.7	2.4	3.2	3.8
College outcomes					
Attended Barron's most competitive	13.4%	8.6%	12.3%	20.4%	32.2%
Graduated within six years	77.8%	74.0%	80.2%	84.5%	87.4%
Majored in STEM field	14.3%	10.9%	14.3%	18.5%	24.0%
N	3,148,598	317,481	519,340	553,804	385,777

Notes: Summary statistics are calculated using deduplicated, individual-level data. The first column includes the full sample of all students who took one of the 19 most taken AP exams. All other columns include students within five points below the relevant threshold on at least one exam. Some students do not provide demographics.

Taking the previous table one step further, Table 3 reports the distribution of college majors for students with scaled score of 3 or higher on each of 19 most popular AP exams, indicating a conspicuous correlation between AP exam performance and choice of college major.[18] Typically, the most popular college major for students who score 3 or higher on a particular AP exam is the major most closely associated with that exam. For example, students with scaled score of 3 or higher in AP Biology were more than twice as likely to major in biology (18.9%) than in any of the other tabulated subjects.

The two previous tables demonstrate the predictive power of AP scores in determining major for all students. The next section focuses on students just around the integer thresholds so as to compare students who are identical across all dimensions and to estimate the impact of receiving higher AP scores, independent of differences in student attributes.

18. Students with scaled scores of 3 or higher on more than one AP exam are counted multiple times in this table.

Table 2
Probability of Majoring in Core CIP Code or STEM by AP Exam Subject

	Major in Core Subject					Major in STEM				
	AP Score					AP Score				
AP Exam	1	2	3	4	5	1	2	3	4	5
Biology	5.4%	9.5%	13.7%	18.5%	24.6%	8.0%	14.4%	20.9%	28.7%	40.9%
Calculus AB	4.4%	7.1%	9.2%	12.0%	18.4%	11.8%	17.5%	21.3%	26.0%	35.4%
Calculus BC	8.7%	11.7%	14.8%	18.3%	26.4%	20.4%	25.9%	30.8%	35.8%	46.4%
Chemistry	1.9%	3.5%	4.7%	6.5%	9.6%	14.7%	24.6%	32.1%	40.2%	51.7%
English Language & Comp.	1.0%	1.9%	3.4%	5.5%	8.1%	5.6%	10.4%	15.0%	18.3%	20.6%
English Literature & Comp.	0.9%	2.0%	3.8%	6.4%	9.8%	5.6%	10.2%	14.5%	17.4%	18.5%
Environmental Science	2.2%	3.9%	5.5%	7.6%	11.3%	4.6%	7.6%	10.7%	15.7%	26.1%
European History	1.2%	1.9%	3.7%	6.5%	9.8%	7.1%	10.9%	14.4%	18.0%	20.2%
French Language and Culture	2.0%	3.9%	6.4%	9.0%	10.9%	11.4%	15.0%	17.7%	18.3%	18.7%
Macroeconomics	7.0%	9.4%	11.1%	13.0%	16.8%	8.4%	13.6%	17.6%	23.6%	32.4%
Microeconomics	7.9%	10.0%	11.6%	14.5%	17.8%	8.5%	12.3%	16.9%	22.9%	34.3%
Physics B	1.4%	2.3%	3.6%	5.0%	8.7%	15.2%	22.7%	30.4%	38.8%	49.8%
Physics C: Mechanics	1.9%	2.9%	3.8%	5.1%	9.3%	20.9%	31.0%	38.0%	45.2%	55.9%
Psychology	4.6%	5.8%	7.1%	9.1%	11.2%	4.4%	6.8%	9.1%	13.1%	20.3%
Spanish Language	2.1%	4.0%	5.3%	6.4%	7.4%	12.4%	15.1%	15.2%	14.7%	14.4%
Statistics	1.9%	3.6%	6.3%	11.9%	22.4%	5.9%	9.6%	14.8%	24.1%	39.5%
U.S. Gov. and Politics	5.0%	8.0%	11.1%	14.7%	18.1%	6.1%	11.1%	16.0%	20.6%	24.9%
U.S. History	0.6%	1.4%	2.5%	4.2%	7.0%	7.0%	12.1%	16.0%	19.9%	23.3%
World History	0.5%	1.1%	2.2%	4.0%	6.5%	6.8%	11.0%	16.0%	20.4%	24.7%

Notes: Each cell indicates the probability of majoring in the two-digit CIP code categorization most closely associated with the AP exam.

Table 3
Probability of Majoring in CIP Code by AP Exam Subject, Student Scoring 3 or Higher

	College Major and CIP Code							
AP Exam	English Language and Literature/ Letters 23	History 54	Social Sciences 45	Psychology 42	Foreign Languages, Literatures, and Linguistics 16	Biological and Biomedical Sciences 26	Physical Sciences 40	Engineering/ Mathematics and Statistics 14/27
Biology	2.6%	2.0%	8.7%	4.9%	2.3%	18.9%	2.7%	7.8%
Calculus AB	2.0%	1.5%	7.9%	3.6%	2.1%	9.7%	3.1%	13.4%
Calculus BC	1.5%	1.4%	9.1%	3.0%	2.1%	11.4%	4.7%	21.7%
Chemistry	1.5%	1.4%	7.8%	3.1%	2.0%	14.2%	6.7%	18.2%
English Language & Comp.	4.8%	2.5%	10.0%	4.9%	2.8%	7.4%	1.9%	6.6%
English Literature & Comp.	5.4%	2.7%	10.0%	4.9%	2.8%	7.3%	1.9%	6.0%
Environmental Science	3.2%	2.6%	11.5%	4.5%	2.1%	7.6%	1.9%	5.7%
European History	4.5%	5.7%	13.5%	4.0%	3.1%	7.0%	2.1%	6.5%
French Language and Culture	4.8%	3.1%	14.3%	4.8%	7.9%	8.3%	2.5%	6.5%
Macroeconomics	2.3%	2.4%	13.4%	3.3%	2.0%	8.2%	2.4%	11.7%
Microeconomics	2.1%	2.2%	14.3%	3.1%	2.0%	7.7%	2.4%	11.8%
Physics B	1.6%	1.4%	7.9%	2.6%	1.7%	9.4%	5.2%	20.4%
Physics C: Mechanics	1.1%	1.0%	7.3%	1.7%	1.3%	7.6%	6.2%	29.9%
Psychology	3.2%	1.9%	8.5%	9.2%	2.2%	7.1%	1.4%	4.9%
Spanish Language	2.9%	1.9%	11.0%	5.0%	6.3%	6.9%	1.6%	5.7%
Statistics	2.2%	1.8%	10.0%	4.4%	2.0%	7.8%	2.3%	11.8%
U.S. Gov. and Politics	3.5%	3.7%	13.5%	3.9%	2.4%	7.6%	2.2%	8.2%
U.S. History	4.0%	4.0%	12.2%	4.3%	2.8%	7.9%	2.2%	7.8%
World History	3.6%	3.7%	11.4%	4.1%	2.7%	7.9%	2.2%	7.9%

Notes: Each cell indicates the probability of majoring in the two-digit CIP code categorization. Bolded and underlined cells are the outcome major used in all regressions.

IV. Methodology

In this section, we describe the methodology to estimate the effect of a marginal change in AP exam scores on major choice. This notation and methodology is similar to that of Smith, Hurwitz, and Avery (2017). Each student i on AP exam j receives a continuous score C_{ij}. This continuous score maps into the scaled score, T_{ij} as follows[19]:

$$T_{ij} = \begin{cases} 1 \text{ if } C_{ij} < t_j^2 \\ 2 \text{ if } t_j^2 \leq C_{ij} < t_j^3 \\ 3 \text{ if } t_j^3 \leq C_{ij} < t_j^4 \\ 4 \text{ if } t_j^4 \leq C_{ij} < t_j^5 \\ 5 \text{ if } t_j^5 \leq C_{ij} \end{cases}$$

where t_j^{n} are the thresholds for each scaled score n on exam j. For each value of $n \in \{2,3,4,5\}$, we create two variables. The first is the forcing variable:

$$Dist_{ijn} = C_{ij} - t_j^n$$

which captures how far student i's score on exam j is from threshold n. A $Dist_{ijn} \geq 0$ implies that the student has a scaled scores of at least an n. This leads to the second variable for each value of n, the dichotomous threshold variable:

$$Threshold_{ijn} = \begin{cases} 1 \text{ if } Dist_{ijn} \geq 0 \\ 0 \text{ if } Dist_{ijn} < 0 \end{cases}$$

After generating these variables, our basic empirical framework is shown by the standard regression discontinuity equation presented in Equation 1, where X_{ij} is a vector of fixed effects for the student's year of high school graduation and the interaction of the AP exam subject and year the exam is taken.

$$(1) \quad Outcome_{ijn} = \alpha_0^n + \alpha_1^n Threshold_{ijn} + \alpha_2^n Dist_{ijn} + \alpha_3^n Threshold_{ijn} \times Dist_{ijn} + X_{ij} + \varepsilon_{ijn}$$

We are primarily interested in the estimate of α_1^n, which is the coefficient on $Threshold_{ijn}$ that represents the discontinuous effect of being above the AP scaled n threshold on the outcome of interest. In practice, we separately estimate the effects of each scaled threshold.

The dependent variable in Equation 1 is often an indicator variable for an outcome at each threshold n, which is typically whether a student majors in the same subject or the same field as the AP exam subject. In order to capture trends in the forcing variable that exist on either side of the boundary, we fit a local linear regression with a triangular kernel. The triangular kernel puts more weight on the observations closest to the threshold. In all regressions, we use a bandwidth of 10, which is roughly equal to the optimal bandwidth suggested by Imbens and Kalyanaram (2012).[20]

19. Technically, $T_{ij,}$ varies by year, but for ease of exposition, we omit a year subscript.

20. We test the sensitivity to bandwidth and kernel choices and find no measurable differences. These robustness tests are presented in Online Appendix Table 4. We obtain the IK-estimated optimal bandwidth using software designed by Calonico, Cattaneo, and Titiunik (2014).

Researchers implementing regression discontinuity designs may confront challenges if score manipulation or gaming takes place in the vicinity of thresholds. In this context, such manipulation is essentially impossible, as grading standards and score thresholds vary from year to year and are never reported to students. Score thresholds are also predetermined by psychometricians and do not result from natural variation in students' scoring patterns on an exam within a given year. In the next section, we run covariate balancing tests with similar specifications to Equation 1, but using the covariates as outcomes, and find no indication of manipulation of raw scores near scaled score thresholds.

V. Main Results

A. Testing the Assumptions of Regression Discontinuity

In Figure 1, we show the density of raw scores near each threshold. For each of the 19 exams in each of the years the exam is offered, the threshold is centered at zero, and then

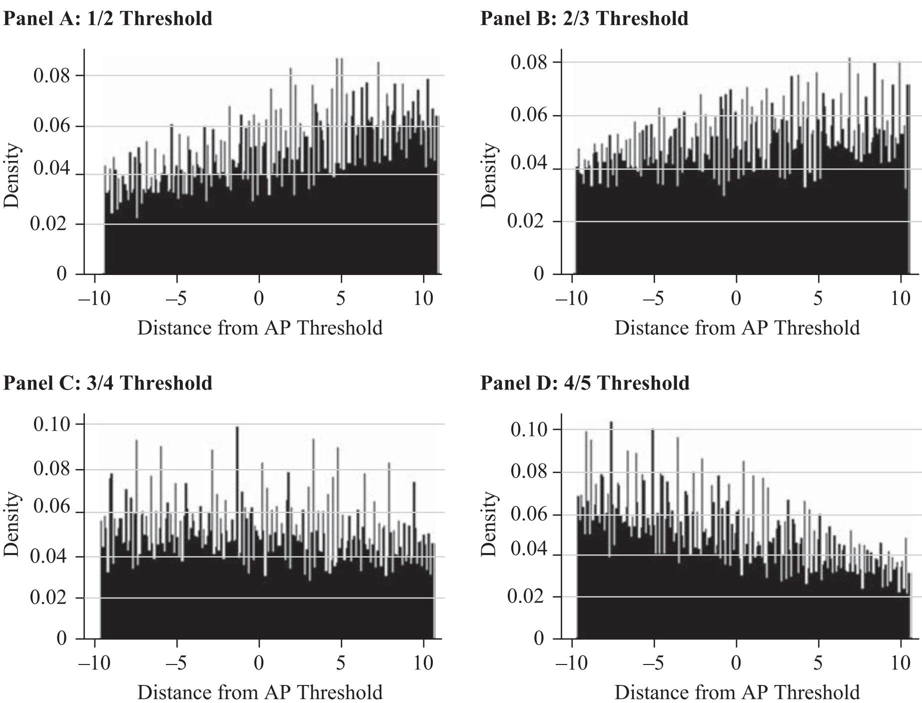

Figure 1

Density of Students around Thresholds

Notes: Figure includes all student–exam observations within ten points of the integer AP score threshold for the 19 AP exams listed in Table 2 for the years 2004 through 2009 graduating high school cohorts. Sample sizes for the 1/2, 2/3, 3/4, and 4/5 thresholds are 1,473,612, 2,383,844, 2,472,178, and 1,679,162 observations, respectively.

Table 4
Covariate Balancing Tests

						Parent Education		
	Male	White	Asian	Black	Hispanic	Less Than HS	HS graduate	BA or higher
Panel 1: 1/2 Threshold								
Above Threshold	0.0017 (0.0018)	–0.0009 (0.0017)	0.0015 (0.0012)	0.0002 (0.0010)	–0.0005 (0.0012)	–0.0004 (0.0011)	–0.0001 (0.0013)	0.0017 (0.0018)
N	1,473,612	1,473,612	1,473,612	1,473,612	1,473,612	1,473,612	1,473,612	1,473,612
Panel 2: 2/3 Threshold								
Above Threshold	–0.0010 (0.0014)	–0.0020 (0.0013)	0.0013 (0.0009)	–0.0005 (0.0006)	0.0015+ (0.0008)	0.0000 (0.0007)	0.0028** (0.0010)	–0.0021 (0.0014)
N	2,383,844	2,383,844	2,383,844	2,383,844	2,383,844	2,383,844	2,383,844	2,383,844
Panel 3: 3/4 Threshold								
Above Threshold	–0.0034* (0.0014)	–0.0035** (0.0012)	0.0017+ (0.0010)	–0.0008+ (0.0005)	0.0012+ (0.0007)	–0.0007 (0.0006)	0.0014 (0.0009)	0.0010 (0.0013)
N	2,472,178	2,472,178	2,472,178	2,472,178	2,472,178	2,472,178	2,472,178	2,472,178
Panel 4: 4/5 Threshold								
Above Threshold	–0.0016 (0.0016)	–0.0011 (0.0015)	0.0032* (0.0012)	–0.0004 (0.0005)	–0.0012 (0.0007)	–0.0001 (0.0006)	–0.0002 (0.0009)	0.0020 (0.0016)
N	1,679,162	1,679,162	1,679,162	1,679,162	1,679,162	1,679,162	1,679162	1,679,162

Notes: + $p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. All students in the sample first attended a four-year college within 180 days of high school graduation. An observation is a student AP exam. Results based on local linear regressions with triangular kernels of bandwidth 10 that include fixed effects for AP exam year and high school graduation year.Other variables include the Distance from the threshold and the interaction of Distance and Above Threshold.Standard errors clustered by individual.

the raw scores from the stacked exams are collapsed into one-point bins. Continuous density in the vicinity of each of the 1/2, 2/3, 3/4, and 4/5 thresholds is evident in this figure.[21]

21. The formal approach recommended by (McCrary 2008) to test for continuous density around thresholds may not be appropriate in light of the scoring rubric of most AP exams. Raw scores generally extend out to four decimal places, but most raw scores are simply unattainable based on the combination of correct and incorrect responses. Moreover, the distances between consecutive attainable raw scores appear to differ within AP exams, as does the probability of achieving these raw scores based on combinations of points earned/deducted from the multiple choice and free response sections. To illustrate, among students who took the 2008 administration of AP Biology and were just on the cusp of 2/3 threshold, 18 had forcing variable values of –0.0435, followed by one student who had a forcing variable value of exactly 0, 12 students with forcing variable values of 0.0008, and so on. As is also the case when the data are discrete (see Frandsen 2017), this type of clustering, which is obviously not reflective of score manipulation, presents a challenge to the traditional McCrary test.

Income								
<$50k	$50k–$100k	Income > $100k	Took SAT	SAT Score	High School FRPL	Total AP within Year	Average AP Score within Year	Predicted Score
0.0002	–0.0014	–0.0004	0.0004	0.1553	–0.0504	0.0028	–0.0054	–0.0032
(0.0013)	(0.0014)	(0.0013)	(0.0014)	(0.5107)	(0.0804)	(0.0038)	(0.0041)	(0.0058)
1,473,612	1,473,612	1,473,612	1,473,612	1,195,599	1,223,338	1,473,612	931,157	1,473,612
0.0011	0.0014	–0.0022*	0.0015	0.7791*	0.1047+	0.0024	–0.0020	–0.0092
(0.0009)	(0.0011)	(0.0011)	(0.0011)	(0.3811)	(0.0569)	(0.0031)	(0.0033)	(0.0066)
2,383,844	2,383,844	2,383,844	2,383,844	1,972,409	1,928,003	2,383,844	1,623,492	2,383,844
–0.0006	–0.0003	0.0022+	0.0009	–0.9662**	–0.0266	0.0021	–0.0021	0.0058
(0.0008)	(0.0011)	(0.0012)	(0.0010)	(0.3649)	(0.0516)	(0.0032)	(0.0030)	(0.0063)
2,472,178	2,472,178	2,472,178	2,472,178	2,113,990	1,933,772	2,472,178	1,848,836	2,472,178
–0.0001	0.0016	–0.0003	0.0020+	0.6089	–0.0307	0.0017	0.0020	0.0205
(0.0009)	(0.0013)	(0.0015)	(0.0011)	(0.4268)	(0.0601)	(0.0041)	(0.0033)	(0.0129)
1,679,162	1,679,162	1,679,162	1,679,162	1,485,920	1,277,910	1,679,162	1,351,907	1,679,162

Covariate balancing tests in Table 4 generally show balance across the thresholds. Among the 68 separate covariate balancing tests, which examine student sex, ethnicity, parental education and income, high school poverty, SAT scores, and AP course-taking, seven yield statistically significant parameter estimates at the 0.05 level.[22] All seven precisely estimated differences are extremely small in magnitude. In addition, some of the imbalanced covariates are highly related (that is, an imbalance in one race category naturally lends itself to an imbalance in another race category). Online Appendix Figure 1 shows graphical results for the few variables that exhibit the highest level of imbalance, and none of the pictures are suggestive of large jumps in student characteristics at the threshold.

22. Earlier work finds that a higher AP score in 11th grade produces a small causal increase in the likelihood that a student takes AP courses in 12th grade (Smith, Hurwitz, and Avery 2017). In order to avoid this endogeneity we only focus on within-year course-taking, as crossing an AP score threshold should have no impact on the amount of APs taken or scores in the same year.

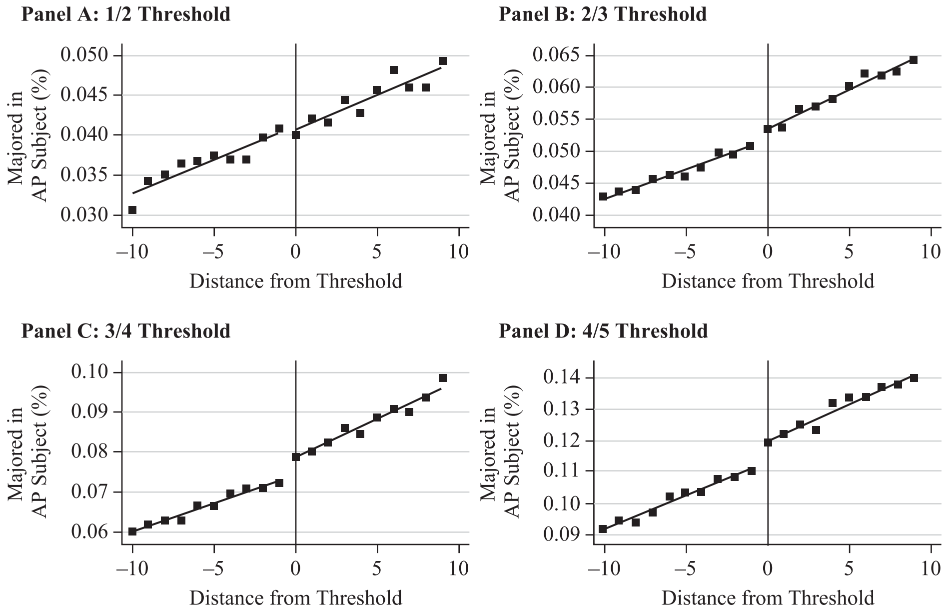

Figure 2
Main Results

Notes: Figure includes all student–exam observations within ten points of the integer AP score threshold for the 19 AP exams listed in Table 2 for the years 2004 through 2009 graduating high school cohorts. Sample sizes for the 1/2, 2/3, 3/4, and 4/5 thresholds are 1,473,612, 2,383,844, 2,472,178, and 1,679,162 observations, respectively.

We test overall balance two ways. First, we construct threshold-specific predicted raw scores based on a model that includes SAT scores, ethnicity, sex, parental education level, and income level, using only those observations below the relevant threshold. Estimates of changes in this predicted score at the threshold are null, and they are included as the last column in Table 4. Second, running a seemingly unrelated regression to jointly test for balance shows small imbalances on the 2/3 and 3/4 thresholds, although there is no evidence of imbalance on the 4/5 threshold, where we have our most prominent results. Finally, though not shown, results are largely the same when using IK bandwidths for each covariate.[23]

B. Main Regressions

Figure 2 presents our primary set of results on whether receiving higher AP exam scores causes students to major in the same subject as the AP exam. There are clear, observable

23. Despite the overwhelming evidence that there is no manipulation around the threshold, we estimate regressions that remove donut holes and heaps around the thresholds. These robustness tests are described at the end of the next section.

Table 5
Effect of Attaining Higher AP Exam Scores on Major

	Outcome = Majored in Same Subject as AP Exam			
	Threshold			
	1/2	2/3	3/4	4/5
Panel 1: Full Sample				
Above threshold	–0.0003	0.0018**	0.0038**	0.0064**
	(0.0007)	(0.0006)	(0.0007)	(0.0011)
Mean at cutoff	4.2%	5.4%	8.1%	12.2%
N	1,473,612	2,383,844	2,472,178	1,679,162
Panel 2: Only AP STEM Exams				
Above threshold	–0.0006	0.0011	0.0022	0.0053**
	(0.0012)	(0.0013)	(0.0015)	(0.0019)
Mean at cutoff	5.4%	7.5%	10.8%	15.7%
N	626,287	770,240	803,432	635,615
Panel 3: Only AP Non-STEM Exams				
Above threshold	–0.0001	0.0022**	0.0045**	0.0073**
	(0.0008)	(0.0007)	(0.0008)	(0.0012)
Mean at cutoff	3.3%	4.4%	6.8%	10.0%
N	847,325	1,613,604	1,668,746	1,043,547

Notes. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$. All students in the sample first attended a four-year college within 180 days of high school graduation. An observation is a student AP exam. Results based on local linear regressions with triangular kernels of bandwidth 10 that include fixed effects for AP exam year and high school graduation year. Other variables include the Distance from the threshold and the interaction of Distance and Above threshold. Standard errors clustered by individual. Means at cutoff are based on all students within one point below the designated threshold.

differences in student major at the thresholds, particularly as students cross into AP scores of 4 or 5. Table 5 provides regression estimates for the magnitude of these effects, with each coefficient from a separate regression that represents the causal effect of receiving a higher AP score on the corresponding threshold. Results in the first row show parameter estimates for the full sample, with separate results in subsequent rows for the set of STEM and non-STEM exams. Results for STEM and non-STEM exams are also shown graphically in Figure 3.

The first coefficient shows that receiving a score of 2 over a 1 on the sampled AP exams does not shift students' college majors into the AP exam field. This finding is unsurprising because scores of 1 and 2 are both considered nonpassing scores, and colleges rarely offer credit for either score. A score of 1 could be construed as an extremely negative signal and result in a disincentive to major in the subject. Each successive integer jump above the 1/2 margin leads to a larger boost in the probability

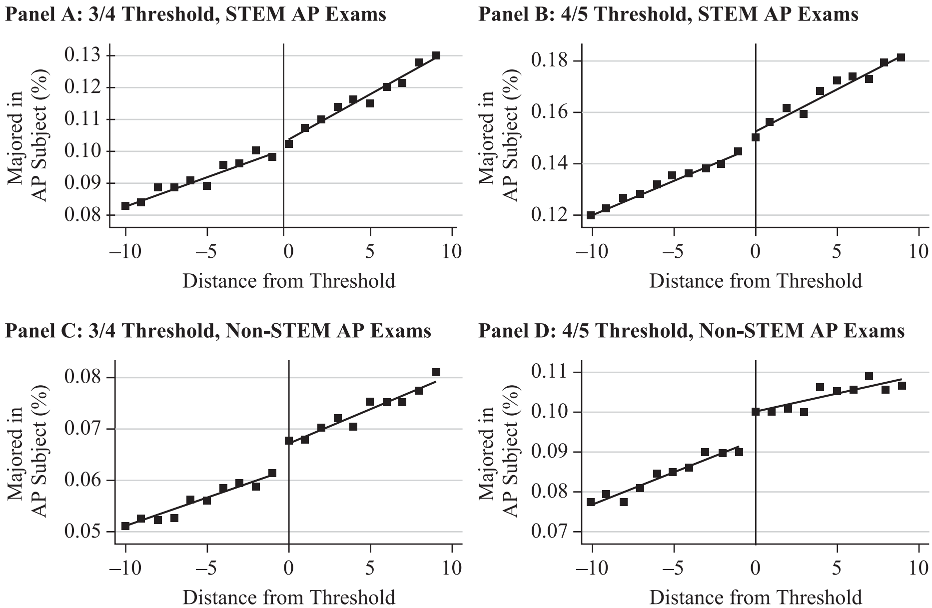

Figure 3

Main Results: STEM and Non-STEM Fields

Notes: Figure includes all student–exam observations within ten points of the integer AP score threshold for the 19 AP exams listed in Table 2 for the years 2004 through 2009 graduating high school cohorts. Sample sizes for the STEM AP exams at the 3/4 and 4/5 thresholds are 803,432 and 635,615 observations, respectively. Sample sizes for the Non-STEM AP exams at the 3/4 and 4/5 thresholds are 1,668,746 and 1,043,547 observations, respectively.

that a student will choose a major in the same subject as the AP exam. Across all sampled exams, jumps in the probability that the student major matches that AP exam subject increases by approximately 0.2 pp (3.3 percent), 0.4 pp (4.7 percent), and 0.6 pp (5.2 percent) from receiving AP scores of 3, 4, and 5, respectively. Subject-by-subject results are presented in Online Appendix Table 3 and demonstrate that there appears to be a distribution of effects, with upper bound estimates in the range of two percentage points (and 30 percent). This is more succinctly demonstrated in Figure 4, which plots the coefficient estimates of the 19 exams at each threshold. There is a clear pattern of positive results, particularly at the 4/5 threshold and in the non-STEM subjects.

When AP exams are separated into STEM and non-STEM exams, two different stories emerge. Receiving a higher integer AP score on a STEM exam tends to yield a statistically insignificant change in student major, except at the 4/5 threshold, where students are 0.5 pp (3.4 percent) more likely to major in the AP subject. By contrast, coefficients for non-STEM exams are statistically significant at all margins other than the 1/2 threshold and are larger than the STEM results.

In a set of robustness tests, we repeat the analysis reported in Table 5, while imposing some changes in the underlying empirical specification. Online Appendix Table 4 reports the results of analysis with different choices of bandwidth, kernel, the choice of

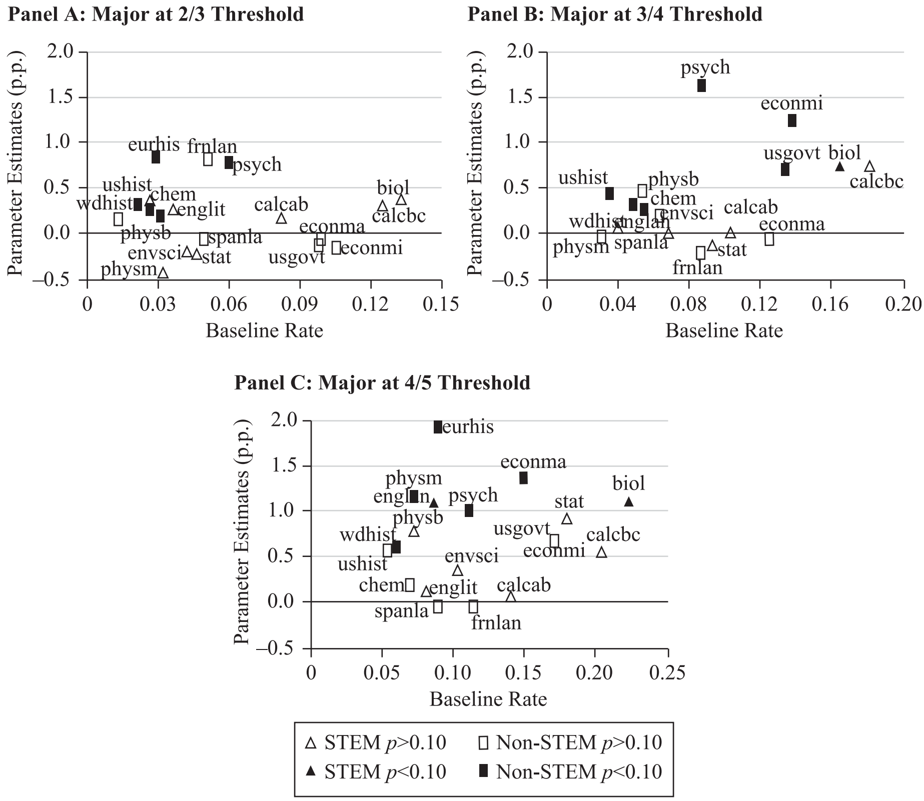

Figure 4
Subject by Subject Results

controls, and the number of higher order expressions of the forcing variable (thereby altering the functional form). Results are very similar to those in Table 5.[24]

Online Appendix Table 5 reports the results of analyses using different rules for inclusion and exclusion of students from the sample, restricting analyses to: students who graduated from high school in the 2005–2007 cohorts so that all students are tracked for six years (Panel 1), students who majored in a field where the CIP codes were provided by NSC and not hand coded by the researchers (Panel 2), and students with a unique rather than a "double" major (Panel 3). All results are similar to those reported in Table 5.

Online Appendix Table 6 deals with any potential concerns related to the bunching of the exam distributions. The left panel utilizes a "donut hole" approach and removes all observations within one point of the threshold.[25] The right panel removes any "heaps"

24. Using smaller bandwidths and triangular kernels alleviates the small issue that a ten point bandwidth includes two thresholds on three exams in a few years (never 1/2 or 4/5). In nearly identical results not shown, we also drop these exam–years.

25. Alternative sized donut holes produce similar results.

within the distribution that are outsized relative to their neighboring raw scores.[26] Neither robustness check changes the results.

B. Mechanisms

Higher AP scores may alter college major through multiple mechanisms, which we explore in two subsections. First, we separately explore the contributions of endogenous college enrollment and that of producing more graduates; the former of which we rule out and the latter of which is only marginally altered by a higher AP integer score among students with otherwise similar exam performance. Second, we decompose the estimates into the mechanical effect of credit receipt versus the behavioral response to a strong signal. We find strong support that the behavioral responses to higher AP scores are the primary drivers of our estimates, though we cannot discount the possibility that credit-granting policies are playing a small role in the shifting of majors.

1. College Enrollment and Graduation

The first four columns of Table 6 indicate whether students alter their college enrollment decisions from the receipt of a higher integer AP score. These results help differentiate whether the impacts on student major arise from changes in the schooling environment, as opposed to shifts in individual preferences. The first two columns show trivial differences in college choice arising from different integer scores, with small and often insignificant effects on school quality, as measured by average SAT or Barron's ranking.[27]

To further allay any concerns that our primary results are being driven by shifts in college choice, we repeat the primary analyses from Table 5 in two distinct ways. First, we refit our main models using only students taking AP exams in their senior year (Column 3), after college enrollment decisions have already been made, and we continue to find positive and statistically significant results comparable to those shown in Table 5. In fact, results that differentiate between senior and nonsenior exams (Online Appendix Table 7) indicate that higher AP scores earlier in high school have significant impacts on college quality, and shifts in college major for this group are roughly twice as large as for senior exams. These results suggest that earlier signaling allows students more opportunities to make decisions prior to graduation (for example, senior year courses and college application and enrollment) that facilitate major choice. Second, we then refit our main models using college fixed-effects specification (Column 4), and, again, our results are unchanged from those shown in Table 5. Taken together, these results suggest that there are impacts of the AP score that go beyond the college environment.

Finally, we test whether students with virtually similar exam performance, but different integer scores, endogenously choose colleges that offer credit for the scores they

26. In order to implement this method, we bin raw scores to reduce the variance, run linear regressions over short bandwidths across each test-by-year distribution, and remove heaps that are more than 25% greater than their predicted value by the linear regression. This identifies roughly 15% of the distribution. Choice of initial binning, regression bandwidth, or deviance from predicted value do not change the results.

27. We also test other measures of college quality, including each Barron's ranking individually, college graduation rates, and other potentially relevant measures, such as attending school out of state. We similarly find insignificant results, which are available upon request. Our earlier paper (Smith, Hurwitz, and Avery 2017) studies this possibility in more detail and similarly finds no evidence that AP scores influence the choice of colleges by students.

Table 6
Potential Mechanisms for Impacts of Higher AP Exam Scores on College Major

	College Choice					College Graduation	
	College's Average SAT (1)	Barrons Most, Highly, or Very Competitive (2)	Major in Subject, Senior Exams Only (3)	Major in Subject, College FE (4)	Schools Offers Credit at Threshold (5)	Bachelor in Six Years (6)	Major in Subject, College Graduates (7)
Panel 1: 1/2 Threshold							
Above threshold	0.4266	0.0026	–0.0001	–0.0004		–0.0037+	–0.0000
	(0.3993)	(0.0017)	(0.0010)	(0.0007)		(0.0021)	(0.0010)
Mean at cutoff	1,140	59.0%	5.3%	4.2%		75.0%	6.2%
N	1,427,550	1,473,612	928,304	1,473,612		831,234	982,436
Panel 2: 2/3 Threshold							
Above threshold	0.4230	0.0009	0.0013	0.0020**	0.0001	0.0027+	0.0022**
	(0.3131)	(0.0013)	(0.0009)	(0.0006)	(0.0015)	(0.0014)	(0.0008)
Mean at cutoff	1173	69.4%	6.3%	5.4%	59.3%	82.0%	7.1%
N	2,325,531	2,383,844	1,475,603	2,383,844	1,956,213	1,386,828	1,768,736
Panel 3: 3/4 Threshold							
Above threshold	0.6023+	0.0026*	0.0033**	0.0035**	–0.0031*	0.0029*	0.0044**
	(0.3219)	(0.0011)	(0.0010)	(0.0007)	(0.0015)	(0.0013)	(0.0009)
Mean at cutoff	1213	78.7%	9.1%	8.1%	55.6%	85.7%	10.0%
N	2,421,632	2,472,178	1,553,288	2,472,178	2,045,180	1,435,783	1,966,483
Panel 4: 4/5 Threshold							
Above threshold	0.9922*	0.0006	0.0042**	0.0068**	0.0019	0.0032*	0.0076**
	(0.4087)	(0.0012)	(0.0014)	(0.0011)	(0.0014)	(0.0014)	(0.0012)
Mean at cutoff	1258	85.8%	13.3%	12.2%	20.4%	88.7%	14.5%
N	1,650,548	1,679,162	1,057,689	1,679,162	1,402,697	953,759	1,401,213

Notes: $^{+}p<0.10$, $^{*}p<0.05$, $^{**}p<0.01$. All students in the sample first attended a four-year college within 180 days of high school graduation. An observation is a student AP exam. Results based on local linear regressions with triangular kernels of bandwidth 10 that include fixed effects for AP exam year and high school graduation year. Other variables include the Distance from the threshold and the interaction of Distance and Above threshold. Standard errors clustered by individual.

attain. Column 5 of Table 6 tests whether students are more likely to enroll at a college that offers additional credit for a higher AP score. We find no statistical evidence to support this at the integer score thresholds, other than a small negative coefficient on the 3/4 threshold, which paradoxically suggests that a student is less likely to attend a college if that college offers the student additional credit for a score of 4 over a 3. Combined, the first five columns suggest that there is no evidence that endogenous college enrollment is driving the main results.

As we only can identify a student's choice of major for those students listed with a BA degree in the NSC data, the results in Table 5 could conceivably reflect an effect of AP credit on college graduation rather than on the choice of college major. Column 6 of Table 6 shows small increases in six-year completion rates at the 2/3, 3/4, and 4/5 thresholds, which is consistent with Smith, Hurwitz, and Avery (2017).[28] However, when we condition on bachelor's completion (Column 7), we find nearly identical point estimates to those shown in Table 5. This provides reassuring evidence that we can isolate the effects of higher AP exam scores on shifting college major from the documented effects on the production of more college majors.

2. *Signal versus College Credit*

With the mechanism(s) largely unexplained as of yet, we explore two alternatives: college-specific credit policies that reduce major course requirements, which we label "mechanical," or the behavioral response to higher scores. The behavioral response may be a result of positive affirmation of a student's ability to succeed in a subject, but could be reaffirmed by other actors driving the decision process, such as parents, counselors, or even the college itself. An alternative behavioral response may be simply that students use the high score as a guidepost to in the course selection process, with no impact on self-confidence.

We exploit the rich variety in AP credit policies across postsecondary institutions and compare students at the 4/5 threshold (for example) who attend institutions where a score of 5 results in additional credit with similar students who attend institutions where no such credit is offered. Note that students with higher scores at colleges that give credit for those scores may benefit from the mechanical and behavioral impact of higher scores. Students only benefit from the behavioral impact of higher scores if their higher scores do not come with credit, and thus, we can compare the relative impacts across sets of institutions.

To separate the behavioral from the mechanical effects, in Column 1 of Table 7 we reproduce our main results using only our "policy sample" of 500 largest colleges, for which we collected detailed AP credit policy information. The results mimic those for the full sample in Table 5. The second column then shows results for the subsample of colleges that *offer additional credit or placement* for scores above versus below a particular scaled score threshold, whereas Column 3 shows effects at colleges that do not offer credit or placement (henceforth referred to as credit for the sake of brevity). Thus, Column 3 represents the pure behavioral effect, whereas the estimated effects reported in Column 2 represent a combination *of behavioral and mechanical effects* from receiving an increased AP integer score.

We find statistically significant increases in the probability of a matched college major due to the pure behavioral effect at the 3/4 and 4/5 thresholds. The behavioral effect is slightly smaller than the combined behavioral and mechanical effect at the 2/3 and 3/4 thresholds and slightly larger than the combined effect at the 4/5 threshold. The evidence in Table 7 suggests a strong behavioral effect from receiving higher AP integer scores, particularly at the 4/5 threshold, where the signal is strongest and changes in credit

28. We can only use 2004–2007 in these analyses. Using the full sample and four-year graduation rates, results are consistent with Smith, Hurwitz, and Avery (2017)—strong effects on the 2/3 and 3/4 thresholds, where college credit is often at stake.

Table 7
Effect of Attaining Higher AP Exam Scores on Major—Credit or Signal?

	Policy Sample			STEM AP Exams			Non-STEM AP Exams		
	All	With AP Policy	Without AP Policy	All	With AP Policy	Without AP Policy	All	With AP Policy	Without AP Policy
Panel 1: 2/3 Threshold									
Above threshold	0.0021**	0.0028**	0.0012	0.0014	0.0030	–0.0012	0.0025**	0.0026**	0.0022+
	(0.0007)	(0.0009)	(0.0011)	(0.0014)	(0.0018)	(0.0023)	(0.0008)	(0.0010)	(0.0012)
Mean at cutoff	5.4%	5.4%	5.4%	7.6%	7.9%	7.2%	4.3%	4.1%	4.6%
N	1,956,213	1,164,325	791,888	638,044	391,487	246,557	1,318,169	772,838	545,331
Panel 2: 3/4 Threshold									
Above threshold	0.0036**	0.0042**	0.0028*	0.0019	0.0027	0.0010	0.0044**	0.0049**	0.0037**
	(0.0008)	(0.0011)	(0.0012)	(0.0016)	(0.0022)	(0.0023)	(0.0009)	(0.0012)	(0.0014)
Mean at cutoff	8.1%	8.2%	8.0%	11.1%	11.8%	10.3%	6.7%	6.6%	6.8%
N	2,045,180	1,137,975	907,205	674,183	363,711	310,472	1,370,997	774,264	596,733
Panel 3: 4/5 Threshold									
Above threshold	0.0068**	0.0053*	0.0072**	0.0069**	0.0024	0.0085**	0.0069**	0.0080*	0.0066**
	(0.0012)	(0.0026)	(0.0013)	(0.0021)	(0.0041)	(0.0024)	(0.0014)	(0.0032)	(0.0015)
Mean at cutoff	12.4%	12.5%	12.4%	16.2%	15.2%	16.6%	9.9%	10.0%	9.9%
N	1,402,697	277,830	1,124,867	540,619	130,449	410,170	862,078	147,381	714,697

Notes: $^{+}p<0.10$, $^{*}p<0.05$, $^{**}p<0.01$. All students in the sample first attended a four-year college within 180 days of high school graduation. Policy sample includes 500 colleges where credit policies are collected. Colleges are defined as "with a policy" if there is any alteration in the units or courses offered at the threshold. An observation is a student AP exam. Results based on local linear regressions with triangular kernels of bandwidth 10 that include fixed effects for AP exam year and high school graduation year. Other variables include the Distance from the threshold and the interaction of Distance and Above threshold. Standard errors clustered by individual. Means at cutoff are based on all students within one point below the designated threshold.

receipt are uncommon. However, we are unable to rule out completely the possibility that the mechanical effect of receiving a higher AP score plays a small role in influencing a student's choice of college major.

Recall, the main results show no impact on major selection among students of similar exam performance who receive an integer score of 2 rather than 1. This implies that the signaling effect of a score of 1, as opposed to 2, does not cause students to shy away from majoring in the AP subject. Since there is almost never credit on the line, the impact (or lack thereof) should be considered behavioral and not mechanical. Given the strongest impact on the 4/5 margin and the null impact on the 1/2 margin, students are responding to positive signals and not responding to negative signals.

We further investigate the behavioral and mechanical effects of receiving higher AP integer scores separately for STEM and non-STEM AP exams. The middle set of columns in Table 7 report the results for STEM AP exams. We find a strong behavioral effect from receiving a score of 5 over a 4 on STEM AP exams. We report the results for non-STEM AP exams on the right of Table 7. In these specifications, we estimate that effects on college major that are of similar magnitudes, regardless of whether or not the higher scaled AP score earns the student more college credit.[29] The consistent similarity between these two sets of estimates suggests that the effect of an increased AP score on the choice of major is primarily behavioral in nature.[30]

3. Robustness Tests of the Behavioral Effect

As AP policies may have changed over time, we test the robustness of these results by using only the set of colleges and subjects whereby the minimum credit-granting AP exam score as reported in ASC in 2004 matches the data we collected from the colleges' websites in 2015.[31] Using the approximately 70% of exams that agree perfectly between the sources, estimates are largely unchanged and can be found in Online Appendix Table 8.

We next consider whether the behavioral effect is in fact students responding to higher scores or, rather, students responding to college-specific credit policies, even when students are on the cusp of an AP integer margin where there is no difference in credit. As an example, a college may give additional credit for a 3 (over a 2) and 5 (over a 4) in a subject but not a 4 (over a 3). In this setting, does the student infer from the college-specific policy that scores of 3 and 4 represent the same level of preparation in that AP subject? If this scenario played out in the data, we might expect null findings at these colleges for the impact of receiving higher AP scores. Removing students attending these types of colleges from the analyses may expose even larger behavioral responses among students attending colleges where the student is not primed to interpret scores of 3 and 4 (for example) as representing identical ability.

29. We also explore whether STEM and non-STEM students have access to earlier signals via the SAT II subject exams. We do not find results consistent with the idea that SAT II taking students are less susceptible, though results are fairly noisy.

30. Subject-by-subject results are in Online Appendix Table 3, but they are individually too imprecise to distinguish between these two mechanisms. In general, we find positive effects irrespective of whether the college does or does not have an AP policy at the threshold.

31. The ASC data contain only the minimum credit-granting thresholds, and it is unclear whether colleges interpret that to include instances of placement without credit, which is our approach in collecting data on the policy sample.

Table 8

Effect of Attaining Higher AP Exam Scores on Major—Credit or Signal with Uniform Credit Policies

Policy	Sample		2/3	3/4	4/5
No credit given for exam at any threshold	311 school–exam combinations (106 schools) that give no credit at any threshold. Environmental Science, World History, and European History are one-half of the sample.	Above threshold	–0.0011	0.0084+	0.0047
			(0.0050)	(0.0045)	(0.0050)
		N	34,048	56,323	62,216
Credit given for all thresholds	435 school–exam combinations (203 schools) that given credit at the 2/3, 3/4, and 4/5 thresholds. Biology, Chemistry, Spanish Language, and French Language are four-fifths of the sample.	Above threshold	0.0087*	0.0099+	0.0074
			(0.0042)	(0.0054)	(0.0078)
		N	58,256	50,920	31,721
Uniform credit policy across AP subjects	71 schools that offer AP credit at either the 2/3 or 3/4 threshold for every exam	Above threshold	0.0009	0.0022	0.0088**
			(0.0019)	(0.0020)	(0.0029)
		N	283,142	315,673	222,507
Uniform credit policy across AP subjects (2/3 threshold only)	52 schools that offer AP credit at the 2/3 threshold for every exam	Above threshold	0.0019	0.0026	0.0083*
			(0.0021)	(0.0023)	(0.0033)
		N	233,759	256,960	180,238

Notes: $^{+}p<0.10$, $^{*}p<0.05$, $^{**}p<0.01$. All students in the sample first attended a four-year college within 180 days of high school graduation. Analyses use credit policy sample, whereby policies are collected for 500 colleges. An observation is a student AP exam. Results based on local linear regressions with triangular kernels of bandwidth 10 that include fixed effects for AP exam year and high school graduation year. Other variables include the Distance from the threshold and the interaction of Distance and Above threshold. Standard errors clustered by individual.

To address this issue, we restrict attention to colleges that have uniform credit policies in two senses (Table 8). First, we look at the subsample of college–exam combinations for which the college does not offer credit for any AP scaled score. Repeating the analysis from Table 5 for this subsample provides a clean test of the behavioral effect described above. Not only is there no mechanical effect from credit, but students cannot infer anything from lack of credit offered at one score versus another.[32] As shown in the first row of Table 8, the estimated effect of an increase in AP score at the 3/4 and 4/5 thresholds is positive and of similar magnitude to our estimated effects from earlier results. However, these coefficients are also imprecisely estimated because of the relatively small subsample for college–exam pairs where there is no possibility of AP credit.

Second, we repeat this analysis for the subset of college–exam combinations where students receive credit at each of the 3, 4, and 5 thresholds. Once again, as reported in Row 2 of Table 8, the estimated effects of increased AP scaled score on choice of college major are positive, generally large in magnitude, but still somewhat imprecisely estimated.

Finally, some colleges have a blanket policy on their credit policies across all subjects, for example, by awarding credit for scoring a 3 on all exams with no additional credits offered at higher integer scores. Assuming students are aware of the blanket policy, they may not infer anything from the absence of credit increases on the other margins. Using only the subsample of colleges that have these blanket policies, we find consistent evidence, as reported in Rows 3 and 4 of Table 8, with the main results. Combined, these analyses provide evidence supporting the general accuracy of our earlier estimates in that students are responding to the positive signal, and this behavioral response is not dampened from the unique college-specific credit policies where they enroll.

VI. Additional Results

In this section we examine three sets of additional results pertinent to our findings: heterogeneous results across important demographic groups, how students respond to multiple signals, and overall impacts on STEM degree attainment.

A. Heterogeneous Effects of AP Credit

In this subsection, we investigate whether AP credits have heterogeneous effects by types of student or college. On the student side, we are especially interested in the effects of AP credits on subgroups, such as women, low-income families, and minority students traditionally underrepresented in STEM fields (Turner and Bowen 1999; Zafar 2015). We report the results of our analyses for each of these subgroups in the first eight rows of Table 9. One immediate challenge is that these subgroups of students are underrepresented in our AP samples (as evidenced by the fairly small sample sizes for these groups), thereby limiting the precision of our estimated effects for each of these subgroups. Subject to this caveat, we find only limited evidence of differential responses for

32. This is somewhat rare, but the most common exams include AP Environmental Science and AP World History.

Table 9
Effect of Attaining Higher AP Exam Scores on Major—Heterogeneity

	All AP Exams			STEM AP Exams			Non-STEM AP Exams		
	2/3	3/4	4/5	2/3	3/4	4/5	2/3	3/4	4/5
Male	0.0023*	0.0033**	0.0067**	0.0022	0.0028	0.0055*	0.0022*	0.0037**	0.0078**
	(0.0010)	(0.0011)	(0.0015)	(0.0020)	(0.0021)	(0.0027)	(0.0010)	(0.0012)	(0.0018)
N	1,024,316	1,140,410	836,630	369,339	418,061	359,260	654,977	722,349	477,370
Female	0.0016*	0.0040**	0.0062**	0.0001	0.0016	0.0050+	0.0022*	0.0050**	0.0068**
	(0.0008)	(0.0010)	(0.0015)	(0.0016)	(0.0019)	(0.0027)	(0.0009)	(0.0011)	(0.0017)
N	1,359,528	1,331,768	842,532	400,901	385,371	276,355	958,627	946,397	566,177
White	0.0020**	0.0040**	0.0075**	0.0012	0.0026	0.0051*	0.0024**	0.0047**	0.0091**
	(0.0007)	(0.0009)	(0.0013)	(0.0015)	(0.0017)	(0.0023)	(0.0008)	(0.0010)	(0.0015)
N	1,656,577	1,789,749	1,204,179	538,372	573,035	449,550	1,118,205	1,216,714	754,629
Asian	0.0023	0.0028	0.0029	0.0000	0.0016	0.0040	0.0037*	0.0036+	0.0022
	(0.0017)	(0.0019)	(0.0025)	(0.0034)	(0.0037)	(0.0044)	(0.0018)	(0.0021)	(0.0029)
N	308,939	345,783	273,790	119,323	137,425	125,712	189,616	208,358	148,078
Minority (black/ Hispanic)	0.0023	0.0041+	0.0026	0.0051	−0.0002	0.0026	0.0013	0.0056*	0.0026
	(0.0017)	(0.0023)	(0.0036)	(0.0040)	(0.0053)	(0.0082)	(0.0018)	(0.0025)	(0.0038)
N	323,417	239,491	134,227	83,214	63,093	36,986	240,203	176,398	97,241

(*continued*)

Table 9 (*continued*)

	All AP Exams			STEM AP Exams			Non-STEM AP Exams		
	2/3	3/4	4/5	2/3	3/4	4/5	2/3	3/4	4/5
Income <$50k	0.0045* (0.0018)	0.0046* (0.0023)	–0.0020 (0.0036)	0.0059 (0.0038)	0.0031 (0.0048)	–0.0047 (0.0068)	0.0039* (0.0019)	0.0055* (0.0025)	–0.0003 (0.0041)
N	299,503	248,206	144,763	91,086	79,173	52,964	208,417	169,033	91,799
Income $50k–$100k	0.0017 (0.0014)	0.0038* (0.0016)	0.0098** (0.0025)	0.0016 (0.0029)	0.0031 (0.0034)	0.0137** (0.0046)	0.0018 (0.0015)	0.0043* (0.0018)	0.0080** (0.0029)
N	485,175	483,721	308,762	153,935	153,702	115,658	331,240	330,019	193,104
Income >$100k	0.0001 (0.0014)	0.0051** (0.0015)	0.0089** (0.0021)	–0.0008 (0.0027)	0.0036 (0.0030)	0.0090* (0.0039)	0.0006 (0.0016)	0.0058** (0.0017)	0.0090** (0.0025)
N	474,661	557,844	409,834	157,629	179,685	151,643	317,032	378,159	258,191
Took one AP	–0.0007 (0.0017)	0.0035 (0.0025)	0.0041 (0.0047)	–0.0025 (0.0034)	–0.0017 (0.0045)	–0.0016 (0.0075)	0.0003 (0.0019)	0.0070* (0.0030)	0.0090 (0.0057)
N	328,563	225,638	101,150	114,393	86,042	45,804	214,170	139,596	55,346
Took two or more AP	0.0023** (0.0007)	0.0038** (0.0008)	0.0066** (0.0011)	0.0017 (0.0014)	0.0026+ (0.0015)	0.0058** (0.0020)	0.0025** (0.0007)	0.0044** (0.0008)	0.0073** (0.0013)
N	2,055,281	2,246,540	1,578,012	655,847	717,390	589,811	1,399,434	1,529,150	988,201
Bottom third SAT	0.0024* (0.0010)	0.0062** (0.0019)	0.0030 (0.0041)	0.0033 (0.0024)	0.0023 (0.0039)	0.0025 (0.0084)	0.0021+ (0.0011)	0.0080** (0.0021)	0.0033 (0.0045)
N	834,173	434,741	131,157	218,393	123,701	41,141	615,780	311,040	90,016

Middle third SAT	0.0005	0.0035**	0.0052*	–0.0010	0.0021	0.0019	0.0012	0.0043**	0.0077**
	(0.0011)	(0.0012)	(0.0022)	(0.0021)	(0.0024)	(0.0037)	(0.0012)	(0.0014)	(0.0026)
N	798,336	908,774	469,918	28,0039	29,6351	18,4270	518,297	612,423	285,648
Top third SAT	–0.0003	0.0016	0.0083**	–0.0014	0.0002	0.0096**	0.0005	0.0023+	0.0076**
	(0.0017)	(0.0012)	(0.0014)	(0.0028)	(0.0024)	(0.0025)	(0.0022)	(0.0014)	(0.0016)
N	339,900	770,475	884,845	143,872	265,633	331,807	196,028	504,842	553,038
Bottom third college quality (Avg. SAT)	0.0030**	0.0026*	0.0064**	0.0052*	–0.0011	0.0065	0.0021*	0.0042**	0.0063*
	(0.0009)	(0.0013)	(0.0023)	(0.0021)	(0.0028)	(0.0044)	(0.0009)	(0.0014)	(0.0026)
N	975,176	734,177	341,542	278,649	220,467	123,894	696,527	513,710	217,648
Middle third college quality (Avg. SAT)	0.0016	0.0039**	0.0083**	0.0002	0.0037	0.0094*	0.0023*	0.0040**	0.0079**
	(0.0011)	(0.0013)	(0.0020)	(0.0023)	(0.0026)	(0.0037)	(0.0012)	(0.0014)	(0.0023)
N	822,881	838,322	493,909	270,060	274,458	192,741	552,821	563,864	301,168
Top third college quality (Avg. SAT)	–0.0001	0.0043**	0.0054**	–0.0025	0.0029	0.0028	0.0014	0.0051**	0.0071**
	(0.0015)	(0.0013)	(0.0015)	(0.0024)	(0.0023)	(0.0027)	(0.0019)	(0.0015)	(0.0018)
N	527,474	849,133	815,097	204,439	293,389	309,072	323,035	555,744	506,025

Notes: $^{+}p<0.10$, $^{*}p<0.05$, $^{**}p<0.01$. Each estimate is a separate regression that is restricted to the identified sample. All students in the sample first attended a four-year college within 180 days of high school graduation. An observation is a student AP exam. Results based on local linear regressions with triangular kernels of bandwidth 10 that include fixed effects for AP exam year and high school graduation year. Other variables include the Distance from the threshold and the interaction of Distance and Above threshold. Standard errors clustered by individual.

any subgroup of students, regardless of the threshold or field, in the probability of majoring in the AP subject in response to a higher AP score.[33]

We consider separately the possibility of an interaction between SAT score and AP exam score. Specifically, when we split the sample into three SAT score ranges, as reported in Rows 9–11 of Table 9, we find similar estimated effects of AP score on college major for each of these subsamples. These results suggest that a change in AP scores has a similar effect on all students, regardless of that student's academic ability (as measured by SAT score).

On the college side, we split the sample by average SAT of all enrolled students at the colleges and report the results in Rows 12–14 of Table 9. Once again, we find little evidence of differential effects across the subsamples of colleges. These results suggest that the effects of higher AP scores are not localized to certain types of colleges.

B. Multiple Signals

In this section, we address how students shift majors when they receive multiple signals of ability. Students differ substantially in the number of AP exams taken and their performance on these exams, and both of these factors likely influence the extent to which an additional score of 5 (for example) alters student major. In the presence of many other positive signals through high AP exam scores, we hypothesize that receiving an additional AP score of 5 (for example) is less likely to shift a student's major into the focal AP exam subject relative to the effect such a signal might have on the student with no additional AP signals.

In Table 10, we present the results of a pooled regression in which the threshold dummy variable in Equation 1 is interacted with student's average performance on all other AP exams, while including fixed effects that control for the exact combination of AP exams taken by the student. We focus on the 4/5 margin because it is this threshold on which we find the largest effects throughout the paper, and we only include the multiple exam takers in this table. The first column of Table 10 demonstrates that the main effect of receiving a 5 over a 4 is similar in magnitude for multiple-exam takers, compared to the entire sample of students.

Interacting the average of a student's other AP exams with the 4/5 threshold indicator, we find that magnitude of the shift in college major into the focal AP subject is highly sensitive to the average of AP scores on the other exams taken by the student. For ease of interpretation, the average AP score on other exams is centered at three, indicating that a student with an average score of 3.0 on all other exams would be 0.9 percentage points more likely to major in the AP subject with a score of 5 over a 4. Across all exams, the coefficient of –0.0029 in Row 2, Column 2 suggests that each one point increase in average AP score on other AP exams mutes the focal exam's pull by about 0.3 percentage points. So for a student with an average of 4 on all other AP exams, scoring a 5

33. There is some evidence in Table 7 of larger estimated coefficients for "white" students than for other subgroups of students. However, the estimated effects for Asian and other minority students (Black and Hispanic) are also positive, and the standard errors are sufficiently larger that it does not seem plausible to conclude that there are differential effects across these groups. Similarly, the estimated effects at the 4/5 scaled score threshold appear to be smaller for students from lowest-income families (less than $50,000 in family income) than for others, but this is not the case at other score thresholds.

Table 10
Impact of Earning Multiple High Scores on AP Exams

	Multiple AP Takers	Other Exam = Average of Exam Scores (Centered at 3)		
	(1)	(2)	(3)	(4)
Above 4/5 threshold (Primary exam)	0.0068*** (0.0011)	0.0091*** (0.0013)	0.0077*** (0.0014)	0.0093*** (0.0014)
Above 4/5 threshold (Primary exam)* Other exam		–0.0029*** (0.0008)		
Above 4/5 threshold (Primary exam)* Other STEM exam			–0.0011+ (0.0007)	
Above 4/5 threshold (Primary exam)* Other non-STEM exam				–0.0030*** (0.0008)
N	1,578,216	1,578,012	1,256,934	1,449,178

Notes: $^{+}p < 0.10$, $*p < 0.05$, $**p < 0.01$. All students in the sample first attended a four-year college within 180 days of high school graduation. An observation is a student AP exam but only for students near the 4/5 threshold on at least one exam. Results based on local linear regressions with triangular kernels of bandwidth 10 that include fixed effects for AP exam subject of the forcing variable, AP exam year and high school graduation year. These regressions also contain fixed effects for the exact set of total AP exams taken by the student. Other variables include the Distance from the threshold and the interaction of Distance and Above Threshold. Standard errors clustered by individual.

over a 4 on an additional AP exam would increase the probability that she majors in that subject by about 0.6 percentage points (calculated as 0.0091 – 0.0029). It is also clear from Columns 3 and 4 that high average scores on other non-STEM exams have a notably stronger muting effect than do high average scores on STEM exams.[34]

We used this analytic strategy because of data limitations that come from the recommended approach by Papay, Murnane, and Willett (2016), but it comes at the cost of introducing alternative hypotheses.[35] One leading alternative explanation is that students with multiple positive signals, who are typically high-ability students, have a

34. Regressions that interact the threshold variable with alternate definitions of alternate AP exam performance, such as counts of the number of exams with scores of 3, 4, or 5, produce similar results.
35. With no data limitations, the preferred analytic strategy would be to follow Papay, Willett, and Murnane (2011) and test the simultaneous effects of crossing one or two thresholds. This approach is challenging in our context for a number of reasons that severely limit our statistical power. Predominantly, many students take three or more exams, requiring higher order interaction terms that increase exponentially, and interaction effects differ across four different boundaries, rather than only one. Results using the Papay, Willett, and Murnane (2011) approach for students taking two AP exams (not shown here) produce noisy estimates that are uninformative.

Table 11
Effect of Attaining Higher AP Exam Scores on Majoring in Any STEM Field

	Outcome = Majored in STEM			
	Threshold			
	1/2	2/3	3/4	4/5
Panel 1: Full Sample				
Above threshold	–0.0002	0.0021*	0.0001	–0.0021
	(0.0011)	(0.0010)	(0.0011)	(0.0014)
Mean at cutoff	11.7%	15.8%	20.2%	26.0%
N	1,473,612	2,383,844	2,472,178	1,679,162
Panel 2: Only AP STEM Exams				
Above threshold	–0.0012	0.0028	0.0010	0.0006
	(0.0019)	(0.0019)	(0.0021)	(0.0026)
Mean at cutoff	15.5%	20.5%	26.7%	34.9%
N	626,287	770,240	803,432	635,615
Panel 3: Only AP Non-STEM Exams				
Above threshold	0.0005	0.0018	–0.0003	–0.0035*
	(0.0014)	(0.0012)	(0.0013)	(0.0017)
Mean at cutoff	9.0%	13.5%	17.1%	20.4%
N	847,325	1,613,604	1,668,746	1,043,547

Notes: $^{+}p < 0.10$, $*p < 0.05$, $**p < 0.01$. All students in the sample first attended a four-year college within 180 days of high school graduation. An observation is a student AP exam. Results based on local linear regressions with triangular kernels of bandwidth 10 that include fixed effects for AP exam year and high school graduation year. Other variables include the Distance from the threshold and the interaction of Distance and Above threshold. Standard errors clustered by individual. Means at cutoff are based on all students within one point below the designated threshold.

difficult time interpreting the signals. However, Table 9 shows us that the highest ability students, as measured by SAT, are the most responsive to these positive signals.

C. STEM Degree Attainment

Both descriptive statistics in Table 2 and the causal estimates above imply that higher scores increase the likelihood that a student majors in a specific subject. However, major choice is typically a zero sum game—if a student majors in one subject, then she is likely forgoing the opportunity to major in a different subject. This is a key differentiator between this study and Smith, Hurwitz, and Avery (2017), which examined bachelor's degree completion. We show that receiving a higher integer score on the AP Biology exam increases the likelihood of majoring in Biology, but STEM production increases only if the student's counterfactual degree was in a non-STEM field, such as English or social sciences, rather than an alternate STEM degree, such as chemistry. This is important because a shortage of STEM majors is frequently cited as a deficit in our current educational system, and multiple policy levers have been enacted to combat this problem.

Table 11 suggests that, in general, we are unable to conclude that simply receiving a higher integer AP score on a STEM AP exam positively impacts STEM major completion, although we do observe a positive and statistically significant effect in the full sample at the 2/3 threshold. By contrast, we find some suggestive evidence that higher integer scores on non-AP STEM exams may draw students away from STEM fields in other non-STEM disciplines. These results show that positive signals of high AP scores alone may not be enough to shift students into STEM fields, as STEM-focused students may enter college with stronger major intentions. Yet, it is important to remember that exposure to any subject may have independent effects on majoring in that subject (Fricke, Grogger, and Steinmayr 2015), and this includes exposure to STEM curriculum in the promotion of STEM degrees, which we cannot test here.

VII. Discussion and Conclusion

This paper shows that students incorporate signals of their relative academic performance in determining an important human capital decision: choice of college major. Although high school graduates have received countless sources of feedback over their lifetime, our results suggest that performance labels provided late in secondary school can have large impacts on subsequent educational investment decisions. This paper adds to the growing evidence that educational interventions late in students' high school careers can have significant impacts on their subsequent postsecondary decisions. Although not as consequential towards the immediate college enrollment process as interventions that provide information or structured guidance (Carrell and Sacerdote 2017; Hoxby and Turner 2013), students are using academic information gathered during high school to help make long-term career decisions. Similar to Hoxby and Turner (2013), our results are concentrated among high-performing students, though they are not restricted to specific geographic regions.

We find that the shift in college major is predominately a behavioral response and is strongest when the students have few other competing signals of academic excellence on AP exams. Our strong results may come from the signal salience, as AP exams are nationally recognized and known to be consequential in the college acceptance process and for college course credits. Other studies of academic signaling have produced mixed results. Papay, Murnane, and Willett (2016) find that students are more likely to attend college when they have a positive label that summarizes their score performance on a standardized test, though they cannot distinguish between the impact of the signal on student attitudes versus how the signal is used by schools as a sorting mechanism that might promote college-going behaviors. As in their paper, we also find the strongest impacts to occur at the highest levels of the test-score distribution. Foote, Schulkind, and Shapiro (2015) and Jackson (2015) find no impacts on any relevant college-going measures from signals of college-readiness in high school, perhaps because few students or institutions recognize the signal as particularly consequential, or the signals are poorly communicated.[36]

The impacts on student major found in this research, in combination with the decrease in time to degree found in Smith, Hurwitz, and Avery (2017), show that that AP scores affect the postsecondary choices and outcomes of different students in different ways. The earlier study finds that at the minimum credit-granting margin (generally the 2/3 scaled score margin), students receiving the higher integer score, despite otherwise similar performance on that AP exam, are more likely to complete a BA degree in four years, principally because credit receipt generally reduces the minimum credits for BA completion. By contrast, this study finds that the AP integer score primarily influences the choice of college major for students at the higher scaled score cutoff of 4/5 on most exams, when performance on that AP exam is otherwise similar to students who simply fell on the other side of the cutoff score. Although the magnitudes of these effects are generally less than 1 percentage point per test, they are not negligible by comparison to the cross-sectional correlations between AP score and college major. On average, the signaling effect of the higher score explains approximately 16 percent of the difference in the probability of majoring in the subject for students who receive a 5 versus a 4. Also, given the national scope of AP, small magnitudes in parameter estimates translate into thousands of students in each high school cohort.

Our results highlight that timely signals of academic preparation can impact major choice, yet we generally find statistically significant evidence of changes in college major within the broader classifications of "STEM" versus "Non-STEM," but not across these broad classifications. Our estimates of the effect of a higher integer score on an AP STEM exam and the probability of choosing a STEM major are consistently positive (Table 11), even though they lack statistical precision. That is, there may be small positive effects of AP integer scores on the choice of a STEM major that are beyond the power of the tests we can perform on existing data. In addition, we only observe student major when a student graduates from college, as opposed to intended college major at each stage of postsecondary enrollment. Given the evidence that many college students who aspire to complete STEM degrees switch to non-STEM disciplines (Arcidiacono, Aucejo, and Spenner 2012; Stinebrickner and Stinebrickner 2014), this may lead us to underestimate the potential impact of these signals on total effort towards STEM majors. Instead, our results are indicative of the power of signals to obtain achievable goals.

Why might we find differences in the effects of ability signals between STEM and non-STEM AP exams? Signaling effects may be weaker for STEM AP takers because these students may have already received many alternate and perhaps competing signals of preparation in that AP subject. For example, STEM AP takers may have received more consistent feedback from frequent tests that use grading standards on which the student might place more weight. In other words, students may perceive their evaluations in these subjects to have greater objectivity. This then suggests that developing skills in rigorous high school courses can help promote STEM completion. In addition, STEM students tend to take more AP exams, which we show mitigates any one signal, so it is certainly possible that variations on signal strength and timeliness in STEM fields can have sizable impacts. Stinebrickner and Stinebrickner (2014) find that "students

36. Other papers within the broader literature on the effects of positive signals of ability include Diamond and Persson (2016); Kosfeld and Neckermann (2011); Fryer, Levitt, and List (2008); and Steele and Aronson (1995).

enter school quite optimistic about obtaining a science degree, but that relatively few students end up graduating with a science degree, … [primarily due to] misperceptions about their ability to perform well academically in science." As AP takers enter college amongst the most highly prepared students in the nation, these results underscore the challenge of carrying out a plan to complete a major in a STEM field. Interventions that help students navigate introductory courses, perhaps through counseling or psychological supports (see, for example, Walton and Cohen 2011), may help retain these high-achievers in STEM fields.

Overall, the results in this paper suggest that positive signal of students' ability can change their major, and that timely provision of signals might produce larger shifts in outcomes. For example, providing students similar feedback earlier within their high school careers might increase subsequent effort or spur additional course-taking within desired fields. More research that identifies what aspects of various signals students find salient could help identify ability signals that yield the largest changes in student behavior. This may be a particularly desirable strand of research because these signals are nearly costless as compared to more traditional methods of producing STEM majors, such as outreach activities or financial incentives. As there are many opportunities for individuals and organizations to incentivize strategic goals, such as efforts to increase STEM majors, these results in this paper show promising evidence of low-cost signaling interventions to shift the distribution of college majors.

References

Altonji, Joseph G., Peter Arcidiacono, and Arnaud Maurel. 2015. "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects." *Handbook of the Economics of Education* 5(7):305–96.

Arcidiacono, Peter, Estaban M. Aucejo, and Ken Spenner. 2012. "What Happens after Enrollment? An Analysis of the Time Path of Racial Differences in GPA and Major Choice." *IZA Journal of Labor Economics* 1(5):1–24.

Arcidiacono, Peter, V. Joseph Hotz, and Songman Kang. 2012. "Modeling College Major Choices Using Elicited Measures of Expectations and Counterfactuals." *Journal of Econometrics* 166(1):3–16.

Beffy, Magali, Denis Fougère, and Arnaud Maurel. 2011. "Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter?" *Review of Economics and Statistics* 94(1):334–47.

Betts, Julian R. 1996. "What Do Students Know about Wages? Evidence from a Survey of Undergraduates." *Journal of Human Resources* 31(1):27–56.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik. 2014. "Robust Data-Driven Inference in the Regression-Discontinuity Design." *Stata Journal* 14(4):909–46.

Carnevale, Anthony P., Ban Cheah, and Andrew R. Hanson. 2015. *The Economic Value of College Majors*. Washington, DC: Center on Education and the Workforce, Georgetown University.

Carrell, Scott E., and Bruce Sacerdote. 2017. "Why Do College Going Interventions Work?" *American Economic Journal: Applied Economics* 9(3):124–51.

Castleman, Benjamin L., Bridget Terry Long, and Zachary Mabel. 2018. "Can Financial Aid Help to Address the Growing Need for STEM Education? The Effects of Need-Based Grants on the Completion of Science, Technology, Engineering, and Math Courses and Degrees." *Journal of Policy Analysis and Management* 37(1):136–66.

Chajewski, Michael, Krista D. Mattern, and Emily J. Shaw. 2011. "Examining the Role of Advanced Placement Exam Participation in 4-Year College Enrollment." *Education Measurement: Issues and Practice* 30:16–27.

Denning, Jeffrey T., and Patrick Turley. 2017. "Was That SMART?: Institutional Financial Incentives and Field of Study." *Journal of Human Resources* 52(1):152–86.

Diamond, Rebecca, and Petra Persson. 2016. "The Long-term Consequences of Teacher Discretion in Grading of High-Stakes Tests." NBER Working Paper 22207. Cambridge, MA: National Bureau of Economic Research.

Dodd, Barbara G., Steven J. Fitzpatrick, R.J. De Ayala, and Judith A. Jennings. 2002. "An Investigation of the Validity of AP Grades of 3 and a Comparison of AP and Non-AP Student Groups." College Board Research Report 2002-9. New York, NY: The College Board.

Dougherty, Chrys, Lynn Mellor, and Shuling Jian. 2006. "The Relationship between Advanced Placement and College Graduation." *Education* 501:1–35.

Dynarski, Susan M., Steven W. Hemelt, and Joshua M. Hyman. 2015. "The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes." *Educational Evaluation and Policy Analysis* 37:53S–79S.

Evans, Brent. 2013. "How Do College Students Use Advanced Placement Credit?" Unpublished.

———. 2017. "SMART Money: Do Financial Incentives Encourage College Students to Study Science?" *Education, Finance, and Policy* 12(3):342–68.

Foote, Andrew, Lisa Schulkind, and Teny M. Shapiro. 2015. "Missed Signals: The Effect of ACT College-Readiness Measures on Post-Secondary Decisions." *Economics of Education Review* 46:39–51.

Frandsen, Brigham R. 2017. "Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete," in *Regression Discontinuity Designs, Advances in Econometrics* 38, 281–315. Bingley, UK: Emerald Publishing Ltd.

Fricke, Hans, Jeffrey Grogger, and Andreas Steinmayr. 2015. "Does Exposure to Economics Bring New Majors to the Field? Evidence from a Natural Experiment." NBER Working Paper 21130. Cambridge, MA: National Bureau of Economic Research.

Fryer, Roland G., Steven D. Levitt, and John A. List. 2008. "Exploring the Impact of Financial Incentives on Stereotype Threat: Evidence from a Pilot Study." *American Economic Review* 98(2):370–5.

Goldin, Claudia. 2015. "Gender and the Undergraduate Economics Major: Notes on the Undergraduate Economics Major at a Highly Selective Liberal Arts College." Harvard University. Unpublished.

Hargrove, Linda, Donn Godin, and Barbara Dodd. 2008. "College Outcomes Comparisons by AP and Non-AP High School Experiences." Research Rep. 2008-3. New York, NY: The College Board.

Hastings, Justine, Christopher Neilson, and Seth D. Zimmerman. 2013. "Are Some Degrees Worth More than Others? Evidence from College Admission Cutoffs in Chile." NBER Working Paper 19241. Cambridge, MA: National Bureau of Economic Research.

Hoxby, Caroline, and Sarah Turner. 2013. "Expanding College Opportunities for High-Achieving, Low Income Students." Stanford, CA: Stanford Institute for Economic Policy Research.

Imbens, Guido, and Karthik Kalyanaram. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3):933–59.

Jackson, Jacob S. 2015. "Does an Early College Readiness Signal Discourage College Application and Enrollment?" *Journal of Research on Educational Effectiveness* 8(3):380–99.

Jackson, C. Kirabo. 2010. "A Little Now for a Lot Later: A Look at a Texas Advanced Placement Incentive Program." *Journal of Human Resources* 45(3):591–639.

Karp, Melinda Mechur, Jaun Carlos Calcagno, Katherine L. Hughes, Dong Wook Jeong, and Thomas Bailey. 2007. "The Postsecondary Achievement of Participants in Dual Enrollment: An

Analysis of Student Outcomes in Two States." St. Paul, MN: National Research Center for Career and Technical Education, University of Minnesota.
Keng, Leslie, and Barbara G. Dodd. 2008. "A Comparison of College Performances of AP and Non-AP Student Groups in 10 Subject Areas." College Board Research Report 2008-7. New York, NY: The College Board.
Kirkebøen, Lars, Edwin Leuven, and Magne Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *Quarterly Journal of Economics* 131(3):1057–111.
Kosfeld, Michael, and Susanne Neckermann. 2011. "Getting More Work for Nothing? Symbolic Awards and Worker Performance." *American Economic Journal: Microeconomics* 3(3):86–99.
Long, Mark C., Dylan Conger, and Patrice Iatarola. 2012. "Effects of High School Course-Taking on Secondary and Postsecondary Success." *American Educational Research Journal* 49(2):285–322.
Long, Mark C., Dylan Conger, and Raymond McGhee. 2014. "Evaluation of the Sustainability and Effectiveness of Inquiry-Based Advanced Placement Science Courses: Evidence from an In-Depth Formative Evaluation & Randomized Controlled Study." Arlington, VA: National Science Foundation.
Long, Mark C., Dan Goldhaber, and Nick Huntington-Klein. 2015. "Do Completed College Majors Respond to Changes in Wages?" *Economics of Education Review* 49:1–14.
Mattern, Krista D., Jessica P. Marini, and Emily J. Shaw. 2013. "Are AP Students More Likely to Graduate from College on Time?" College Board Research Report 2013-5. New York, NY: The College Board.
Mattern, Krista D., Emily J. Shaw, and Maureen Ewing. 2011. "Advanced Placement Exam Participation: Is AP Exam Participation and Performance Related to Choice of College Major?" College Board Research Report 2011-6. New York, NY: The College Board.
McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2):698–714.
Morgan, Stephen L., Dafna Gelbgiser, and Kim A. Weeden. 2013. "Feeding the Pipeline: Gender, Occupational Plans, and College Major Selection." *Social Science Research* 42(4):989–1005.
Murphy, Daniel, and Barbara Dodd. 2009. "A Comparison of College Performance of Matched AP and Non-AP Student Groups." College Board Research Report 2009-6. New York, NY: The College Board.
Oreopoulos, Philip, and Ryan Dunn. 2013. "Information and College Access: Evidence from a Randomized Field Experiment." *The Scandinavian Journal of Economics* 115(1):3–26.
Ost, Ben. 2010. "The Role of Peers and Grades in Determining Major Persistence in the Sciences." *Economics of Education Review* 29(6):923–34.
Papay, John P., Richard J. Murnane, and John B. Willett. 2016. "The Impact of Test-Score Labels on Human-Capital Investment Decisions." *Journal of Human Resources* 51(2):357–88.
Papay, John P., John B. Willett, and Richard J. Murnane. 2011. "Extending the Regression-Discontinuity Approach to Multiple Assignment Variables." *Journal of Econometrics* 161(2):203–7.
Patterson, Brian F., and Maureen Ewing. 2013. "Validating the Use of AP Exam Scores for College Course Placement" Research Rep. 2013-2. New York, NY: The College Board.
Shu, Pian. 2013. "Are the 'Best and Brightest' Going into Finance? Career Choice and Skill Development of MIT Graduates." Cambridge, MA: Harvard Business School.
Smith, Jonathan, Michael Hurwitz, and Christopher Avery. 2017. "Giving College Credit Where It Is Due: Advanced Placement Exam Scores and College Outcomes." *Journal of Labor Economics* 35(1):67–147.
Stange, Kevin M. 2015. "Differential Pricing in Undergraduate Education: Effects on Degree Production by Field." *Journal of Policy Analysis and Management* 34(1):107–35.
Steele, Claude M., and Joshua Aronson. 1995. "Stereotype Threat and the Intellectual Test Performance of African Americans." *Journal of Personality and Social Psychology* 69(5):797–811.

Stinebrickner, Ralph, and Todd R. Stinebrickner. 2012. "Learning about Academic Ability and the College Dropout Decision." *Journal of Labor Economics* 30(4):707–48.

Stinebrickner, Ralph, and Todd R. Stinebrickner. 2014. "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout." *Review of Economic Studies* 81(1):426–72.

Tai, Robert H., Christine Qi Liu, John T. Almarode, and Xitao Fan. 2010. "Advanced Placement Course Enrollment and Long-Range Educational Outcomes." In *AP: A Critical Examination of the Advance Placement Program*, ed. Phillip M. Sadler, Gerhard Sonnert, Robert H. Tai, and Kristin Klopfenstein, 109–18. Cambridge, MA: Harvard Education Press.

Theokas, Christina, and Reid Saaris. 2013. "Finding America's Missing AP and IB Students." Washington, DC: The Education Trust.

Turner, Sarah E., and William G. Bowen. 1999. "Choice of Major: The Changing (Unchanging) Gender Gap." *Industrial and Labor Relations Review* 59(2):289–313.

Walton, Gregory M., and Geoffrey L. Cohen. 2011. "A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students." *Science* 331(6023):1447–51.

Wiswall, Matthew, and Basit Zafar. 2015a. "Determinants of College Major Choice: Identification using an Information Experiment." *Review of Economic Studies* 82:791–824.

Wiswall, Matthew, and Basit Zafar. 2015b. "How Do College Students Respond to Public Information about Earnings?" *Journal of Human Capital* 9(2):117–69.

Zafar, Basit. 2015. "How Do College Students Form Expectations?" *Journal of Labor Economics* 29(2):301–48.