Fischer Christian (Orcid ID: 0000-0003-0856-6407)
Nguyen Ha (Orcid ID: 0000-0001-7138-1427)

Advanced Placement Course Credit and Undergraduate Student Success in Gateway Science Courses

Christian Fischer,[1,*], Eben Witherspoon,[2] Ha Nguyen,[3] Yanan Feng,[4] Stefano Fiorini,[4] Paulette Vincent-Ruz,[5] Chris Mead,[6] William Bork,[7] Rebecca L. Matz,[8] Christian Schunn[2]

[1] *Hector Research Institute of Education Sciences and Psychology, University of Tübingen, 72072 Tübingen, Germany.*

[2] *Learning Research & Development Center, University of Pittsburgh, Pittsburgh, PA, 15260.*

[3] *School of Education, University of California, Irvine, Irvine, CA 92697.*

[4] *College of Arts and Sciences, Indiana University, Bloomington, IN 47405.*

[5] *College of Literature, Science, and the Arts, University of Michigan, Ann Arbor, MI 48109.*

[6] *School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287.*

[7] *Hub for Innovation in Learning and Technology, Michigan State University, East Lansing, MI 48824.*

[8] *Center for Academic Innovation, University of Michigan, Ann Arbor, MI 48104.*

*Corresponding author: christian.fischer@uni-tuebingen.de

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/tea.21799

Acknowledgments: This work received support from Sloan Equity and Inclusion in STEM Introductory Courses (SEISMIC) project funded by the Alfred P. Sloan Foundation and participating institutions. The SEISMIC website is https://www.seismicproject.org. SEISMIC is managed at the University of Michigan for the participating institutions which include Arizona State University, Indiana University, Michigan State University, Purdue University, University of California Davis, University of California Irvine, University of California Santa Barbara, University of Michigan, University of Minnesota, and University of Pittsburgh. The authors would like to thank the following people for their support of this research: members of the larger SEISMIC research project (Timothy McKay and Nita Kedharnath), research team members contributing to early iterations of this work (Mark Umbricht, Yasemin Kalendar, Kyle Whitcomb, and Peter Radcliffe), as well as the many educational stakeholders and administrators involved in the collection, preparation, and curation of the institutional data sets used in this project. The views contained in this article are those of the authors, and not their institutions, the Alfred P. Sloan Foundation, or the College Board.

Abstract: Approximately two million students take Advanced Placement (AP) examinations annually. However, departmental policies that allow students to replace introductory courses with AP credit greatly vary within and across universities, even across relatively similar universities. This study examines the impact of AP credit policies on second course success in introductory Biology, Chemistry, and Physics course sequences at six large public research universities (N = 48,230 students). Applying logistic regression analyses, we found that general performance indicators and measures of systemic inequities during high school were predictive of students earning skip-eligible scores. Interestingly, we descriptively discovered wide variation across institutions and disciplines in the percentage of students who chose to skip when meeting their local policies. However, logistic regression analysis did not find general trends of inequalities for skipping courses for students with skip-eligible scores. We then applied inverse-probability weights with regression adjustment to examine the effects of course skipping. We found that students who skipped their first course with AP credit actually performed similarly well or better in subsequent courses than did students who did not skip, even in contexts where lower AP scores were accepted. Therefore, our study suggests administrators to increase coherence in AP credit policies across departments and disciplines. In particular, we encourage modifying AP credit policies and academic advising to encourage students to skip when they have eligible AP scores. Benefits may include reductions in unnecessary coursework that is burdensome for both students and universities, potentially freeing up resources to better support students who were not privileged to enroll in AP courses during their high school education.

Keywords: AP Program, Educational Policy, Gateway Courses, Higher Education, Science Education, Student Success

Success in postsecondary education and pathways into the labor market are uncertain for many students in science, technology, engineering, and mathematics (STEM) fields (Allen-Ramdial & Campbell, 2014; Olson & Riordan, 2012). About half of all students seeking bachelor's degrees in STEM between 2003 and 2009 left the STEM fields by 2009 (Chen, 2013). In particular, students historically marginalized in college environments are at risk of leaving

STEM fields (Griffith, 2010). Efforts to address this problem often center individual-level approaches like improving students' learning experiences and their college-readiness (Maltese & Tai, 2011; Means et al., 2016), with the Advanced Placement (AP) program being viewed as a college readiness indicator for student success in gateway courses (Cromwell et al., 2013).

The College Board's AP program is a large and complex system with many intended and unintended effects. In 2020, approximately 2.6 million students completed more than 4.7 million AP examinations (The College Board, 2020b). AP courses could provide rigorous, and inexpensive college-level coursework to high school students, especially if fully supported by appropriate and improved training and discipline alignment of AP instructors (Burnett & Burkander, 2021). The intention is that university departments should allow students to replace introductory college courses with AP credit if the students obtain a passing score (i.e., 3 or above; Cromwell et al., 2013), thereby accelerating time-to-degree, increasing retention and course concentration in STEM, and reducing the costs of college (Gurantz, 2021). That said, AP credit policies substantially vary across universities and even by department within universities (Duffy, 2010; Jagesic & Wyatt, 2018; Wyatt et al., 2018), sometimes requiring higher minimum thresholds of 4 or 5, and sometimes not accepting the equivalence at all. Some of these policies and variation in how students take advantage of AP credit could further exacerbate systemic inequities, especially given concerns about access to AP coursework (Kanno & Kangas, 2014; Price, 2020). However, structural-level policies and systemic inequities of awarding college-level credit are seldomly studied. Hence, independent studies evaluating structural-level policies and systemic inequities of awarding college-level credit, both topics seldomly studied with a comparative approach, can reveal untapped opportunities offered by the AP program.

Background

Impacts of the AP Program

Several studies have examined associations of students' AP program participation with college-level outcomes. For instance, Chajewski and colleagues (2011) found an association of AP participation with higher rates of college enrollment. When examining introductory college course performance, students who earn passing AP scores tend to perform better compared to students who did not take AP examinations (Sadler & Sonnert, 2010; Sadler & Tai, 2007). Similarly, AP course-taking and AP exam performance are associated with higher first-year college GPA, higher graduation rate, and shorter time-to-degree (Ackerman et al., 2013; Evans, 2019; Scott et al., 2010; Smith et al., 2017). AP participation and success are also associated with college major choices and career interests (Avery et al., 2016; Evans, 2019; Warne et al., 2019). Notably, the higher a student's score on an AP exam, the greater the probability of majoring in the corresponding subject (Avery et al., 2016; Gurantz, 2021).

Especially relevant to awarding credit for AP courses is performance in later courses; even when students have earned passing scores on AP exams, university faculty and administrators often worry that students will struggle in later courses if they do not experience the equivalent gateway courses at the given university. Research on the impacts of AP credit policies with subsequent course success is, however, currently limited to four non-peer-reviewed College Board reports, described below.

Wyatt and colleagues (2018) examined 14 AP courses for students at 93 colleges. Mean comparisons without covariates indicated that students who earned AP credit and skipped the introductory college courses tended to score higher in subsequent courses across all subject areas compared to students who did not take the AP exam but completed the respective introductory college courses. Jagesic and Wyatt (2018) conducted a complementary study examining 10 other

AP exams across 50 colleges. Using the same mean comparison approach, students who earned AP credit and skipped the introductory college course tended to score higher in subsequent courses across most subject areas compared to students who did not take the AP exam but completed the respective introductory college courses. Morgan and Klaric (2007) examined data of 10 AP exams from over 70,000 first-year undergraduate students at 27 colleges using a regression analysis with students' SAT score as the sole covariate. These results indicated that students who earned AP credit had significantly higher course grades in subsequent courses for seven out of the ten examined AP exams compared to students who did not take the AP exam but completed the respective introductory college courses. Earning a 4 was usually sufficient to show a positive effect, and sometimes even a 3 showed a positive effect. However, this study did not use high school performance as a covariate, which is a major limitation as it represents a stronger predictor of college performance than SAT scores (Galla et al., 2019). Another study explored data from a single university using simple regression controls of SAT scores and high school rank (Keng & Dodd, 2008). This study found that students who earned AP credit generally did better in later same-subject coursework for 9 of the 10 AP exams explored. However, when focusing on two-course gateway sequences, students in Biology did worse in the second course when compared with students who took the first gateway course at the university, students in English Composition performed equally well, and students in Calculus did somewhat better in the subsequent gateway course. Together, the studies present a mixed story about the impact of AP credit on subsequent course performance.

Educational Equity and the AP Program

Although studies have demonstrated potential benefits of AP program participation for students' college careers, differential access to quality AP programs can exacerbate unequal learning opportunities to students (Price, 2020). Studies have shown that access to AP courses across the nation and performance on AP examinations are stratified by race and class (Kanno & Kangas, 2014; Klugman, 2013; Schneider, 2009), and as such can be related to racial segregation and income inequality (Card & Rothstein, 2007; Tienken, 2012). Further, students with AP credit are more likely to be accepted into elite institutions and are more likely to receive grants from the university (e.g., Lin et al., 2020). Inequitable learning opportunities in AP courses are often attributed to shortages of qualified and experienced teachers, and an overall lower quality of instruction in schools with higher percentages of historically underserved and marginalized students (Hallett & Venegas, 2011; Kyburg et al., 2007; Taliaferro & DeCuir-Gunby, 2008).

In response, the College Board has engaged in efforts to mitigate these inequalities, for instance, by increasing access to the AP program (Conger et al., 2009; Roegman & Hatch, 2016). Moreover, many school systems and several states have allocated substantial funds to pay students' test fees to further improve access (The College Board, 2020a). Further, in terms of potential benefits to students, the fiscal burden of unnecessary course-taking in college will be felt most severely by students from lower-income families, and university policies which broadly discount AP credit could disproportionately disadvantage students from low-income families. Therefore, to better shape policies that impact historically underserved students who do currently earn AP credits and also to shape future policies in order to expand access to AP, this study examines key aspects of advanced placement decisions for and by students.

Research Questions

Institutions of higher education become the gatekeepers of the potential benefits that earning a successful AP score affords students. This is affected by a combination of effective and fair credit policies, and student advising. This study represents the first independent and crossinstitutional effort to examine the impact of replacing introductory science courses with AP credit on subsequent course performance applying advanced statistical research methods. As equity is a central concern in undergraduate science policies, we also carefully examine who currently tends to have qualifying AP scores and who has tended to claim AP credit when eligible, a potentially important but previously unexamined topic. This research is carried out with students attending public research institutions in STEM programs, where inequities in access may be quite different from inequities within national samples because of the non-representativeness of who attends those institutions. The goal is to provide policy-relevant research findings for those kinds of institutions while also providing a model for how to conduct analyses to shape policies at other institutions. Specifically, this study is guided by three research questions (RQs) focused on students enrolled in selected introductory science course sequences:

RQ1: What student characteristics, within selective public research universities, are associated with having earned AP scores that qualify for skipping introductory science courses?

RQ2: What student characteristics, within selective public research universities, are associated with skipping introductory science courses for students who have earned skip-eligible AP scores?

RQ3: What are the impacts of replacing an introductory course with AP course credit on subsequent science course grades?

Methodology

Data Sources and Sample

This study was conducted at six large public research universities in the United States. We focused on AP Biology, AP Chemistry, and AP Physics C: Mechanics as their corresponding college-level courses are central to most undergraduate STEM degrees. All universities are members of the SEISMIC collaboration, a multi-university project aimed at improving equity and inclusion in STEM (SEISMIC Collaboration, 2022). More details on the context of the institutions participating in this study is provided in Table A1 in the supplementary materials. Notably, this study only included years after the redesign of the AP science curricula in their respective disciplines (AP Biology in 2013, AP Chemistry in 2014, and AP Physics in 2015) so that our findings are more relevant to current policy makers. Further, we focused our analyses on the first two courses in each course series irrespective of whether institutions were on a quarter or semester schedule (Biology: Cell Biology and Evolution and Ecology; Chemistry: General Chemistry I and II; and Physics: Force and Motion, and Electricity and Magnetism). The course sequences at each institution are described in more detail in Table A2 in the supplementary materials.

Data was provided at each of the six institutions from their respective on-campus services that collect and curate institutional data including the Offices of Admissions and Institutional Research and the Registrar's Offices. Notably, each institution had to conduct the analyses at their respective campus because raw data at the detailed level required by our analytic approach could not be shared across institutions due to FERPA regulations. This study only included degree-seeking students on their first attempt taking courses in these three large science gateway course series. We excluded visiting students and those who did not attend a U.S. high school as they were unlikely to have had AP experiences. We also excluded transfer students as they were unlikely to have participated in these science course sequences. This led to a full analytical sample of 48,230 students within 15 course sequences across the six institutions.

Measures

The core *dependent variables* indicated whether a student earned an AP score that allowed for skipping the first course in the course sequence (RQ1), whether a skip-eligible student actually skipped the first course in the course sequence (RQ2), and a continuous measure describing second course performance on the typical 0–4 grade scale (RQ3).

Independent variables for RQs 1 and 2 included dichotomous variables indicating students' self-reported sex (1: female, 0: male), first-generation college student status (1: neither parent/guardian holds a bachelor's degree, 0: at least one parent/guardian holds a bachelor's degree), and low-income status (1: flagged as low-income based on either family household income and household size using 185% of the U.S. poverty line or Pell grant eligibility; 0: not flagged as low-income). We also included a categorical variable representing students' self-reported racial/ethnical background (White: student only self-identifies as White; Racially marginalized students (RMS): student's self-identity includes Latino or Hispanic, Black or African American, American Indian, Alaska Native, or Pacific Islander Students; Asian: students' self-identity includes Asian or Asian American). In the case of multiple racial/ethnic backgrounds, White/Asian students were grouped with Asian students, and all other combinations were grouped with RMS given how race/ethnicity is commonly experienced in the U.S. (Asai, 2020). In addition, we included continuous variables that capture prior academic achievement in high school including students' high school GPA, mathematics proficiency (highest score on math ACT/SAT exam; ACT scores were converted to SAT scores using common concordance tables), and English proficiency (highest score on Reading ACT/SAT exam). The year of the course was included as a covariate to remove variance related to features groups of students in those courses that year (e.g., grade curves).

For RQ3, the core independent variable represented a dichotomous variable indicating whether a student skipped the first course in the course sequence. The independent variables described for RQs 1 and 2 served in RQ3 as covariates and matching variables for creating propensity score measures.

Table 1 provides descriptive information for all variables at each institution. Descriptive details on demographic characteristics in this study are described in Table A3 in the supplementary materials. Missingness across variables was very low (i.e., usually below 10%) and assumed to be completely random. To verify that potential patterns of missingness do not affect the outcomes, we examined correlations between our demographic variables, cohort, and an indicator variable of missingness on our outcome variable. These analyses showed no significant or meaningful associations between these variables and missingness (e.g., all correlation coefficients were $r = .06$ or below). Therefore, we conducted all analyses using listwise deletion instead of the more complex missing data strategies such as Multiple Imputations (MI) or Full Information Maximum Likelihood (FIML; Graham, 2009).

[Table 1 about here]

Statistical Analysis

For RQ1 (i.e., to identify whether different groups of students were more or less likely to be skip- eligible), we conducted simple bi-serial correlations between demographic variables and the binary variable that indicated whether the student earned a skip-eligible score at their respective institution. Afterwards, we conducted multiple logistic regression models focused upon predicting skip-eligibility using high school performance and achievement variables (keeping the same demographic variables as control variables). These models used the following reduced form equation:

$$Y(skip\_eligible)_i = \alpha + \gamma X_i + \mu$$

Y(skip_eligible) is the probability that student i was skip eligible. Xi represents a vector of student variables, including gender, racial/ethnic background, first-generation student status, low-income status, standardized high school GPA, standardized Math and English ACT scores, and the cohort (year of class taking). α is the intercept and μ represents the error terms, respectively. Notably, these logistic regression analyses were only conducted for institutions and disciplines with sufficient variation on the dependent variable (i.e., at least 100 skip-eligible students), which led to the exclusion of institutions for Biology (excluding Institutions E, F), Chemistry (excluding Institution F), and Physics (excluding Institution A). As a reminder, each institution conducted these analyses separately with their own data. Therefore, we applied meta-regression techniques for both the bi-serial correlations and the logistic regression analyses to synthesize the extent to which there was sufficient evidence for a consistent relationship across institutions, similar to synthesizing findings across studies. Meta-regressions were conducted using the following form equation:

$$Y(skip\_eligible)_i = \theta + \gamma X_i + \epsilon_i + \zeta_i$$

The meta-regression equation introduced two additional error terms. ϵ_i indicates the sampling error where the estimates of an institution deviate from their true estimates. ζ_i indicates a random effect, denoting that the estimates are sampled from an overall distribution.

For RQ2 (i.e., were those same demographic and performance variables related to students' propensity to skip the course once they had earned an eligible score), we limited the dataset to students who earned a skip-eligible score at their corresponding institution and discipline. We then conducted multiple logistic regression analyses predicting skipping when eligible. These models also controlled for the students’ cohort (i.e., matriculation year) to remove the variance that was correlated with the particular group of students matriculating across years. In particular, we used the following reduced form equation:

$$Y(skipped\_course)i = \alpha + \gamma Xi + \mu$$

Y(skipped_course) is the probability that student i chose to skip. Xi represents a vector of student covariates, including gender, racial/ethnic background, first-generation student status, low-income status, standardized high school GPA, standardized Math and English ACT scores, and the cohort (year of course enrollment). α is the intercept and μ represents the error terms, respectively. Again, logistic regression analyses were only conducted for institutions and disciplines with sufficient variation on the dependent variable, which led to the exclusion of three institutions for Biology (only included Institutions A, C, D), three institutions for Chemistry (only included Institutions A, D, E), and five institutions for Physics (only included Institution D), respectively. Similarly, we used meta-regression techniques to synthesize estimates of each relationship across the findings from each institution.

For RQ3 (i.e., what is the effect of skipping on course grade in the second course), we applied a propensity score weighting approach, in which the model “matches/weights” students on propensity to receive the “treatment” (in this case, “matching” students on likelihood of skipping the first course) and then controls for that likelihood to better recover the causal effect of the treatment (in this case, skipping the first course). Moving beyond simple multiple regression models, researchers have used propensity score matching and propensity score weighting approaches to reduce sample biases and provide more robust estimates in non-experimental settings where there may exist confounding differences between the comparison groups (Dehejia & Wahba, 2002). Notably, our specific analytic approach involves both inverse probability weights and a “doubly robust estimation” (Funk et al., 2011; Li, Kleinman, &

Gillman, 2014) to estimate the causal effect of skipping the introductory course (Austin, 2011; Funk et al., 2011). Weighing approaches have the advantage of using more data compared to traditional propensity score matching approaches as they tend to only match identical cases across treatment and control groups. To implement this approach, we first calculated the propensity score e in a logistic regression model:

$$e = P(Z = \text{skipped\_course} | Xi)$$

Here, Xi represents a vector of student covariates, including gender, racial/ethnic background, first-generation student status, low-income status, standardized high school GPA, standardized Math and English ACT scores, and the cohort (year of course enrollment). The inverse probability of treatment weight, or the probability of having skipped the courses, is then used as weights in subsequent regression models to determine the effects of skipping on course grade. Intuitively speaking, inverse probability weighting assigns more weight to individuals who skipped courses, although they would have been more likely to not have skipped. The models controlled for demographics and prior performance using the following equation form:

$$Y(\text{course\_grade})i = \alpha + \square(\text{skipped\_course})i + \gamma Xi + \mu$$

Y(course_grade)i is the outcome variable, the standardized grade students received in the subsequent course. Skipped_course is the independent variable that indicates whether students skipped the introductory course (skipped_course = 1) or not (skipped_course = 0). Subjects with skipped_course = 1 received weight 1/e; subjects with skipped course = 0 received weight 1/(1-e). Xi represents a vector of student covariates, including gender, racial/ethnic background, first-generation student status, low-income status, standardized high school GPA, standardized Math and English ACT scores, and the cohort (year of course enrollment). α is the intercept and μ represents the error terms, respectively. Before moving forward with subsequent analyses, we assessed the balance in the matched samples. We examined the sample sizes and the balance for the covariates to ensure that they meet common balance criteria (setting standardized mean differences at the .05 threshold; What Works Clearinghouse, 2016).

Using this inverse probability treatment weighting approach, we determined the effects of skipping on course grade, controlling for demographics and prior performance for two samples: (a) the full sample of students and (b) students with skip-eligible AP scores. Notably, we only conducted analyses for institutions and disciplines that exhibited sufficient variation on the matching term (i.e., fewer than 90% of those who were eligible to skip actually skipped the first course), which led to an exclusion of three institutions for Biology (only included Institutions A, C, D), two institutions for Chemistry (only included Institutions A, D, E), and four institutions for Physics (only included Institutions B, D), respectively.

Afterwards, we conducted meta-regression analyses to synthesize the findings across institutions. The meta-regression applied to propensity scores was similar to the regression model equation, but included additional error terms for sampling error and random effect. Notably, we did not apply multilevel models that account for the influences students might have on one another within specific course sections. Although such an approach is often used in analyses of student performance data, the need for such an approach in large lecture courses is less clear. Further diagnostic assessments using intraclass correlation coefficients (ICCs), found the effects of course sections were below 0.1, which is commonly considered a minimal threshold for conducting multilevel modeling (Field, 2005). At lower ICC levels, the more complex multilevel models produce identical findings.

Results

Student Participation and Success in the AP Program (RQ1)

Student participation in the three AP exams varied widely across institutions within each discipline. The percent of students having AP experiences varied from 16% to 54% in Biology, 14% to 40% in Chemistry, and 6% to 32% in Physics (see Table 1). Overall, having AP experiences ranged from 1 in 16 students to over half of students. Similarly, departmental policies awarding AP credit to replace introductory courses varied greatly across institutions and disciplines (see Table 2). Out of the six institutions included in our study, three had AP credit policies allowing students to skip the first course in all three of the Biology, Chemistry, and Physics courses sequences, and the other three institutions allowed students to skip courses in two of the three studied disciplines. Even for institutions awarding AP credit, the corresponding thresholds varied across disciplines, with three institutions awarding credit for a 3 on an AP exam (one institution for Biology, two institutions for Chemistry, zero institutions for Physics). All six institutions awarded AP credit for a 4 on at least one AP exam (four institutions for Biology, two institutions for Chemistry, two institutions for Physics) and three institutions only for a 5 on one AP exam (zero institutions for Biology, one institution for Chemistry, three institutions for Physics). In sum, even for a relatively homogeneous set of large public research institutions, we find no agreement on departmental AP policies.

[Table 2 about here]

Examining the simple correlations with demographic factors revealed that first-generation college student status, low-income status, and being female tended to be negatively correlated with earning skip-eligible AP scores across disciplines and institutions. Figure 1 shows these correlations graphically along with results of the meta-regressions applied to these correlations. Interestingly, there are no significant correlations between racial/ethnic minority status and students' having skip-eligible AP scores for Chemistry and Physics, and this correlation tends to be negative for Biology across institutions. Importantly, however, the estimated demographic correlations are small (within 0.1 of zero) for all demographic variables and all subjects. A correlation of -0.1 means only 1% of the variance is explained by the demographic variable. The strength of the correlation appeared to be unaffected by institutional cut-offs for accepting AP credit. Having small-sized effects in the relationships between AP performance and demographic factors in university populations is consistent with prior research regarding the localization of inequities which appeared to emerge upstream (Price, 2020). For example, it appears that the national sample's inequities of AP enrollments (e.g., large gender differences in AP Physics enrollments) manifests primarily at earlier steps: in who enrolls at these institutions and in who follows science-related pathways.

[Figure 1 about here]

Academic performance predictors of high school GPA (HS GPA) and SAT performance were often statistically significant predictive of earning skip-eligible AP scores. The meta-regressions coefficients displayed in Figure 2, which control for demographic variables to estimate the unique associations with HS GPA and SAT, indicate that HS GPA is a small positive predictor of having eligible AP scores (in particular for AP Chemistry), SAT Reading is a small predictor (particularly for AP Biology), and SAT Math varies from being a small (AP Biology) to large predictor (AP Chemistry and AP Physics). The one salient exception to the predictiveness of SAT Math comes from two institutions with non-predictive SAT math scores. Notably, one of these institutions had relatively few students meeting the sample size threshold as they required a five on the AP exam.

[Figure 2 about here]

Table 3 presents all meta-regression coefficients for RQ1. Overall, we find that having eligible AP credit scores is somewhat confounded with general academic performance and various systemic inequities apparent during the K-12 years. Thus, these variables need to be accounted for when examining the impact of skipping on subsequent course performance. It is also important to note that there was often statistically significant institutional variation in the relative predictiveness of each predictor (i.e., the point variation by institution in Figure 2 is often not just statistical noise). Thus, it was important to build separate prediction models by institution, both for addressing RQ1 and for accounting for confounding variables in RQ3.

[Table 3 about here]

Course Skipping Patterns (RQ2)

Figure 3 presents the percent of students choosing to skip when eligible on the Y-axis. Across courses and institutions, there was huge variation with a somewhat bi-modal distribution, with some courses having very few students choosing to skip (i.e., <10%) and some courses having 80% to >99% choosing to skip when eligible. On the x-axis is the percent of students eligible to skip given the populations involved and the departmental policies in question. There was no relationship between the percentage of skip-eligibility and the percentage of choosing to skip. These patterns suggest not only variation in departmental policies allowing students to earn AP credit but also in the institutional knowledge and (informal or formal) academic advising available to students in applying their AP credits towards their degree progression. Often, the majority of students, especially within the Biology sequence, seem to have received advice to enroll in courses they could have skipped. This is potentially an inequity problem as the lack of consistent guidelines can allow biases in informal networks or individual advisors to have a strong effect on student decision-making processes. This issue is especially important given research showing systemic problems with college advising (Zhang, 2016).

[Figure 3 about here]

Figure 4 shows the odd ratios from the logistic regressions predicting students' choice to skip using demographic and academic resource variables (see also RQ2 sections of Table 2). While in a few cases for a particular institution for a particular course, there appeared to some predictive relationships with demographic variables, meta-regression results (gray squares with 95% CIs) found that none of the variation across institutions for a predictor was statistically significant (i.e., the institutional variation could just be noise). The average effect was significantly different from zero in only one case: in Biology, first-generation students were more likely to skip when eligible. Note that the confidence intervals were wide for these models because of the relatively low proportions of key student groups in these institutions combined with the bimodal/extreme pattern in student choices: it was either the case that relatively few students skipped or relatively few students did not skip when eligible.

HS GPA, SAT Math, and SAT Reading performance also did not significantly predict course skipping across disciplines either, with SAT Reading representing the only exception in Physics. Overall, we did not find general trends of inequalities for skipping courses for students with skip-eligible scores. Combined with the small demographic effects on having skip-eligible scores at these institutions, if skip eligibility thresholds were lowered at these institutions, we expect that equal proportions of students by demographic variables would take advantage of the increased skip opportunities.

[Figure 4 about here]

Impact of Skipping Introductory Science Courses (RQ3)

Figure 5 presents the estimated effects of course skipping on performance in the second course. When examining all students enrolled in the selected science gateway course series (Figure 5, top panel), we find that performance in the second Biology gateway course is nearly identical between those who skip and those who do not, and, for Chemistry and Physics, students who skipped actually outperformed students who took the first course with a moderate-to-large effect size. Also, there is no indication of significant heterogeneity across the examined institutions within each discipline (see RQ3 section of Table 3).

While the analytical technique controlled for many potentially confounding factors, it could not directly control for depth of interest and experience in the discipline as the full sample analysis could not include the presence of an AP experience or performance on an AP exam. The analyses that focused on skip-eligible students (Figure 5, bottom panel) controlled for these confounders more directly because it only included students with advanced disciplinary experiences in high school by directly matching on their AP performance. In support of the idea that the full sample had confounding variables with skipping that were not properly controlled, the estimates of the effect of skipping using the reduced sample were generally less positive. However, even in this more conservative estimate, the meta-regressions indicate no overall significant differences in subsequent course performance between students who skipped the first course and those who did not across all three disciplines (see RQ3 section of Table 3). There was statistically significant heterogeneity in this effect estimate across institutions for the case of Biology. Specifically, students skipping the first gateway Biology course tended to perform worse in the second course, particularly for a cutoff score of 3, driving this observed heterogeneity.

[Figure 5 about here]

Discussion

This large-scale study offers unique insights into the higher education landscape within large public research institutions. It represents the first independent, peer-reviewed study that uses advanced quantitative research methodologies to examine a common AP credit policy – offering students AP credit to replace introductory college courses. Policy decisions can have complex and unintended consequences for students overall or for groups of students, and thus it is important for policy makers to be informed about potential factors that might exacerbate inequities in the system.

Variation in Department Policies and Student Skip Choices

Variations in choice were particularly salient at the department and student levels. At the department level, there was huge variation in what was accepted (if anything). Here we consider explanations we have heard from colleagues while conducting this research. First, university selectivity is an unlikely explanation since there were many instances in which a pair of universities had one university being more restrictive in one course but less restrictive in another course. Second, tracking could offer an explanation, where sometimes the same course is offered to all majors in one university whereas a different course version is offered to different tracks (e.g., Biology for Biology majors vs. Biology for other students); departments may be more conservative in requiring their own course versions when it is also taken by their own majors because the foundational knowledge will be built upon in a wide range of future courses. Third, physical resource limitations (number of large lecture halls or large lecture halls with specialized demonstration facilities) might offer cause for variation: less restrictive skipping policies may reflect a need to have lower enrollments. Fourth, funding could also matter; more restrictive skipping policies could reflect a need to create more positions for graduate students to serve as teaching assistants or a need to increase departmental seat counts which shape departmental funding levels. Finally, variation could reflect use of data-informed decision policies: some departments may have used theoretical rather than empirical processes to make their decisions, or they may have conducted their analyses based upon the prior versions of the AP courses, which were less likely to be equivalent in value as the revised AP courses. Here we note that many of these concerns are not focused upon student success.

Student Course Enrollment Choices

In terms of student choices, there was huge variation, ranging from some courses being skipped by most students when eligible to some courses being skipped by very few students when eligible. Such variation might come from advising policies, which in turn could be shaped by some of the same factors that shape formal course-skipping department policies. For example, advisors could be asked to encourage or discourage course skipping based upon resource implications. Second, variation might come from variation in level of concern about GPA; students planning to go to very selective graduate degrees (e.g., medical school) may want to retake courses they could have skipped to inflate their GPA, whereas students in professional undergraduate degrees like engineering may be less concerned about GPA. Third, variation may come from reputations regarding the quality in the teaching in the course sequences. If students believe the first course in the sequence is generally taught poorly or generally taught very well, this could have a large influence on their choices. Given the differences by first-generation status, income, and race/ethnicity that we observed in students earning skip eligible AP scores , it is important that future research examine these potential factors influencing student choice.

Impact of Course Skipping

In the current contexts, analyses critically showed only very small demographic differences in which students had eligible AP scores and effectively no demographic differences in terms of which students chose to skip courses when eligible for students enrolled in the examined institutions. However, it is important to emphasize that marginalized students that had access to quality AP instruction may have had very different experiences as compared to marginalized students with different economic or academic backgrounds, and therefore we must be careful about over-generalizing these results across all institutions or all students even within these institutions (Owen, 2021).

Before discussing the implications of these findings, we first note some important limitations that should be considered when interpreting these findings. Methodologically, propensity score weighting approaches can only adjust for observed features. Therefore, it may not account for hidden biases correlated with students who chose to skip a course they are eligible to skip (e.g., motivational variables such as interest in the discipline). That said, those motivational variables are correlated with academic performance variables included in the models and thus are not completely unaccounted for. Future research might use regression discontinuity designs if they have datasets with larger numbers of skip-eligible students and better balances between skippers and non-skippers to generate treatment effect estimations. Also, the nature of public research universities will likely limit the generalizability to all universities in the nation. Given the educational inequalities at the national level for access to the AP program (Ackerman et al., 2013; Fischer, Fishman, et al., 2020), it is important to examine whether the patterns hold for marginalized students at different types of institutions, including ones where overall student performance and retention rates are lower. University policies should, at the very minimum, not exacerbate inequalities that occur in the K-12 environment. We also emphasize the limitations that institutional data has in capturing important nuances in race and ethnic diversity, gender, or the educational contexts in which students grew up. There may be further dimensions of inequity that would emerge with data that allow for a more nuanced analysis.

Overall, the main finding that using AP credit to skip the first course does not adversely affect student performance in subsequent gateway courses has implications for educational stakeholders across the nation as millions of students take AP examinations every year (The College Board, 2020b). Importantly, when considering all students (including students who did not take AP examinations), students who skip tend to perform better on subsequent courses compared to students who did not take AP exams or took AP exams but did not skip. When restricting the analysis to skip-eligible students (they took the corresponding AP exam and earned a passing score), the effect of actually skipping the course is negligible but, importantly, students who skip do not seem to be disadvantaged. Overall, these findings indicate that students obtained relevant introductory disciplinary content knowledge through their AP courses with introductory college courses not offering additional benefits that translate in improved subsequent course performance. This is remarkable as institutions tend to brand their own introductory courses as of higher quality compared to high school level instruction.

The lack of benefits from retaking the first course at the university requires some discussion. How could it be that retaking the course did not produce at least small benefits for students, and what could underlie negative effects, if the less conservative full sample analyses present an accurate picture? From a memory perspective, more recent reinforcing of concepts should have produced benefits, especially considering that students could have taken the corresponding AP course as much as four years earlier. However, from a memory perspective, it is also important to point out that forgetting follows a power function (Anderson, 2000), with initially very rapid forgetting, followed by much slower forgetting. Thus, the difference in effects between 4 months of forgetting and 4 years of forgetting is not linear, and context cues can effectively retrieve memories from long ago (Altmann & Schunn, 2012). Second, the level of conceptual overlap between the two courses in these sequences is often relatively small; research examining relationships between performance in the first course predicting performance in the second course (after controlling for general academic performance) have found very weak relationships for Physics and Chemistry sequences in college (Whitcomb et al., 2020). For the Biology sequence, we observed both orders of the two courses' content areas, suggesting relatively little content dependence. Third, students may not put much effort into a course they are retaking and the learning experience of a large lecture class is often not ideal, so the relative amount of learning from the re-taking experience may be minimal. Finally, the experience of re-taking a course and then doing somewhat poorly may be very demotivating to students, and low self-efficacy for learning in that discipline may undermine future learning and performance in that discipline (Blatt et al., 2020).

Building upon this last point, it is important to note that we only examined the impact on an academic outcome measure but not on other important facets of college life in the first year including interest in the major, network building among peers, or mental health and stress. Retaking a challenging STEM course could increase first-year students' stress or decrease performance in other concurrent courses because the total workload might be increased.

While there was some meaningful heterogeneity in these second-course performance effects across disciplines and institutions, these differences seemed to matter only in Biology. This increased heterogeneity in Biology may be explained through the flexibility in course orders. Whereas the course series in general Chemistry and Physics are mostly stable in terms of what content is covered in the first versus second course, there is a lot more variation in Biology courses sequences. For example, sometimes cell biology and ecology content were in different course sequence order across institutions. As a result, the content knowledge that students may acquire in AP Biology may be differentially important for course performance depending on which content is being skipped. That said, the similarity in our findings is striking especially given the broad range of departmental AP policies and differences in institutional demographics and contexts represented in this study. Differences in AP policies may be explained through the nature of educational decision making on AP policies, for instance, in department meetings serving as have led to these diverse circumstances. With our findings, we encourage administrators to increase coherence in AP credit policies across departments and disciplines through better communication within and across their institutions and through use of data.

Importantly, our findings are consistent with those previously obtained (Jagesic & Wyatt, 2018; Keng & Dodd, 2008; Morgan & Klaric, 2007; Wyatt et al., 2018), which used simple correlation or regression methods containing very few control variables and using methods that are less able to recover causal relationships than propensity score approaches. This consistency suggests that our findings may also apply to the broader set of institutions examined in those prior studies. In other words, it may be generally the case that prior AP experiences allow students to perform satisfactorily in subsequent courses compared to students without comparable AP experience while also avoiding potential negative effects of not taking the first course in science course sequences.

Future Research

As large institutional data sets become increasingly available to universities to aid data-driven educational policy-making processes (Fischer, Pardos, et al., 2020), we encourage institutions to replicate these analyses at their campus to evaluate their AP policies and potential inequities. Further, we encourage institutions to use the methodological approach used in this study to examine other placement decisions affecting large numbers of students based on IB scores, courses taken at community colleges, and internal course placement exams. The GitHub repository of this study (https://github.com/JRST-APCredits/SharedCode) supplementing this article provides in-depth analysis scripts for researchers to aid this process. Included in the repository are methods for: (a) creating aligned variables from institutional data that typically differ in their collection and organization; (b) selecting and aligning courses sequences given the wide variety in alternative pathways institutions often offer (specific to colleges, majors, non-majors, or honors); (c) formally analyzing and representing patterns across institutions without transferring student-level data across institutions, and (d) attending to equity of credit access and use for course skipping.

Although the potential impacts of not taking the first course in a science course sequence are most immediate on the corresponding subsequent course, future research may also investigate impacts of AP credit policies on more distal college success factors including course taking patterns, graduation rates, terms-to-degree, and time-to-degree, as well as labor market outcomes (Fischer et al., 2021; Gurantz, 2021). Each course's specific policy on its own is unlikely to have large effects on the overall population, but collectively changing policies across an institution does have the potential to substantially change those large grain-size variables.

Practical Implications

Implications of this study for educational stakeholders are two-fold and the two parts should be jointly considered. First, departmental and university administrators should carefully review their formal and informal AP credit policies. If only high scores are accepted, the rationale for that decision should be empirically evaluated and potentially reconsidered. Similarly, if few students are choosing to skip courses when eligible to skip, the underlying rationale for those decisions should be investigated; if students are being encouraged to take the first course in the sequence (for instance through academic advising services), it should not be for fear of poor performance in the second course. Instead, it should be for other reasons like peer-network building through similarities in course schedules for the same cohort of freshman students in a given major.

Second, it is important for institutions to use resources to improve instruction for those students that had different high school experiences and are unable to skip introductory courses. If institutions only improve the support for those students that are able to skip, they will essentially create an overprivileged track. By allowing students to skip, institutions may be able to reduce class size, and increase both curricular resources and teaching resources to other students and create a more equitable benefit across the institution. In a similar vein, allowing students to skip can also have benefits to lower financial strain in students and lower curricular load.

All in all, our study encourages departments to continue with policies that award college credit for passing AP scores, or to explore the adoption of such policies as students do not seem negatively affected by skipping the first course in science course sequences.

References

Ackerman, P. L., Kanfer, R., & Calderwood, C. (2013). High School Advanced Placement and Student Performance in College: STEM Majors, Non-STEM Majors, and Gender Differences. *Teachers College Record*, *115*, 1–43.

Allen-Ramdial, S.-A. A., & Campbell, A. G. (2014). Reimagining the pipeline: Advancing STEM diversity, persistence, and success. *BioScience*, *64*(7), 612–618. https://doi.org/10.1093/biosci/biu076

Altmann, E. M. & Schunn, C. D. (2012). Decay versus Interference: A new look at an old interaction. *Psychological Science, 23*(11), 1435-1437.

Anderson, J. R. (2000). *Learning and memory: An integrated approach*. John Wiley & Sons Inc.

Asai, D. J. (2020). Race Matters. *Cell, 181*(4), 754–757. https://doi.org/10.1016/j.cell.2020.03.044

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Avery, C., Gurantz, O., Hurwitz, M., & Smith, J. (2016). *Shifting college majors in reponse to Advanced Placement exam scores* (NBER Working Paper No. 22841). National Bureau of Economic Research.

Blatt, L. R., Schunn, C. D., Votruba-Drzal, E., & Rottman, B. M. (2020). Variation in which key motivational and academic resources relate to academic performance disparities across introductory college courses. *International Journal of STEM Education, 7*, 58.

Burnett, A., & Burkander, P. (2021). *Advanced Placement participation, staffing, and staff training in the District of Columbia Public Schools (REL 2021–077).* U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. http://ies.ed.gov/ncee/edlabs.

Card, D., & Rothstein, J. (2007). Racial segregation and the black–white test score gap. *Journal of Public Economics*, *91*(11–12), 2158–2184. https://doi.org/10.1016/j.jpubeco.2007.03.006

Chajewski, M., Mattern, K. D., & Shaw, E. J. (2011). Examining the role of Advanced Placement exam participation in 4-year college enrollment. *Educational Measurement: Issues and Practice*, *30*(4), 16–27.

Chen, X. (2013). *STEM attrition: College students' paths into and out of STEM fields* (Statistical Analysis Report NCES 2014-001). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Conger, D., Long, M. C., & Iatarola, P. (2009). Explaining race, poverty, and gender disparities in advanced course-taking. *Journal of Policy Analysis and Management*, *28*(4), 555–576. https://doi.org/10.1002/pam.20455

Cromwell, A., McClarty, K. L., & Larson, S. J. (2013). College readiness indicators. *Bulletin*, *25*, 1–8.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics*, *84*(1), 151–161.

Duffy, W. R. (2010). *Persistence and performance at a four-year university. The relationship with advanced coursework during high school* (P. M. Sadler, G. Sonnert, R. H. Tai, & K. Klopfenstein, Eds.; pp. 139–163). Harvard Education Press.

Evans, B. J. (2019). How college students use Advanced Placement credit. *American Educational Research Journal*, *56*(3), 925–954. https://doi.org/10.3102/0002831218807428

Field, A. P. (2005). Intraclass correlation. In B. Everitt & D. C. Howell (eds.), Encyclopedia of Statistics in Behavioral Science (Vol. 2, pp. 948-954). New York: Wiley.

Fischer, C., Baker, R., Li, Q., Orona, G., & Warschauer, M. (2021). Increasing success in higher education: The relationships of online course taking with college completion and time-to-degree. *Educational Evaluation and Policy Analysis*. Advance Online Publication. https://doi.org/10.3102/01623737211055768

Fischer, C., Fishman, B., Levy, A., Eisenkraft, A., Dede, C., Lawrenz, F., Jia, Y., Kook, J., Frumin, K., & McCoy, A. (2020). When do students in low-SES schools perform better-than-expected on a high-stakes test? Analyzing school, teacher, teaching, and professional development characteristics. *Urban Education*, *5*(8–9), 1280–1314. https://doi.org/10.1177/0042085916668953

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining Big Data in Education: Affordances and Challenges. *Review of Research in Education*, *44*(1), 130–160. https://doi.org/10.3102/0091732X20903304

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology, 173*(7), 761-767.

Galla, B. M., Shulman, E. P., Plummer, B. D., Gardner, M., Hutt, S. J., Goyer, J. P., D'Mello, S. K., Finn, A. S., & Duckworth, A. L. (2019). Why High School Grades Are Better Predictors of On-Time College Graduation Than Are Admissions Test Scores: The Roles of Self-Regulation and Cognitive Ability. *American Educational Research Journal*, *56*(6), 2077–2115. https://doi.org/10.3102/0002831219843292

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549-576.

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, *29*(6), 911–922. https://doi.org/10.1016/j.econedurev.2010.06.010

Gurantz, O. (2021). How College Credit in High School Impacts Postsecondary Course-Taking: The Role of Advanced Placement Exams. *Education Finance and Policy*, *16*(2), 233–255. https://doi.org/10.1162/edfp_a_00298

Hallett, R. E., & Venegas, K. M. (2011). Is increased access enough? Advanced Placement courses, quality, and success in low-income urban schools. *Journal for the Education of the Gifted*, *34*(3), 468–487. https://doi.org/10.1177/016235321103400305

Jagesic, S., & Wyatt, J. (2018). *Postsecondary Course Performance of AP® Exam Takers in Subsequent Coursework: Biology, Languages, and Studio Art*. The College Board.

Kanno, Y., & Kangas, S. E. N. (2014). "I'm not going to be, like, for the AP": English language learners' limited access to Advanced College-preparatory courses in high school.

American Educational Research Journal, *51*(5), 848–878. https://doi.org/10.3102/0002831214544716

Keng, L., & Dodd, B. G. (2008). *A Comparison of College Performances of AP® and Non-AP Student Groups in 10 Subject Areas* (No. 2008–7). The College Board.

Klugman, J. (2013). The Advanced Placement arms race and the reproduction of educational inequality. *Teachers College Record*, *115*(5), 1–34.

Kyburg, R. M., Hertberg-Davis, H., & Callahan, C. M. (2007). Advanced Placement and International Baccalaureate programs: Optimal learning environments for talented minorities? *Journal of Advanced Academics*, *18*(2), 172–215. https://doi.org/10.4219/jaa-2007-357

Lin, C.-H., Borden, V. M. H., & Chen, J.-H. (2020). A Study on Effects of Financial Aid on Student Persistence in Dual Enrollment and Advanced Placement Participation. *Journal of College Student Retention: Research, Theory & Practice*, *22*(3), 378–401. https://doi.org/10.1177/1521025117753732

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education*, *95*(5), 877–907. https://doi.org/10.1002/sce.20441

Means, B., Wang, H., Young, V., Peters, V. L., & Lynch, S. J. (2016). STEM-focused high schools as a strategy for enhancing readiness for postsecondary STEM programs. *Journal of Research in Science Teaching*, *53*(5), 709–736. https://doi.org/10.1002/tea.21313

Morgan, R., & Klaric, J. (2007). *AP® Students in College: An Analysis of Five-Year Academic Careers* (No. 2007–4). The College Board.

Olson, S., & Riordan, D. G. (2012). *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics* [Report to the President]. Executive Office of the President.

Owen, S. (2021). Essays in the Economics of Education (dissertation). University of Michigan, Horace H. Rackham School of Graduate Studies, Ann Arbor, Michigan.

Price, H. E. (2020). The College Preparatory Pipeline: Disparate Stages in Academic Opportunities. *American Educational Research Journal*. https://doi.org/10.3102/0002831220969138

Roegman, R., & Hatch, T. (2016). The AP lever for boosting access, success, and equity. *Phi Delta Kappan*, *97*(5), 20–25.

Sadler, P. M., & Sonnert, G. (2010). High school Advanced Placement and success in college and coursework in the sciences. In P. M. Sadler, G. Sonnert, R. H. Tai, & K. Klopfenstein (Eds.), *AP: A critical examination of the Advanced Placement program* (pp. 119–137). Harvard Education Press.

Sadler, P. M., & Tai, R. H. (2007). Advanced Placement Exam Scores as a Predictor of Performance in Introductory College Biology, Chemistry and Physics Courses. *Science Educator*, *16*(2), 19.

Schneider, J. (2009). Privilege, equity, and the Advanced Placement program: Tug of war. *Journal of Curriculum Studies*, *41*(6), 813–831. https://doi.org/10.1080/00220270802713613

SEISMIC Collaboration (2022). SEISMIC Overview. Retrieved from https://www.seismicproject.org/about/overview/

Scott, T. P., Tolson, H., & Lee, Y.-H. (2010). Assessment of Advanced Placement participation and university academic success in the first semester: Controlling for selected high school academic abilities. *Journal of College Admission*, *208*, 26–30.

Smith, J., Hurwitz, M., & Avery, C. (2017). Giving College Credit Where It Is Due: Advanced Placement Exam Scores and College Outcomes. *Journal of Labor Economics*, *35*(1), 67–147.

Taliaferro, J. D., & DeCuir-Gunby, J. T. (2008). African American educators' perspectives on the Advanced Placement opportunity gap. *The Urban Review*, *40*(2), 164–185. https://doi.org/10.1007/s11256-007-0066-6

The College Board. (2020a). *Federal and State AP Exam Fee Assistance*. https://apcentral.collegeboard.org/ap-coordinators/exam-ordering-fees/exam-fees/federal-state-assistance

The College Board. (2020b). *Program Summary Report*. https://secure-media.collegeboard.org/digitalServices/pdf/research/2020/Program-Summary-Report-2020.pdf

Tienken, C. H. (2012). The Influence of Poverty on Achievement. *Kappa Delta Pi Record*, *48*(3), 105–107. https://doi.org/10.1080/00228958.2012.707499

Warne, R. T., Sonnert, G., & Sadler, P. M. (2019). The Relationship Between Advanced Placement Mathematics Courses and Students' STEM Career Interest. Educational Researcher, 48(2), 101–111. https://doi.org/10.3102/0013189X19825811

What Works Clearinghouse (2016.). Procedures and Standards Handbook (Version 3.0). Retrieved from https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_reviewer_guidance_030416.pdf

Whitcomb, K. M., Kalender, Z. Y., Nokes-Malach, T. J., Schunn, C. D., & Singh, C. (2020). Laying a foundation for success in engineering coursework: A predictive curriculum model. *International Journal of Engineering Education, 36*(4), 1340–1355.

Wyatt, J., Jagesic, S., & Godfrey, K. (2018). *Postsecondary Course Performance of AP® Exam Takers in Subsequent Coursework*. The College Board.

Zhang, Y. (Leaf). (2016). An Overlooked Population in Community College: International Students' (In)Validation Experiences With Academic Advising. *Community College Review*, *44*(2), 153–170. https://doi.org/10.1177/0091552116633293

Figures

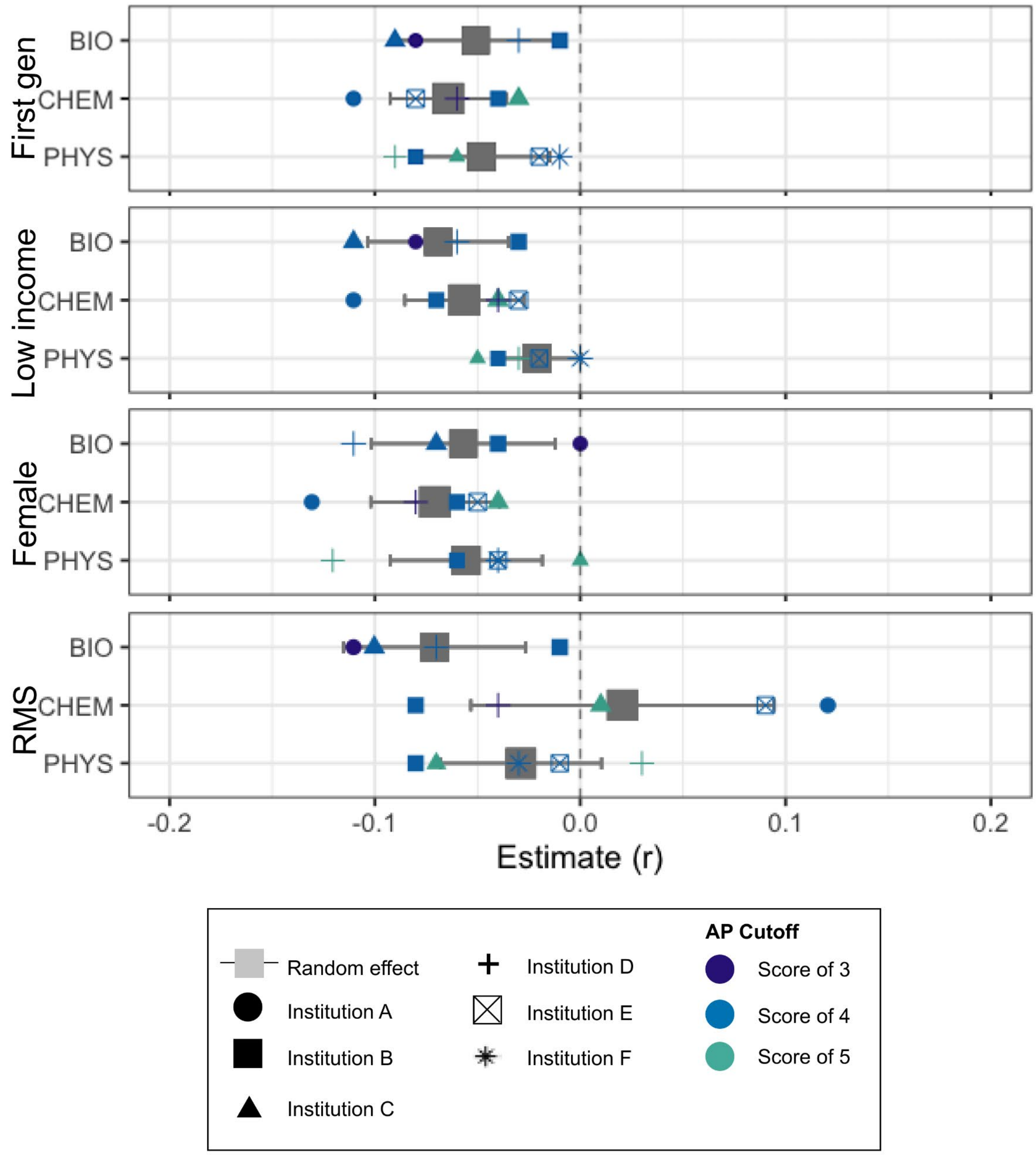

Figure 1. Meta-regression results on biserial correlations between demographic characteristics and earning an AP score that would allow a student to skip the first course in the sequence at their institution. Point shape indicates institution and point color indicates the cutoff value for AP score required by that institution for that discipline (see also Table 2). Large gray squares indicate estimated mean effects across institutions, along with 95% CI in the effect estimate. In all but one case, these estimates were negative, but small. Due to sample size limitations and inclusion criteria, not all institutions contribute to the estimates for each analysis.

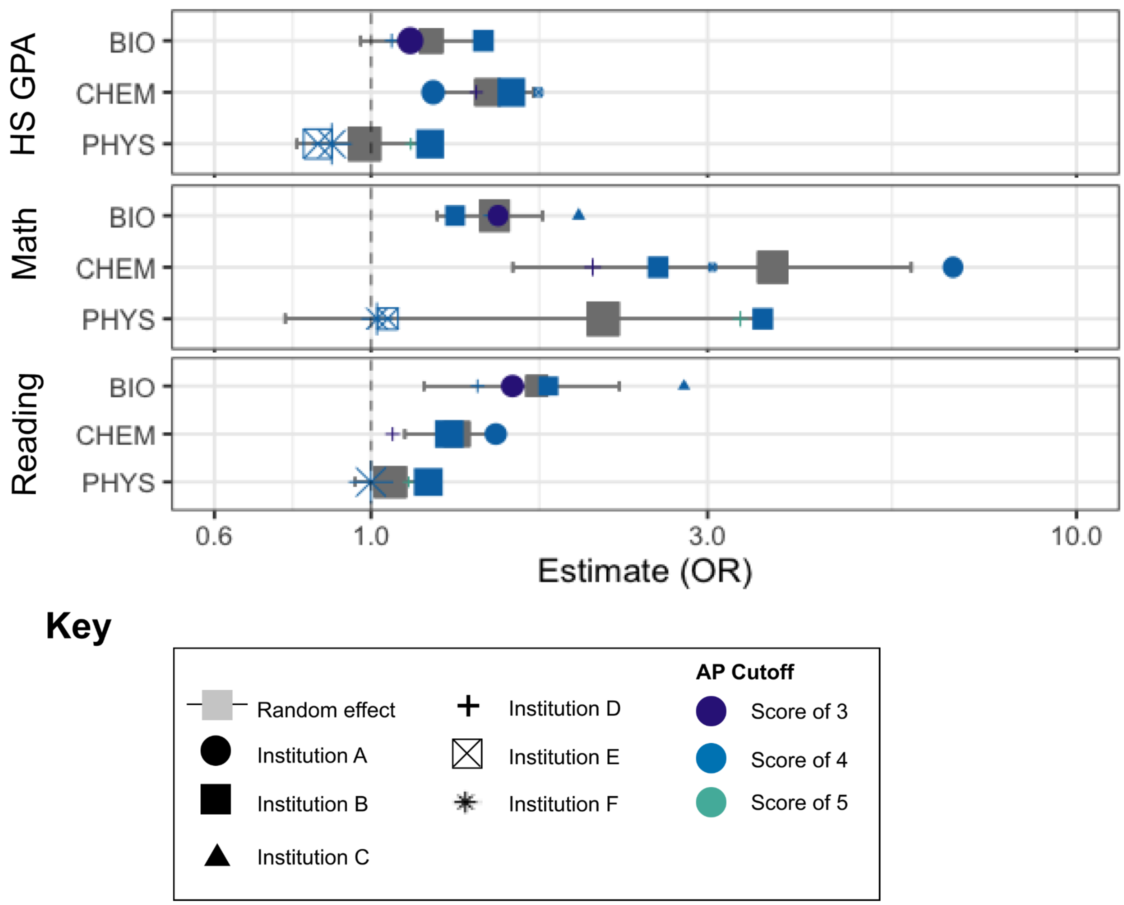

Figure 2. Meta-regression results on Odds Ratios (OR) of HS GPA, SAT Math, and SAT Reading Score predicting skip eligibility (controlling for demographic variables). As in Figure 1, point shape indicates institution and point color indicates the cutoff value for AP score required by that institution for that discipline. Here, point size is proportional to the sample size from each institution. Large gray squares indicate estimated mean effects across institutions, along with 95% CI in the effect estimate. Results are variable across subject and predictor, with HS GPA having a small predictive value on Chemistry skip eligibility, SAT reading and math both having a small predictive value on Biology skip eligibility, and SAT math being relatively strongly predictive of Chemistry and Physics skip eligibility. Due to sample size limitations and inclusion criteria, not all institutions contribute to the estimates for each analysis.

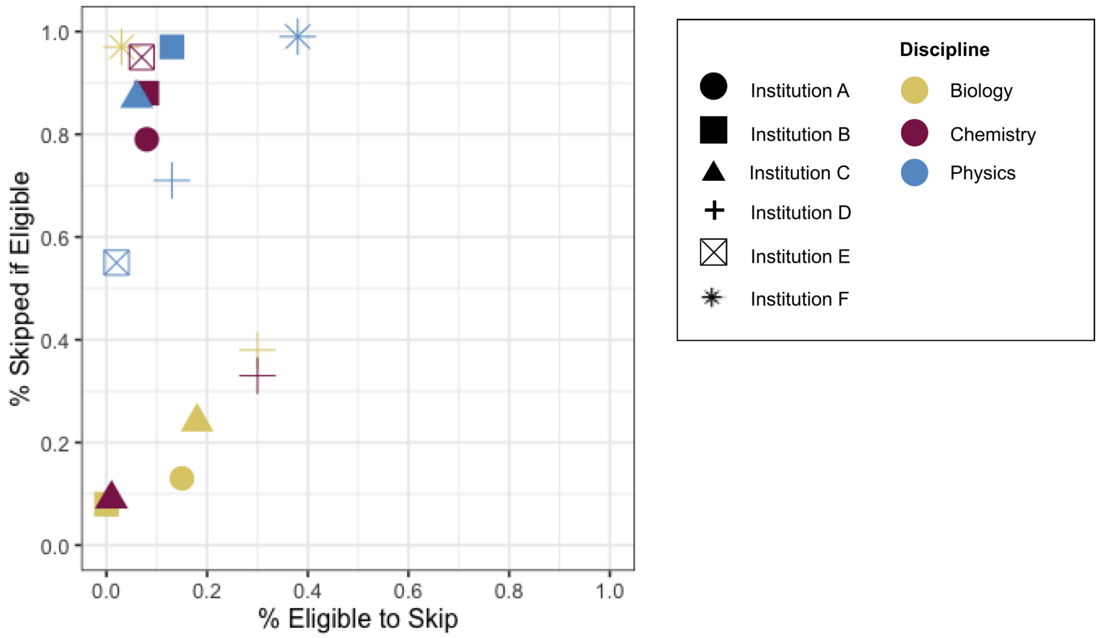

Figure 3. For each institution (indicated by shape) and within each discipline (indicated by color), the x-axis shows the proportion of students that earned an AP score that made them eligible to skip the introductory course, while the y-axis shows the proportion of eligible students that chose to actually skip that course. We found no relationship between skip-eligibility and course skipping.

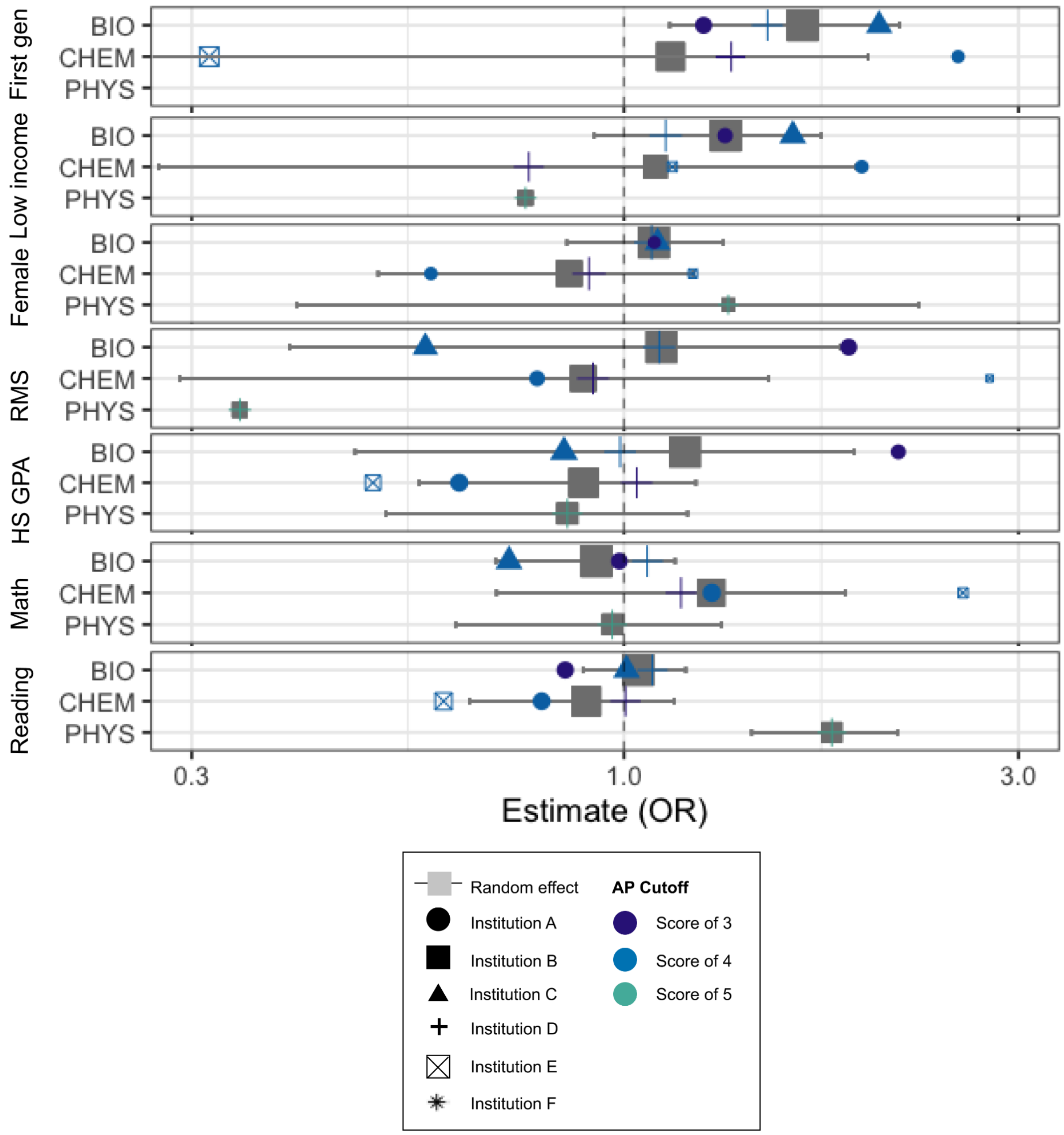

Figure 4. Estimated effect sizes (OR = odds ratios) for each potential demographic and academic resource from logistic regressions predicting whether students actually skip when eligible to skip. Like previous figures, institution is indicated by shape and AP score cutoff value required for skip eligibility is indicated by point color. The mean estimated effect from the meta-regressions across institutions (along with 95% CI bars) is shown with gray squares. In all but one case, these estimates had very large uncertainties and were not significantly different from zero. Only for first-generation students in Biology was the correlation statistically significant. Due to sample size limitations and inclusion criteria, not all institutions contribute to the estimates for each variable.

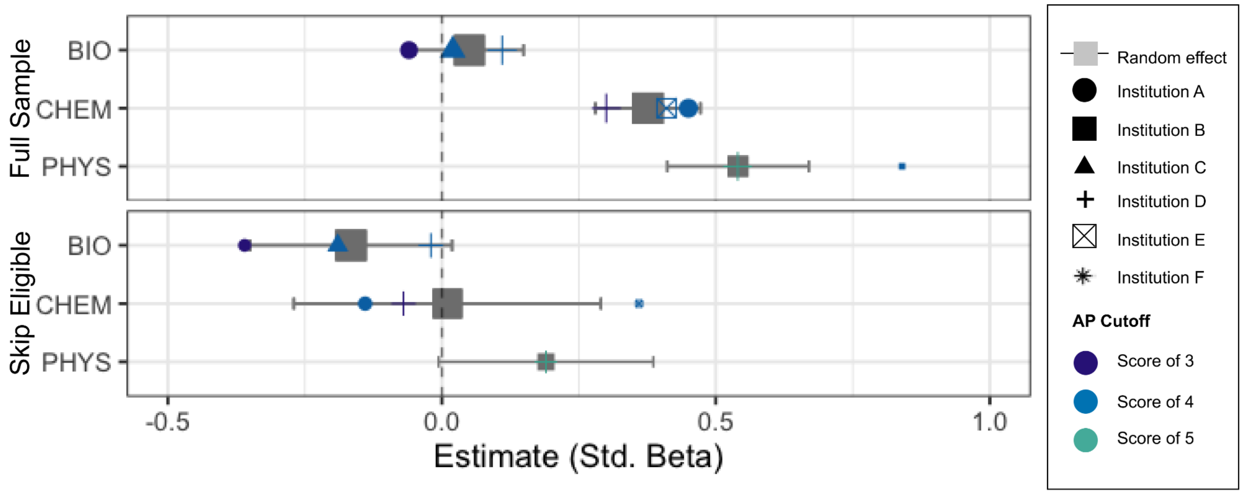

Figure 5. The estimated standardized effect of skipping the first course on students' second course grade in each course sequence at each institution (indicated by shape) and for each AP cutoff score (indicated by color). The estimates are derived from multiple regression controlling for demographics and academic resources and applying inverse probability of treatment weights in the regressions. The mean estimated effect from the meta-regressions across institutions (along with 95% CI bars) is shown with gray squares. The top graph involves the full sample of students, and the bottom graph uses only those students who earned a score that made them eligible to skip the first course. Although the estimates based on the full sample suggest a neutral to positive effect on second course grades, the estimates based on skip eligible students only are less positive. Due to sample size limitations and inclusion criteria, not all institutions contribute to the estimates for each analysis.

Tables

Table 1. Descriptive information for all variables by science discipline for each institution with sample Ns (and corresponding AP credit policies in parentheses).

BIOLOGY	**Inst. A**	**Inst. B**	**Inst. C**	**Inst. D**
	N=1561(≥3)	N=3738(≥4)	N=3250(≥4)	N=3108(≥4)
Demographic				
Female	56%	67%	58%	62%
First generation	28%	24%	14%	9%
Low income	40%	25%	21%	11%
Ethnicity				
White	46%	75%	72%	71%
RMS	33%	13%	14%	9%
Asian	21%	8%	14%	20%
AP				
Took AP	16.3%	24.0%	38.1%	54.0%
Eligible to skip	12.4%	0.4%	18.3%	30.0%
Skipped course	1.7%	0.0%	4.5%	12.0%
Academic	M(SD)	M(SD)	M(SD)	M(SD)
HS GPA	3.21(1.10)	3.74(0.30)	3.85(0.24)	4.10(0.40)
Math proficiency	604(84)	611(66)	658(72)	674(66)
English proficiency	595(97)	622(74)	656(63)	675(61)
Course1 grade	3.24(0.90)	2.85(0.90)	3.46(0.73)	2.9(0.90)
Course2 grade	2.91(1.10)	3.07(0.90)	2.96(1.14)	2.7(0.90)

Table 1 continued

CHEMISTRY	Inst. A	Inst. B	Inst. C	Inst. D	Inst. E
	N=2709(≥4)	N=2722(≥3)	N=3721(≥5)	N=2494(≥3)	N=8266(≥4)
Demographic					
Female	57.10%	64.40%	61.00%	59.20%	64.30%
First generation	27.80%	23.80%	14.10%	8.70%	49.00%
Low income	37.40%	26.20%	22.70%	10.40%	43.60%
Ethnicity					
White	50%	70%	72.	70%	15%
RMS	32%	15%	15%	9%	33%
Asian	18%	10%	13%	21%	52%
AP					
Took AP	17.9%	18.5%	20.0%	40.4%	14.1%
Eligible to skip	6.0%	8.4%	0.6%	30.2%	2.0%
Skipped course	4.7%	7.3%	0.1%	14.0%	1.6%
Academic	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
HS GPA	3.17(1.20)	3.77(0.30)	3.85(0.24)	4.10(0.40)	3.98(0.22)
Math proficiency	618(80)	618(71)	653(71)	682(64)	597(81)
English proficiency	609(97)	623(71)	651(64)	680(57.5)	555(81)
Course1 grade	3.08(0.80)	3.17(0.90)	2.97(1.01)	3.00(0.80)	2.67(0.87)
Course2 grade	2.98(1.20)	2.99(1.10)	2.04(1.52)	2.80(0.90)	2.39(1.10)

Table 1 continued

PHYSICS	**Inst. F**	**Inst. B**	**Inst. C**	**Inst. D**	**Inst. E**
	N=3911(≥5)	N=2122(≥4)	N=547(≥5)	N=1724(≥5)	N=3090(≥4)
Demographic					
Female	24%	24%	31%	31%	26%
First generation	8%	14%	12%	6%	57%
Low income	12%	17%	18%	7%	36%
Ethnicity					
White	62%	80%	80%	85%	17%
RMS	9%	8%	9%	5%	34%
Asian	19%	9%	12%	10%	49%
AP					
Took AP	6.0%	26.0%	12.0%	32.0%	16.5%
Eligible to skip	36.3%	12.7%	5.5%	13.0%	2.5%
Skipped course	36.3%	12.2%	4.8%	14.6%	1.7%
Academic	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
HS GPA	3.80(0.80)	3.81(0.30)	3.86(0.24)	4.20(0.40)	4.03(0.23)
Math proficiency	739(46)	672(66)	700(66)	721(52)	650(76)
English proficiency	521(109)	645(76)	680(58)	697(55)	587(84)
Course1 grade	1.98(1.70)	3.32(0.70)	3.23(0.85)	2.8(0.70)	3.03(0.78)
Course2 grade	2.78(1.10)	3.06(0.80)	2.91(1.19)	2.6(0.90)	2.46(1.03)

Table 2. AP score requirements to be eligible to skip the first course in the course sequence across disciplines and institutions.

	Inst. A	Inst. B	Inst. C	Inst. D	Inst. E	Inst. F
Biology	3	4	4	4	No skip	4
Chemistry	4	3	5	3	4	No skip
Physics	No skip	4	5	5	4	5

Table 3. Meta Regression Results for Each Research Question within Each Discipline: effect estimate along with 95% CI and effect statistical significance; degree of heterogeneity across institutions and statistical significance of that observed heterogeneity.

		Confidence Interval			Test for Heterogeneity (Q)	
Variable	Estimate	Lower 95%	Upper 95%	p	Q	p
			BIOLOGY			
RQ1						
First generation	-0.051	-0.089	-0.013	***	13.74	**
Low-income	-0.069	-0.103	-0.035	***	11.60	*
Female	-0.057	-0.101	-0.012	***	15.29	**
RMS	-0.071	-0.115	-0.037	***	18.61	***
High school GPA	1.217	0.969	1.464	0.141	1.32	0.724
Math score	1.495	1.246	1.743	***	0.43	0.933
Verbal score	1.719	1.188	2.249	**	1.29	0.731
RQ2						
First generation	1.645	1.136	2.154	**	1.81	0.404
Low-income	1.327	0.921	1.733	0.172	1.38	0.502
Female	1.086	0.853	1.319	0.606	0.00	0.998
RMS	1.111	0.394	1.827	0.770	3.13	0.209
High school GPA	1.185	0.473	1.897	0.712	6.44	*
Math score	0.926	0.699	1.153	0.666	4.94	0.085
Verbal score	1.039	0.892	1.187	0.693	0.80	0.669
RQ3						
Full sample	0.050	-0.049	0.149	0.486	2.44	0.295
Skip eligible	-0.166	-0.350	0.019	0.169	7.19	*

Table 3 continued

		Confidence Interval			Test for Heterogeneity (Q)	
Variable	Estimate	Lower 95%	Upper 95%	p	Q	p
		CHEMISTRY				
RQ1						
First generation	-0.064	-0.093	-0.036	***	13.68	**
Low-income	-0.056	-0.085	-0.028	***	14.90	**
Female	-0.071	-0.102	-0.039	***	16.47	**
RMS	0.021	-0.053	0.094	***	96.96	***
High school GPA	1.463	1.229	1.698	***	3.36	0.339
Math score	3.708	1.587	5.829	***	496.96	***
Verbal score	1.324	1.113	1.536	**	0.64	0.886
RQ2						
First generation	1.139	-0.008	2.285	0.795	5.11	0.078
Low-income	1.092	0.274	1.910	0.785	2.22	0.330
Female	0.858	0.504	1.212	0.632	0.56	0.757
RMS	0.893	0.290	1.495	0.769	0.05	0.977
High school GPA	0.894	0.566	1.222	0.677	2.64	0.268
Math score	1.276	0.701	1.850	0.492	1.18	0.554
Verbal score	0.900	0.650	1.150	0.614	1.90	0.387
RQ3						
Full sample	0.376	0.280	0.472	***	3.58	0.167
Skip eligible	0.010	-0.270	0.291	0.796	5.55	0.062

Table 3 continued

Variable	Estimate	Confidence Interval Lower 95%	Confidence Interval Upper 95%	p	Test for Heterogeneity (Q) Q	Test for Heterogeneity (Q) p
PHYSICS						
RQ1						
First generation	-0.048	-0.082	-0.015	***	12.73	0.013*
Low-income	-0.021	-0.043	0.002	***	3.07	0.547
Female	-0.056	-0.093	-0.019	***	10.63	0.031*
RMS	-0.029	-0.068	0.010	***	13.50	**
High school GPA	0.979	0.786	1.172	0.798	13.02	**
Math score	2.133	0.756	3.510	0.123	500.98	***
Verbal score	1.065	0.950	1.180	0.417	5.79	0.122
RQ2						
Low-income	0.760	-0.531	2.051	0.581	-	-
Female	1.337	0.402	2.272	0.677	-	-
RMS	0.343	0.285	1.940	0.492	-	-
High school GPA	0.854	0.515	1.193	0.614	-	-
Math score	0.968	0.625	1.311	0.677	-	-
Verbal score	1.784	1.427	2.141	0.492	-	-
RQ3						
Full sample	0.541	0.412	0.670	***	0.03	0.855
Skip eligible	0.190	-0.006	0.386	0.131	-	-

Notes. * $p < .05$; ** $p < .01$; *** $p < .001$.