The Advanced Placement Program and Educational Inequality

Stephanie Owen*

January 2024

*Assistant Professor, Colby College. Contact: sowen@colby.edu. I thank Sue Dynarski, Sara Heller, Kevin Stange, and Charlie Brown for their invaluable support and guidance. Amanda Griffith and Jeff Smith provided helpful comments on early drafts. I thank Thomas Helgerman for multiple conversations about two-way fixed effects and Soobin Kim for additional information about the Michigan High School Transcript Study. I gratefully acknowledge financial support from the U.S. Department of Education's Institute of Education Sciences through PR/Award R305B150012#. This research result used data structured and maintained by the MERI-Michigan Education Data Center (MEDC). MEDC data are modified for analysis purposes using rules governed by MEDC and are not identical to those data collected and maintained by the Michigan Department of Education (MDE) and/or Michigan's Center for Educational Performance and Information (CEPI). Results, information and opinions solely represent the analysis, information and opinions of the author and are not endorsed by, or reflect the views or positions of, grantors, MDE and CEPI or any employee thereof.

Abstract

The Advanced Placement (AP) program is widely offered in American high schools and has been touted as a way to close racial and socioeconomic gaps in educational outcomes. Using administrative data from Michigan, I exploit variation within high schools across time in AP course offerings to identify the relationship between AP course availability, AP participation, and postsecondary outcomes. I find that students from non economically disadvantaged families, white and Asian students, and higher-achieving students are more likely to take advantage of additional AP courses when they are offered, thus widening existing gaps in course-taking. I find little evidence that additional AP availability is related to improved college outcomes for any students, with the exception of the most academically prepared students. Expanding access to AP courses without additional incentives or support for disadvantaged students to succeed is unlikely to address educational inequality.

1 Introduction

Since its introduction in 1952, the Advanced Placement (AP) program, which provides an opportunity for high school students to take college-level courses and possibly obtain college credit or placement out of introductory college courses, has grown dramatically and is now offered in the majority of American high schools. However, the AP program originally served an extremely elite set of high schools, and disparities in access remain today. Recent policy efforts have called for expanded access of AP courses to an even wider set of students and schools (Tugend 2017; Burnett and Burkander 2021). Despite the popularity of AP and the perception that participation improves college preparation, increases chances of admission to selective colleges, and accelerates degree attainment, there is little convincing causal evidence on how taking AP courses affects human capital investment and later outcomes. It is also not clear if further expanding AP curricula would address existing educational inequality or exacerbate it by primarily serving and benefiting already advantaged students.

In this paper, I study two key aspects of AP. First, I investigate which types of students take advantage of additional AP courses when whey become available. Second, I study whether and how Advanced Placement courses are related to college enrollment and degree receipt. I use administrative data from the state of Michigan and exploit variation within high schools across time in how many AP courses are offered to identify a plausibly causal effect of AP course availability. Although there is obvious selection into which types of schools offer more AP courses, the fixed effects strategy eliminates any bias stemming from fixed characteristics of schools (such as size, geographic location, or a high baseline level of parent involvement) and compares *changes* in course offerings to *changes* in outcomes.

I find that when a high school offers an additional AP course, there is a small increase in participation. The proportion of students taking any AP increases by 1.1 percentage point. The average number number of AP courses taken increases by 0.032, which translates into fewer than 10 enrollments in the typical high school. Participation in AP exams increases by less: 0.4 percentage points on the extensive margin and a statistically insignificant 0.011 exams on the intensive margin. (Estimated effects on course- and exam-taking are of similar magnitude when applying two-way fixed effect bias corrections, but more noisily estimated.) Taking the exam and receiving a sufficiently high score are required for college credit or placement, and selective colleges use AP scores to evaluate applicants, so the changes (or lack thereof) to exam-taking are important for understanding downstream effects.

Not only do expanded AP offerings serve few students, they serve students unequally. Although students of different socioeconomic background and race see similar increases in the probability of taking any AP course (around 1 percentage point), I only detect an increase in the average number of AP courses for students from higher-income families (those not eligible for subsidized school meals) and white and Asian students. These patterns persist even conditional on prior achievement. I find the strongest effects on AP course- and exam-taking for students with the highest levels of academic preparation (measured by performance on a standardized math test in middle school). The results suggest that additional AP courses are mostly taken by students already taking other APs.

I find little evidence that a school offering an additional AP course improves students' outcomes, with the exception of the highest-achieving students. I find precisely estimated null effects of AP course availability on college enrollment, college selectivity, and degree attainment, and no differential effects by family income or race. Students who enter high school with strong academic preparation are the only ones who experience positive benefits of additional course offerings. I estimate that for a student whose middle school math performance puts them in the top 25 percent of my sample, having an an additional AP course available increases their likelihood of enrolling in a four-year college by 0.5 percentage points, of enrolling in a competitive college by 0.5 percentage points, and of earning a bachelor's degree in four years by 0.7 percentage points. However, these effects are not robust to two-way fixed effects bias corrections or to all alternate specifications. With this caveat in mind, a two-stage least squares approach suggests that for a high-achieving student induced to take an additional AP course, doing so increases their chance of enrolling in a competitive college by over 6 percentage points, and their chance of on-time BA completion by 10 percentage points.

Taken as a whole, the results suggest that at best, expanding AP programs may benefit already high-achieving students but do little to close achievement gaps. At worst, increasing AP course offerings may exacerbate socioeconomic and racial inequalities in access to advanced coursework.

The paper proceeds as follows: Section 2 provides history and background on the AP program; Section 3 reviews prior related work; Section 4 describes the methodological approach and data; Section 5 presents findings about the effect of AP course availability on course-taking and exam-taking, as well as on college outcomes; Section 6 discusses threats to identification and presents robustness checks; and Section 7 concludes.

2 Background

The AP program traces its origins to just after World War II, when the Ford Foundation created the Fund for the Advancement of Education and concluded that better coordination between secondary and postsecondary schools would help increase the number of college entrants and graduates in the United States and serve national security interests (Rothschild 1999; College Board 2003; Schneider 2009). A committee was formed "to develop high school course descriptions and assessments that colleges would find rigorous enough to use as a basis for granting credit" and a pilot program in 11 subject areas was launched in 1952 (College Board 2003, p. 1). Since 1955, the AP program has been run by the College Board, the same non-profit organization responsible for the SAT college entrance exam.

Participation has grown dramatically since the program's inception, from 1,229 students at

104 schools nationwide in the 1955-56 academic year (the first year data are available from the College Board) to 2.5 million students and nearly 23,000 schools in 2021 (College Board 2021a). In 2012, 74 percent of all public high schools offered AP courses (Malkus 2016), and these schools serve even more students as a proportion of all public high school students (Theokas and Saaris 2013). In the 2015-16 school year, 85 percent of public high school students attended schools offering at least one AP course (Chatterji, Campbell, and Quirk 2021). However, access is unequal across students and schools. Smaller and rural schools are less likely to offer AP, and Black and Indigenous students are underrepresented in schools that offer high numbers of AP courses. Even at schools with the same number of course offerings, Black, Latinx, and Indigenous students are less likely to enroll (Chatterji, Campbell, and Quirk 2021).

A non-trivial amount of federal, state, and local public funds are dedicated to subsidizing AP teacher training, exam fees, and performance incentives (Klopfenstein 2010). The U.S. Department of Education created Advanced Placement Incentive Program grants in the late 1990s to increase AP participation among students from low-income households and reduce achievement gaps; this program was expanded under No Child Left Behind in 2001 (Klopfenstein 2010). In 2016, the Department of Education awarded over $28 million to subsidize exam fees for students from low-income households in 41 states (including $560,000 to Michigan) plus the District of Columbia (U.S. Department of Education 2016). A number of large school districts, including New York City and Washington, D.C., have adopted policies mandating a minimum number of AP offerings per school (Tugend 2017). Some schools in Washington, D.C. require all students to enroll in at least one AP course (Burnett and Burkander 2021). These policies stem from a belief that expanding access to AP can narrow racial and socioeconomic gaps in educational outcomes (Schneider 2009; Quinton 2015; Tugend 2017).

The AP program serves several ostensible purposes. The College Board describes it as a way to “[enable] willing and academically prepared students to pursue college-level studies—with the opportunity to earn college credit, advanced placement or both—while still in high school” (Rodriguez, McKillip, and Niu 2013, p. 1). The College Board also touts participation as beneficial to college admission and performance, saying that “Taking AP courses demonstrates to college admission officers that students have sought the most rigorous curriculum available to them, and research indicates that students who score a 3 or higher on an AP Exam typically experience greater academic success in college and are more likely to earn a college degree than non-AP students” (Rodriguez, McKillip, and Niu 2013, p. 1). As summarized by Klopfenstein and Thomas (2010), “while the College Board generally makes no explicit statements that AP experience is a cause of college success, their promotional literature readily leads readers to such a conclusion” (p. 170).

As of 2023, the College Board offers 38 AP courses in six subject areas: science, math and computer science, history and social sciences, English, world languages and cultures, and arts. In 2021, the most popular subjects (by exams taken) were English Language and Composition, U.S. History, English Literature and Composition, World History, Psychology, and U.S. Government (College Board 2021b).

3 Related Literature

The current study fits within a number of overlapping literatures. One way to conceptualize the AP program is as a high-ability track within a high school. There is a large literature on ability and achievement tracking that informs theory about the effects of AP participation, particularly differential effects by student type (see Betts 2011 for a review). Theoretically, tracking systems may involve an efficiency-equity tradeoff. Proponents argue that grouping students by ability allows teachers to tailor content and approach, while opponents assert "that it condemns students placed into the lower tracks to lower educational attainment, and [...] aggravates economic inequality and perpetuates economic disadvantage across generations" (Betts 2011, p. 343). The empirical evidence on within-school tracking is mixed (Betts 2011). For example, an experiment by Duflo, Dupas, and Kremer (2011) found that students of all ability levels performed better when sorted into ability-based classrooms. On the other hand, Marascuilo and McSweeney (1972) is an example of a tracking experiment that decreased student learning overall, and for medium- and low-ability students in particular. Others see AP as a type of or alternative to dual enrollment programs (see, for example, Klopfenstein and Lively 2012), which have the explicit goal of reducing the financial and time cost of postsecondary education.

A substantial body of work has focused on inequities in access to AP, and the role of advanced coursework in maintaining educational segregation and inequality by race and income. The program originated in partnership with elite private preparatory schools serving overwhelmingly wealthy, white students (Schneider 2009). Even as participation increased, it was concentrated among white students in affluent private and suburban public schools, and "some people [began to] regard the program as touched with . . . 'institutional racism'" (Hochman 1970, p. 17, quoted in Schneider 2009). Even as education reformers advocated for expanding AP access to underserved students, schools serving economically disadvantaged and minority populations faced constraints in the form of proper teacher training, academic preparation, and low expectations about student ability (Schneider 2009). This history of structural inequity may prevent economically disadvantaged and underrepresented minority students from accessing advanced classes, as well as contribute to lower levels of academic preparation that hinder participation and success in AP courses.

Research by historians, sociologists, and education researchers has argued that the AP program, like many other examples of educational resources, benefits already privileged students and systematically excludes the already marginalized, thus perpetuating inequities even within schools (Schneider 2009; Lewis and Diamond 2015). One framework is race- and class-based opportunity hoarding, wherein a dominant group "gains access to a valuable and renewable resource and precludes others from benefiting from said resource" (Rodriguez and McGuire 2019, p. 650). This could come in the form of privileged parents parents advocating for their own children, and school staff steering students of color away from advanced courses. Rodriguez and McGuire (2019) use cross-sectional national data and instrument for AP availability with per-pupil school expenditures and find that when schools introduce additional AP courses, the Black-white gap in AP course-taking widens. They argue that their results imply opportunity hoarding by white students and families. Course-taking disparities could also stem from students of color opting out due to not feeling welcome in predominantly white classrooms or doubting the quality of programs in underresourced schools (Rodriguez and McGuire 2019). Solorzano and Ornelas (2002) show that Chicana and Latina students in one California district are underrepresented in AP courses, even in schools with strong AP programs; they argue that "school structures, processes, and discourses help maintain racial/ethnic/gender/class discrimination in access to AP/Honors classes" (p. 219). Some argue that racial and income-based disparities in AP course-taking are due to different levels of academic preparation—differences that are a result of lower access to educational resources. Conger, Long, and Iatarola (2009) find that the Black-white and Hispanic-white gaps in advanced course-taking in Florida reverse once middle school test scores are accounted for; this finding remains with the inclusion of school fixed-effects.

There are several key mechanisms by which we might expect participation in AP courses and exams to affect educational choices and outcomes. AP courses are generally considered more rigorous than standard high school classes, so that the experience of taking an AP course and preparing for the exam may directly increase students' knowledge, skills, and college readiness. AP participation can serve as a signal of student ability, motivation, and college readiness, as well as a signal of school quality, which are used by admissions committees at selective colleges in evaluating applicants. Students earning sufficiently high scores on AP exams (a three on a five-point scale, in most cases, though policies vary by institution) can earn college credit and/or placement out of introductory-level college courses, thus shortening time to graduation.

The predicted effects of AP participation are not unambiguously positive. Performance in AP courses and particularly on AP exams may serve as a signal to students that causes them to re-asses their own academic ability and potential for college success; depending on performance, students could revise their self-assessment upwards or downwards.[1] The effort required and stress induced by rigorous AP courses could crowd out effort in other academic and non-academic tasks, depending on the degree of complementarity between the various tasks. This is particularly important in considering policies that subsidize or incentivize AP participation in some way, as they may induce some students to take more than the socially optimal number of AP courses or exams.

1. For evidence that grades and standardized test scores can lead to this type of belief updating, see Jacob and Wilder (2010), Goodman (2016), and Gonzalez (2017); and Avery and Goodman (2022).

Rigorous causal evidence on the effect of AP is fairly limited. Jackson (2010) evaluates the Advanced Placement Incentive Program in Texas, which paid students and teachers for passing AP exams and provided training to teachers. He exploits exogenous variation in when schools implemented the program and finds that it increased participation in AP courses and exams, the number of students scoring highly on the SAT or ACT, and college matriculation. In a longer-term follow up, Jackson (2014) shows positive effects on degree attainment and earnings.

In the only experimental work to date, Conger et al. (2021) randomly assigned high school students into a treatment that included the option to enroll in newly introduced AP Biology or Chemistry course in their schools. Taking an AP science course resulted in a higher self-reported level of course rigor and a higher level of science skill. However, in a longer-term follow-up, Conger, Long, and McGhee (2023) find no effect on SAT or ACT performance, no change in students' self-reported portfolio of college applications, and no ultimate effect on selective college enrollment. They also find suggestive evidence that competitive college enrollment may have *decreased*. These somewhat discouraging findings point to the importance of considering who the marginal students are when expanding access. In a world where nearly all schools have AP courses available, the marginal student may be less prepared and unlikely to benefit.

A related series of studies exploit cutoffs in continuous AP exam scores that translate into the 1-5 integer scores reported to students and colleges. Smith, Hurwitz, and Avery (2017) find that receiving a credit-granting score (a three in most cases) on an AP exam positively affects on-time college graduation. Avery et al. (2018) use the same regression discontinuity design as Smith, Hurwitz, and Avery (2017) and find that receiving a higher score on an AP exam significantly increases the likelihood that a student will major in that subject in college; they argue that "a substantial portion of the overall effect is driven by behavioral responses to the positive signal of receiving a higher score" (p. 918). Gurantz (2021) uses a similar regression discontinuity strategy to examine college course-taking by subject, finding that women who earn credit from AP exams in STEM subjects take more STEM courses.

It is important to note that receiving credit or placement is contingent on taking and passing an AP exam. A significant proportion of students who take an AP course do not take the associated exam; Fazlul, Jones, and Smith (2021) find that 15 percent of AP course enrollments (in four metro Atlanta school districts) do not result in an exam. Even among those who take an exam, many do not receive a passing score. These numbers are likely even higher for schools and students on the margin of offering and taking AP. In the setting of Conger, Long, and McGhee (2023)—schools that had not previously offered AP science—40 percent of treated students opted out of the exams, and 85 percent of those who did take the exams did not pass.

The current study represents, to my knowledge, the first plausibly causal evidence on the long-term effects of AP course offerings on college outcomes. Although Conger, Long, and McGhee (2023) examine effects on enrollment, they do not yet have results on college persistence and graduation. Given the possibility that AP affects college readiness in ways that may not be reflected in standardized test scores, long-term effects may emerge even in the absence of shorter-term ones. Furthermore, I use naturally occurring changes to AP course offerings across a number of subjects, whereas Conger, Long, and McGhee (2023) focus on AP science only. Although advanced science courses are an important part of the curriculum to study, my findings are relevant to a larger set of schools. Jackson (2014) studies college completion, but for a program that financially incentivizes students and teachers to pass exams, paired with teacher training. It is unclear whether his positive findings translate to a program with fewer resources and with less emphasis on exams. Smith, Hurwitz, and Avery (2017) look at on-time college graduation as an outcome, but only for students who take an AP exam and are close to passing. As mentioned above, this population is a small subset of the overall population of students on the margin of taking AP. While the small portion of students who take and pass an AP exam may benefit, a full accounting of the effects of AP must consider any effects (or lack thereof) on a larger population. Thus, the findings of the current paper will be useful to educators making the highly relevant decision of whether to offer an additional AP course or hire or reallocate an additional AP teacher.

4 Method and Data

4.1 Empirical Specification

Simply comparing students or schools with different levels of AP courses will likely give an upwardly biased estimate of the effect on educational outcomes, since students taking AP and schools offering AP tend to be higher-achieving to begin with. For example, in my sample, students who took at least one AP course have average middle school test scores that are a full standard deviation higher than students who never took AP. At the school level, the number of AP courses offered is highly correlated with the prior achievement level of the school's students.

To account for underlying differences in the types of schools (and students attending them) that have more robust AP programs, I exploit time variation in how many AP courses a high school offered each year. I use panel data covering the graduating classes of 2005 through 2012 in a sample of Michigan public high schools. My strategy is similar to that of Darolia et al. (2020), who use what they argue is "plausibly exogenous variation in course offerings within high schools over time" (p. 22) to study the effect of STEM course availability on postsecondary STEM enrollment and degree attainment in Missouri. My identification strategy, like theirs, hinges on year-to-year differences in course offerings within a school being (conditionally) exogenous. This would be the case if the variation is due to factors such as unrelated changes in teaching staff (due to, e.g., retirement or parental leave) and rules governing class size.

By controlling for school fixed effects, I compare a cohort of high school seniors to another cohort from the same school, where one cohort had a higher number of AP courses available to them. School fixed effects account for any fixed (i.e., unchanging over time) underlying characteristics of schools that are related to both AP availability and student outcomes. For example, if more rural schools are both less able to offer AP and send fewer students to college, the school fixed effects would eliminate the omitted variable bias associated with rural/urban status. The fixed effects estimator also controls for underlying school-level differences in school and family resources. The same would be true of harder-to-measure fixed characteristics, such as parental involvement, an underlying college-oriented school culture, or a strong college counseling program—as long as those characteristics don't change along with AP offerings.

I also include year fixed effects to account for the general upward trend in both AP and college outcomes. I include school-specific linear time trends to account for the possibility that schools on an especially steep trajectory in terms of outcomes differentially select into offering more APs. The relationship between AP offerings and outcomes thus captures deviations from trends: in years when a school has a larger change in AP offerings, do student outcomes experience a correspondingly large change?

I start by examining the effect of additional AP course offerings on students' participation in AP courses and exams, by estimating:

$$\mathrm{D}_{ijt,t-1} = \alpha_0 + \alpha_1(\#\text{ AP courses available})_{jt,t-1} + \delta_j + \lambda_t + \tau_j t + \varepsilon_{jt} \tag{1}$$

where i refers to a student, j to a high school, and t to year of high school graduation. The treatment is the count variable $(\#\text{ AP courses available})_{jt,t-1}$: the number of AP subjects available to cohort t at school j during their junior and senior year.[2,3] D refers to four different measures of AP participation: an indicator for taking any AP courses; the number of AP courses taken; an indicator for taking any AP exams; and the number of AP exams taken. As with course availability, I measure AP courses and exams taken in a student's junior and senior year. δ_j are school fixed effects; λ_t are cohort fixed effects; and τ_j are school-specific linear time trends. I estimate all equations with ordinary least squares and cluster standard errors at the school level. The parameter α_1 identifies within-school changes in AP participation when the number of courses offered changes.

2. As an example, if school j offered AP Biology and U.S. History in 2006 and Biology and U.S. Government in 2007, $AP_{j,2007,2006}$ would equal 3. This variable can take values between 0 and 26 AP subjects. I collapsed a number of subjects that the transcript data didn't allow me to distinguish between. For example, microeconomics and macroeconomics are two distinct subjects, but many schools just listed "AP economics." Appendix Figure B1 summarizes these decisions.

3. By focusing on AP courses offered and taken junior and senior year rather than all four years of high school, I am able to include more cohorts. At the course level, 91 percent of AP courses in the sample are taken by juniors or seniors. 81 percent of students who take AP take all of their AP courses in junior or senior year.

A focus of the analysis is not just whether increasing AP offerings increases access, but for whom. To test for heterogeneity by socioeconomic status, I subset the data and estimate Equation 1 separately for students who are and are not eligible for free or reduced-price lunch (FRPL) in 12th grade.[4] To test for heterogeneity by race, I estimate separate regressions for underrepresented minority (URM) students (i.e., Black, Hispanic, or Native) and non-URM (white or Asian).

To test for heterogeneity by academic preparation, I use students' standardized score on the statewide math test students take in middle school (which I describe in more detail in Section 4.2 and Appendix B). For the heterogeneity analysis by test score, students missing test scores are omitted. I test for prior achievement heterogeneity in two ways. The first is with an interaction term between the number of available AP courses and standardized score on the Michigan math test in middle school:

$$D_{ijt,t-1} = \eta_0 + \eta_1 AP_{jt,t-1} + \eta_2 Math_i + \eta_3 AP_{jt,t-1} \cdot Math_i + \delta_j + \lambda_t + \tau_j t + \varepsilon_{ijt} \tag{2}$$

In Equation 2, I've abbreviated the treatment variable—number of AP courses available—to $AP_{jt,t-1}$. Here, η_1 is the effect for a student with an average middle school math score, and $\eta_1 + \eta_3$ is the effect for a student with a math score one standard deviation above the mean. (Scores are standardized among the full population of test takers, within subject, year, and grade.) As a second approach to academic preparation heterogeneity, I sort students by their math score, subset the bottom 75 percent from the top 25 percent of performers, and run separate regressions. These percentiles are based on the analysis sample, not the original sample of middle school test takers.

After reporting how increasing AP offerings increases access, I examine the effects of AP course availability on college outcomes, using:

$$Y_{ijt} = \beta_0 + \beta_1 (\text{\# AP courses available})_{jt,t-1} + \delta_j + \lambda_t + \tau_j t + \varepsilon_{ijt} \tag{3}$$

where Y_{ijt} is the outcome of interest for student i graduating from school j in year t. The college outcomes I measure are (1) whether a student enrolled in any postsecondary institution within one year of high school graduation; (2) whether they enrolled at a two-year institution; (3) whether they enrolled at a four-year institution; (4) whether they enrolled at a college that is classified as competitive or higher by the Barron's selectivity index; (5) whether they earned a bachelor's degree within four years of graduating high school; and (6) whether they earned a bachelor's degree within six years. All college outcomes are unconditional on initial enrollment, so that students not attending college are assigned zeroes for all outcomes. I test for heterogeneity in changes to college outcomes with separate regressions by family income, race, and prior test score (bottom 75 vs. top 25 percent). β_1 in Equation 3 represents the intent-to-treat effect of AP course availability, which is a policy-relevant parameter for

4. In Michigan, the threshold for subsidized lunch is family income up to 185 percent of the federal poverty line. In 2019, this was equivalent to $47,638 for a family of four.

schools and districts considering introducing or expanding an AP program.

Another relevant treatment effect parameter would be the effect of an additional AP course for the students who actually take the course (the treatment effect on the treated). This suggests an instrumental variable (IV) strategy using course availability as an instrument for course-taking. However, the validity of the IV estimates relies on the exclusion restriction that the presence of AP courses at a school affects students only so far as it encourages them to take more AP courses and exams. This would be violated in the presence of within-school spillovers, such as positive spillovers of AP content and a more college-oriented culture, or negative spillovers due to diversion of resources. The direction of the bias here is theoretically ambiguous. For this reason, I consider the intent-to-treat effects more internally valid. Furthermore, as I show below, the first stage is on the margin of being considered too weak for valid IV estimation. With these caveats in mind, I implement a two-stage least squares (2SLS) strategy, described in more detail in section 5.5, for the overall sample as well as for the students with the strongest first stage.

4.2 Data

The data I use are provided by the Michigan Department of Education (MDE) and accessed through the Michigan Educational Data Center (MEDC). The Data Appendix (Appendix B) describes the various data sources, key variables, sample restrictions, and coding of transcript data in more detail.

My first data source is the Michigan High School Transcript Study (MTS), which attempted to collect longitudinal transcript data from a random sample of 150 Michigan public high schools. At the time I received access to the data, February 2017, the MTS research team had received data from 138 schools, but only 87 of those had provided the identifying information (name and birthdate) required to match students to the unique ID variable used in all other MEDC data sources. As I show in Section 5.1 and Table 1, the analytic sample includes schools that are somewhat larger, more urban, and higher-achieving than the state as a whole. The MTS dataset includes, for each school in the sample, every course taken by students at that school in a given year.

In order to measure the treatment I am interested in—AP courses available by school and AP courses taken by students—I systematically identified which courses were AP based on course title in the transcript data. The way in which schools list courses is not standardized across schools. Flagging courses as AP was an iterative process that started with more obvious course titles (e.g. "AP Calculus" or "Advanced Placement Biology") and continued by searching for other phrases associated with AP and with one of the recognized AP subjects (e.g. "AP CMP GOV" for comparative government and politics). While some courses were obviously AP, others were more ambiguous. If I wasn't reasonably sure a course was AP, I erred on the more conservative side and did not classify it as AP.[5] I assign

5. I provide a full list of AP course titles in Appendix B, and test sensitivity to classifying ambiguous courses as AP in Appendix Tables B4 and B5.

course availability at the school level and course-taking at the student level, counting by number of subjects.

For a subset of the students for whom I have course-taking data, I can also observe how many AP exams they took. MEDC has access to all AP exams taken by Michigan students between 2006 and 2013; these data come directly from the College Board. Since most students take AP in their junior and senior years, I can count AP exams for the classes of 2007 onward.[6]

To identify cohorts of high school seniors by school, I use demographic and enrollment data from the Michigan Student Data System (MSDS). This student-by-year panel dataset contains demographic information (including race and free and reduced-price lunch eligibility) as well as the school and district each student attends each year. I limit my sample to students who appear in both the MTS and MSDS.

For heterogeneity analysis by prior student achievement, I use K-12 student assessment data containing standardized test scores. I use a student's eighth grade math test score if it is available, and their seventh grade score if not. Test scores are not available for all students; they would be missing if the student attended middle school in a different state or at a private school, or if they were exempt from the test. The grades in which the state of Michigan tests students by subject have changed over time. I use math scores because the other subject tests were not offered in the relevant years for the full sample. These scores are standardized within subject (math), year, and grade. See the Data Appendix for more detail on test score data.

Information on college outcomes comes from the National Student Clearinghouse (NSC). The NSC provides information on college enrollment at and degrees awarded by any four- or two-year school in the country (with a few exceptions), by date of enrollment and institution. As of 2011, the NSC covered 95 percent of postsecondary institutions in Michigan (Dynarski, Hemelt, and Hyman 2015); it currently covers 98 percent of all students in U.S. institutions (National Student Clearinghouse 2023).[7]

My final sample includes 173,151 students who were seniors at 87 public Michigan high schools between 2005 and 2012.

5 Results

5.1 Descriptive Results

I begin with descriptive statistics about the students and schools in the sample, summarized in Table 1. Roughly half of the students are female. The majority, 75 percent, are white, 17 percent are Black, four percent are Asian, three percent are Hispanic, and fewer than one percent are Native (a category which

6. As is standard in the education literature, years refer to the spring of the academic year. For example, 2006 refers to the 2005-2006 school year.

7. For a detailed description of the NSC and its coverage, see Dynarski, Hemelt, and Hyman (2015).

includes American Indian, Alaska Native, Native Hawaiian, and Pacific Islander students). Given historical racial/ethnic differences in advanced course-taking and college attainment, for analyses by race and ethnicity I collapse the categories into underrepresented minority (URM) students (Black, Hispanic, and Native) and non-URM (white and Asian). Around a quarter of students in the sample are eligible for free or reduced-price lunch, which I use as a proxy for family income. At the school-cohort level, the average school in the panel enrolls around 1,400 students, has a student-to-teacher ratio of 21, spends $6,300 per student, and has a local unemployment rate of 9 percent.

As shown in the final two columns of Table 1, these means generally resemble the full population of Michigan students and high schools during this time. However, the students and schools in my sample are somewhat more advantaged than the state average. They are five percentage points less likely to be eligible for subsidized meals, and have middle school test scores 0.17 standard deviations higher than the state average. My sample of schools have higher average test scores, enroll more students, are less rural, and are in areas with lower unemployment than the average Michigan high school. There were 1,251 unique public high schools and 973,383 public high school seniors in Michigan over the 2005 to 2012 period; the schools and students in my sample are 7 percent and 18 percent of the statewide population, respectively. In any given year between 2005 and 2012, there are between 699 and 1,008 Michigan public high schools. The 87 schools in my sample represent 9 to 12 percent of high schools in a given year.

The average student in the sample has just under ten AP courses available to them during their junior and senior year, takes 0.79 courses, and takes 0.74 exams. The average school offers 8.56 AP courses to a cohort. I provide more detail on the variation in AP course offerings by school and across time, as well as AP course- and exam-taking, in Appendix Table B1 and Appendix Figures A1 through A9. Over time, the most common AP course offerings are English, Calculus, U.S. History, Biology, and Chemistry (see Appendix Table B1). The most common courses taken are English, Calculus, U.S. Government, Biology, and Psychology; and the most popular exams are English, Calculus, U.S. History, U.S. Government, and Biology.[8]

While the vast majority of schools offered at least one AP course to their juniors and seniors over the entire period, there is considerable variation in the number offered. The number of AP courses varies both across and within schools over time (see Appendix Figures A2, A3, and A4), and the changes go in both directions. My identifying variation comes from within-school increases and decreases in AP course offerings. These changes are driven by particular courses. The most common subjects to be introduced are Psychology, World History, Economics, Biology, and Statistics; the most likely to be taken away are Psychology, U.S. History, European History, Computer Science, and World

8. Recall that English and Calculus are each actually two separate courses: English Literature and Composition and English Language and Composition, and Calculus AB and BC. (See Appendix Figure B1.) Still, the hierarchy in Table B1 corresponds to national and Michigan AP exam data from the College Board.

History. The most marginal subjects—meaning those that experience the most changes in both directions—are Psychology, World History, and Economics. Appendix Table B2 lists the number of course changes by subject.

5.2 Effect of AP Course Availability on AP Course- and Exam-Taking

To explore whether and how students take advantage of expanded AP curricula, I estimate Equation 1 on the sample of seniors in Michigan public high schools. Table 2 shows the effect of AP course availability on both the extensive margin (probability of taking any AP course or exam) and the intensive margin (number of AP courses or exams taken). The point estimates suggest that an additional AP course offering increases the probability of a student taking any AP course by one percentage point, and the number of AP courses the average student takes by 0.032. There is a small, 0.4 percentage point increase on the extensive margin of exam-taking, but no detectable effect on the number of exams.

To put these magnitudes in context, the average student in my sample takes 0.79 AP courses, so the 0.032 effect on number of courses taken represents an increase in course-taking of four percent. Put differently, the average senior class has around 250 students, so these numbers translate into roughly eight additional AP course enrollments. The effect on the extensive margin (one percentage point) implies between 2 and 3 additional students taking AP who didn't previously; together, these effects mean that the additional courses are mostly taken by students already taking AP. It is notable that the effects on exams are smaller than the effects on courses, implying that many students induced into an additional AP course do not take the associated exam. Taking the point estimates at face value, the effect on number of exams (0.011) divided by the effect on courses (0.032) imply that fewer than 40 percent of marginal courses convert to exams.[9]

5.3 Heterogeneity in the Effect of AP Course Availability on AP Course- and Exam-Taking

Given documented inequities in the availability of AP by race and income (Solorzano and Ornelas 2002; Rodriguez and McGuire 2019) as well as tracking systems that segregate students within schools (Lewis and Diamond 2015), it is crucial to understand which types of students take advantage of expanded AP course offerings. As I show below, in Michigan there are large gaps in AP participation by family income, race, and academic preparation. For example, students from low-income households are half as likely to take any AP courses compared to their peers from higher-income families, and the typical student from a low-income family has an AP courseload a third the size of their more advantaged peer

9. Recall that the measures of AP courses and exams come from different data sources (high school transcripts for courses and College Board exam data for exams; see the Data Appendix for more detail). Courses are likely measured with more error than exams. This "conversion rate" is intended as a rough back-of-the-envelope calculation.

(see the group means in Table 3). The sizes of the gaps between URM and white and Asian students are very similar to the gaps by family income (see Table 4). In this section, I investigate whether a school offering more AP courses widens or shrinks gaps in participation.

Tables 3, 4, and 5 show effects on course- and exam-taking estimated by family income, race, and prior academic achievement. The estimates in Table 3 suggest that for students from both low- and higher-income families, an additional course offering increases the probability of taking any AP courses by around one percentage point. However, for economically disadvantaged students (those eligible for subsidized school meals), there is no detectable effect on the number of courses taken, and no effect on exam-taking. In contrast, non economically disadvantaged students increase their average number of AP courses significantly (0.038 courses), as well as the extensive and intensive margin of exam-taking (0.5 percentage point increase in any AP exam, and 0.018 increase in number of exams). Together, these results imply that most of the overall increases in Table 2 are due to students from higher-income households moving on the intensive margin of course-taking. Students from low-income families have much lower rates of AP participation (reflected in the lower group means in Table 3), so these results imply a widening of the income-based gap in the number of AP courses and exams.

The patterns by race in Table 4 are similar to the patterns by income. While both URM and non-URM students increase their probability of taking any AP courses by around one percentage point when an additional one becomes available, only white and Asian students significantly increase their average number of AP courses (by 0.036) and probability of taking an AP exam (by 0.4 percentage points). Since white and Asian students take more AP courses and exams, these results again imply a widening gap in AP participation.

Finally, Table 5 indicates that higher-achieving students are more likely to take advantage of additional AP courses. I show this with two alternative specifications. First, I estimate an interaction term between AP course availability and a student's middle school math test score (Panel A). For all four outcomes, the interaction is positive and significant, suggesting that more academically prepared students increase their course- and exam-taking more than their less prepared peers when additional AP courses are offered. Second, I split the sample into the bottom 75 versus top 25 percent of prior achievement (Panel B). (These percentiles are based on the analysis sample, not the original sample of middle school test takers.) Regardless of achievement, students increase the extensive margin of AP course-taking by approximately one percentage point. However, effects on the other outcomes diverge by prior achievement. While the bottom three-quarters of students do increase the number of AP courses they take by 0.023, the highest performing quarter of students increase the number of courses by more than three times as much, 0.074. The effect on the probability of taking any AP exam is also higher for higher achieving students (0.8 vs. 0.3 percentage points), as is the effect on the number of AP exams (0.048 for the top 25 percent and an insignificant 0.002 for the bottom 75 percent). The conversion rate from courses to exams is also much higher for this group: 0.048/0.074 = 65 percent.

The above results show that students—especially those from more advantaged groups—increase both their likelihood of any AP and the number that they take when more become available. Since the more advantaged, higher-achieving students have higher rates of AP participation to begin with, expanding AP offerings may primarily add one more AP to the transcripts of students already taking multiple, which may have a minimal marginal return. To investigate this further, I estimate a version of Equation 1 where, rather than looking at the number of AP courses students take, the left-hand side variables are indicators for exhaustive and mutually exclusive bins of AP courses taken: 0, 1-2, 3-4, and 5 or more. Appendix Table A1 reports the results, for all students and disaggregated by prior achievement (bottom 75 versus top 25 percent of middle school test scores). Panel A shows that overall, students are less likely to take no APs (this is the same as the positive "any AP" results above), and more likely to be in all of the positive categories, suggesting students are increasing across the distribution of courses taken. However, the results by prior achievement in Panel B shows that while lower-achieving students become more likely to take 1-2 or 3-4 AP courses (by 0.9 and 0.3 percentage points, respectively), for higher-achieving students only the increase in the top bin of 5 or more APs is significant (1 percentage point). The null effects on 1-2 and 3-4 courses for high achievers could mean some students shift into while others shift out of those bins, for no net change. On net, this implies that additional AP availability causes higher achieving students to add APs to an already high AP courseload. A correlational study by Beard et al. (2019) found a pattern consistent with diminishing marginal returns to AP, with no additional association between number of exams and BA attainment beyond 4-6 exams. If this is true in my sample, the results in Appendix Table A1 might lead us to expect an equalizing effect on later outcomes; I investigate this in the next section.

In a complementary analysis, I investigate the extent to which additional AP availability induces students already taking AP to shift *between* AP courses rather than take additional courses. If this is happening, then the above results are underestimating how much students react to newly available courses. I identify AP courses that are marginal or newly available to a given cohort of students within a high school (that is, offered to them but not the cohort prior) as opposed to inframarginal (also available to the prior cohort). I then regress the number of both marginal and inframarginal AP courses a student takes on the count of newly available AP courses. In this exercise, the treatment is newly available AP courses; the outcomes are the number of new and old AP courses a student takes. The results, in Appendix Table A2, find a very small negative and statistically insignificant decrease (-0.008 courses) in the number of inframarginal AP courses a student takes. This suggests that new AP courses do not simply lead students to shift which AP courses they take; rather, they strictly increase their AP course load.

Together, these results suggest that additional AP course availability does induce a small proportion (about one percentage point) of students of all backgrounds to cross the extensive margin of AP participation. However, the students who were already taking the most APs—economically advantaged, white and Asian, and higher-achieving students—increase their total number of AP courses and exams more, thus widening gaps.

5.4 Effect of AP Course Availability on College Enrollment and Graduation

How do the changes to AP course-taking affect students' longer-term educational attainment? To measure the effect of AP course availability on longer-term college outcomes, I estimate Equation 3. The estimates appear in Table 6. On average, there is no effect of an additional AP course offering on any of the outcomes; all of the treatment coefficients are close to zero, none are significant, and they are estimated precisely. For example, the effect on enrolling in any college is 0.2 percentage points, with a standard error of 0.2 percentage points. The effects on four-year and six-year BA attainment are both a statistically insignificant 0.1 percentage points. For all outcomes, I can rule out (with 95 percent confidence) effects greater in magnitude than a single percentage point.[10]

There is reason to believe that the effects of AP vary by student type. As I showed above, certain types of students are more likely to take AP in the first place. Even conditional on participation, more academically prepared students might be more likely to reap the benefits of a college-level curriculum. Less prepared students might have more to gain from more rigorous courses; on the other hand, they could fall further behind their peers if they're pushed into courses beyond their preparation level, or might be harmed by the diversion of resources towards AP students and teachers. The structural inequities discussed above may prevent economically disadvantaged and underrepresented minority students from accessing advanced classes, as well as contribute to their lower levels of academic preparation. To test this, I estimate effects of AP course availability on college outcomes, this time estimating effects by family income, student race, and academic performance in middle school.

Table 7 displays the effect of AP availability by family income, which I proxy by eligibility for subsidized school meals. The group means (in square brackets for each group) indicate that students from low-income households have lower rates of all college outcomes. Compared to non-economically disadvantaged students, they are much less likely to attend college at all (gap of 19 percentage points) or a four-year college in particular (gap of 21 percentage points). Economically disadvantaged students are slightly more likely (by 2 percentage points) to attend a two-year institution. They are also half as likely to attend a selective college (23 vs. 45 percent) and a third as likely to earn a BA in six years (13

10. In an attempt to limit the number of outcomes, I collapsed institutions considered competitive, very competitive, highly competitive, and most competitive into the category of "competitive+." In Appendix Table A3, I estimate effects on each of the six Barron's categories (which include the four just listed, as well as non-competitive and less competitive.) All point estimates are small, between -0.1 and 0.2 percentage points. The effect on competitive enrollment, 0.2 percentage points, is significant at the 10 percent level, as is the effect on very competitive enrollment, -0.2 percentage points.

vs. 38 percent).

There is no evidence that additional AP course offerings improve college enrollment or graduation for students in low- or higher-income households. All of the effects are close to zero, precisely estimated, and statistically insignificant. Even if we took the point estimates at face value, they would translate into not even one additional low- or higher-income student per school enrolling in college. (The average cohort has 61 economically disadvantaged students and 192 non-economically disadvantaged students in its senior class.) Examining heterogeneity by race (Table 8) suggests a similar story. All estimated effects are small (less than half of a percentage point) and statistically insignificant.

I also examine heterogeneity by prior academic achievement, which I measure using standardized scores on the state middle school math test. Table 9 summarizes the effect of AP course availability on college outcomes by prior test scores, where I have again split the sample into the bottom 75 and top 25 percent of test scores. Here, there is some evidence of positive effects for the most academically prepared students. While all of the effects on the bottom 75 percent of test takers are null, with an additional available AP course the highest achieving students increase their probability of enrolling in a four-year college by 0.5 percentage points and the probability of enrolling in a competitive or higher college by the same magnitude. They also increase their chances of earning a BA in four years by 0.7 percentage points. The effect on six-year graduation is a statistically insignificant 0.4 percentage points, implying that more AP courses allowed some high-achieving students to decrease their time to degree if not change their ultimate educational attainment.

The above analyses imply that expanding AP course availability has little discernible effect on outcomes for most students. The exception is a small but significant effect on enrollment and graduation outcomes for students who enter high school with the strongest academic preparation; though there may be small positive effects for these students, it would come at the expense of widening existing achievement gaps.

5.5 Two-Stage Least Squares Analysis: Effect of AP Course-Taking on College Enrollment and Graduation

In section 5.4, I found that when schools offer more AP courses, the postsecondary outcomes (enrollment and graduation) of only the highest-achieving students improve. These are also the students who increased their course-taking the most. In this section, I use an IV (2SLS) approach to estimate the effect of an additional AP course for the high-achieving students induced to increase their AP courseload when it becomes available at their school. I also explore different ways to define "high-achieving" with alternative cuts of the test score distribution; doing so provides further evidence on which students take advantage of AP and which ones benefit the most from doing so.

Interpreting 2SLS estimates causally requires the standard assumptions. The exclusion restriction requires that the presence of AP courses at a school affects students only so far as it encourages them to take more AP courses and exams. If changes to AP courses are not due to exogenous factors such as teacher retirements and are instead due to underlying changes to students and families in a school, this would be violated. Spillovers within a school that affect students not taking AP, such as changes to college-going culture or diversion of resources, would also violate the exclusion restriction. Interpreting the parameter as a local average treatment effect (LATE) on students induced to take an additional AP course further requires that there are no defiers: students who take *fewer* AP courses when an additional one is offered. Violations of these assumptions are theoretically possible but untestable. Thus, I consider the results in this section to be suggestive rather than definitive.

Table 10 shows the results of the 2SLS analysis. Each column includes estimates for a different sample: the entire sample, then different top X percents of the sample by test score, where $X = 50$, 25, 10, and 5. Panel A reveals how the first stage, i.e., take-up of expanded AP availability, varies by prior achievement. The more selective the sample (e.g., looking at the top quarter of students compared to the top half), the stronger the change to AP course-taking. As we saw in Table 2, an additional AP course available within a school leads the typical student to increase their AP courseload by 0.032 courses. An above-average (top 50 percent) student increases by 0.048 courses; a student in the top quarter, by 0.74 courses; a student in the top 10 percent, by 0.117 courses; and a student in the top five percent by a similar 0.108 courses.

IV analysis also requires that the instrument—in this case, AP availability—have sufficiently strong correlation with the treatment (in this case, AP course-taking). For a model with one endogenous regressor and one instrument, the critical value for a weak instrument test with a 5 percent significance level and a test of 10 percent maximal size is 16.38; for a 15 percent maximal size, the critical value is 8.96 (Stock and Yogo 2005).[11]. The older rule of thumb proposed by Staiger and Stock (1997) is an F-statistic of at least ten. Thus, depending on the tolerable level of false rejection, the first stage is sufficiently strong for the overall sample, the top 25 percent, and top 10 percent, which have F-statistics above ten.

The reduced form and LATE effects in Panels B and C of Table 10 show a different pattern by achievement than the first stage. The reduced form effects—i.e., the effect of additional available courses on college outcomes—are largest and only statistically significant for students in the top 25 and 10 percent of test scores. (Note that the reduced form effects on all students and the top 25 percent are identical to those reported in Tables 6 and 9.) The top 25 percent of students increase their four-year college enrollment by 0.5 percentage points, competitive enrollment by 0.5 percentage points, and

11. As summarized by Baum, Schaffer, and Stillman (2007), the maximal size refers to the false rejection rate of a hypothesis test that a researcher is willing to tolerate: "Under weak identification, the Wald test rejects too often. The test statistic is based on the rejection rate r (10%, 20%, etc.) that the researcher is willing to tolerate if the true rejection rate should be the standard 5%. Weak instruments are defined as instruments that will lead to a rejection rate of r when the true rejection rate is 5%."

four-year BA attainment by 0.7 percentage points for each additional available course. For the top ten percent, they increase their six-year BA attainment by 0.6 percentage points for each additional available course. (The reduced form effects for the top five percent sample are similar to or larger than those for the top ten percent, but the smaller sample means none are statistically significant.)

Correspondingly, the estimated LATEs—i.e., the effect of *taking* an additional available course—are significant only for these groups, as well. For a student in the top quarter of the test score distribution who was induced to take an additional AP course, doing so increases their probability of enrolling in a four year college by 6.9 percentage points, and in enrolling in a competitive or higher institution by a similar magnitude (6.4). The effect on earning a BA in four years is 10 percentage points. For a student in the top ten percent, taking an additional AP increases their chances of four-year BA receipt by 6.2 percentage points and their chances of six-year BA receipt by 5.6 percentage points. Again, the effects for the top five percent are similar but less precise.

These results suggest that not only do the highest-achieving (top quarter) of students take more APs when they become available, but they are also more likely to benefit when they do. Although I don't find any effects on the extensive margin of college enrollment, additional AP courses seem to shift high-achieving students into higher quality colleges, speed up time to degree, and possibly increase ultimate (six-year) BA receipt. However, given the weaker first stage effects for the full and top 50 percent samples, I cannot say for certain whether lower-achieving students would benefit if they could be induced into taking more AP courses.

5.6 Additional Heterogeneity Analyses

In this section, I explore two different dimensions of AP course offerings. The first is school-level heterogeneity in the initial strength of the AP program. It is not clear ex-ante whether schools with an initially weaker or more robust AP program would experience larger changes from further expansion. If there are diminishing returns to AP, the former set of schools may benefit more; on the other hand, the latter set of schools might have infrastructure and experience that make additional APs more successful. To test this, I characterize schools by whether they offered fewer than five versus five or more AP courses in the first year of the panel. I then estimate Equations 1 and 3 separately for students in the two groups of schools. Appendix Table A4 shows the effects on course- and exam-taking, and Appendix Table A5 shows the effects on college outcomes. Appendix Table A4 reveals a similar pattern as the student-level heterogeneity in Section 5.3. Students in both types of schools cross the extensive margin of AP course-taking when the number of available courses increases. However, in the schools with initially robust AP programs and higher AP participation, the number of AP courses students take increase more, thus widening participation gaps. The effects on college outcomes by initial AP availability in Appendix Table A5 are generally null for both groups. The exception is a positive 1.2 percentage point increase in any college enrollment for students in schools with a low initial number of APs. It appears that additional AP course offerings may make more of a difference in preparing students for college (academically or mentally) in schools that historically offer fewer APs and send fewer students to college.

In all of the above analysis, I have grouped all types of AP courses together. The AP curriculum spans over 30 subjects, ranging from studio art to languages to computer science. It is possible that different subjects affect students' outcomes differently. This could be because performance in different subjects provide different signals to students and colleges about a student's college readiness, or because different subjects are more likely to earn college credit. To test this, I disaggregate the AP course offering variable (number of AP courses available in a student's junior and senior year) into STEM courses (which include all math, science, and computer science subjects) and non-STEM. I then estimate a version of Equation 3 with two treatment variables:

$$\begin{aligned} \mathrm{Y}_{ijt} &= \beta_0 + \beta_1(\text{\# AP STEM courses available})_{jt,t-1} \\ &+ \beta_2(\text{\# AP non-STEM courses available})_{jt,t-1} + \delta_j + \lambda_t + \tau_j t + \varepsilon_{jt} \end{aligned} \tag{4}$$

The results appear in Appendix Table A6. The effects of both STEM and non-STEM AP offerings are generally small and statistically insignificant, and similar to each other. Out of 12 hypothesis tests (two subject groups times six outcomes), I find one effect that is significant at the 10 percent level: a 0.2 percentage point increase in four-year BA attainment for an additional non-STEM AP.

I further disaggregate courses into the six groupings used by the College Board: English, science, math and computer science, history and social sciences, languages, and arts. For this, I estimate a version of Equation 4 with six AP availability variables, one for each subject group (number of AP art courses available, number of AP English, number of AP languages, etc.). The results are in Appendix Table A7. In general, none of the subject-specific effects on college outcomes are statistically significant; the exception is a 0.8 percentage point effect of language APs on four-year college enrollment and a 0.6 percentage point effect of art APs on four-year graduation. Both are significant only at the $\alpha = 0.10$ level. Given the number of tests (six outcomes times six subjects = 36) in Appendix Table A7, I do not put much weight on these differences.

6 Threats to Identification and Robustness Checks

Because I am not able to randomly assign schools to offer AP courses, I have to worry about whether my results are picking up a true causal effect or are driven by some spurious correlation. There are several main threats to identification. Perhaps both AP participation and gaps in college enrollment are growing over time, but the former is not causing the latter. My empirical strategy addresses this in several ways: first, I include year fixed effects to allow for a time trend in college outcomes, and capture deviations from trends: in years when a school has a larger change in AP offerings, do student outcomes experience a correspondingly large change? Second, as I show in Appendix Figure A4, while there is an upward trend in number of AP courses at most schools, it is by no means strictly monotonic, meaning I am identifying off of changes in AP offerings in both directions. A related issue of confounding endogeneity is that it is possible that longer-term, systematic changes to the student population and the demand for AP courses are occurring, and that these are correlated with student outcomes. This would be the case if, for example, schools offer more AP courses in order to attract higher-achieving students. I test for this type of endogeneity in two ways.

First, I re-estimate all effects on course- and exam-taking and college outcomes with additional controls for student- and school-level characteristics. Student characteristics include sex, race (indicators for white, Asian, Black, Hispanic, and Native), free or reduced-price lunch eligibility in 12th grade, and standardized score on the middle school math test. I also add time-varying school characteristics: average middle school math test score, school size, student-to-teacher ratio, per-student spending, and local unemployment; these are measured in the student's sophomore year so that they are unaffected by the treatment. The sample means of all additional control variables are reported in Table 1. Versions of Tables 2 through 9 estimated with additional student- and school-level controls are included as Appendix Tables A8 through A15. The results are nearly identical. Of course, time-varying unobservable factors could still be driving the relationship between AP availability and student outcomes, and the results might change if I were able to control for all relevant factors not picked up by school fixed effects. Nevertheless, it is reassuring that the results are insensitive to a rich set of student and school controls. Furthermore, most of the effects on college outcomes that I estimate are null, while most plausible stories about selection (such as a new school principal who changes parent and student attitudes about college) would imply an upward bias.

As a further check, I estimate effects on course- and exam-taking by income and race using a single regression where I interact course availability with an indicator for either low family income or URM status, and include all of the student- and school-level controls. This specification, shown in Appendix Tables A16 and A17, leads to the same conclusion: students from higher-income families and white and Asian students increase their AP participation more when new courses become available, even accounting for academic preparation. This result differs from Conger, Long, and Iatarola (2009), who find that gaps in advanced course-taking reverse in sign after conditioning on eighth grade test scores.

I further test for the sensitivity of results to alternate specifications, by including or excluding various combinations of student controls, school controls, and school-specific time trends. Appendix Table A18 shows estimates of overall effects on college outcomes using alternative specifications; Appendix Table A19 does so only on the highest-achieving 25 percent of the sample, since this is the subgroup with the strongest results. Appendix Table A18 shows that the overall effects on college outcomes (competitive enrollment and four- and six-year BA attainment) are similar for any combination of controls, i.e., small in magnitude and statistically insignificant. However, specifications with no school-specific time trends and no other controls result in positive, significant effects on competitive enrollment (0.4 percentage points) and four-year degree receipt (0.2 percentage points). The fact that these effects disappear with any additional controls suggests, first, that there may be some bias not captured by school fixed effects alone, which school-specific time trends and/or additional controls account for; and second, that controlling for school-specific time trends is sufficient, since additional controls don't further change the estimates. Appendix Table A19 repeats this robustness check, limiting to the highest-achieving 25 percent of students. The positive, significant effect on competitive college enrollment for this group holds up under almost all sets of controls and school-specific time trends (or lack thereof). While the effect is not statistically significant with the full set of controls and school-specific time trends, (specification (4)), the point estimate is nearly identical to the other specifications. The positive effect on four-year BA receipt for this high-achieving group is significant only when school-specific linear time trends are included (with and without other controls). While my preferred specification is Equation 3, the sensitivity of these results to controls warrants some caution.

As another robustness check, I directly test for positive selection of students into schools with more AP courses by estimating a version of Equation 3 where the left-hand-side variable is the average middle school test score of the senior class:

$$\text{(Average middle school test score)}_{jt} = \alpha + \sum_{k=-2}^{2} \beta_k AP_{j,t+k} + \delta_j + \lambda_t + \tau_j t + \varepsilon_{jt} \tag{5}$$

Note this is done at the school-year level. Positive β_k's, particularly for $k \leq 0$, would suggest that a stronger AP curriculum attracts higher-achieving students, and would cause me to worry that my findings are driven by students with better outcomes coming into schools with more AP rather than more AP causing improved outcomes. Figure 1 graphically depicts the estimated β_k coefficients. There is no evidence that higher-achieving students are positively selecting into schools with more AP courses.

6.1 Corrections for Two-Way Fixed Effects Estimates

Several recent papers have highlighted potential issues with linear regressions that estimate policy treatment effects using time and group fixed effects (two-way fixed effects, or TWFE), such as Equation 3 above (see de Chaisemartin and D'Haultfœuille (2023) and Roth et al. (2023) for reviews). TWFE approaches may, unless researchers are willing to make implausibly strong assumptions, produce estimates that are misleading or hard to interpret. The key problem in extending the canonical two-period difference-in-difference design with a binary treatment to an equation more like Equation 3 comes from what both sets of authors refer to as "forbidden comparisons." A treatment parameter from a TWFE model is a weighted average of all possible comparisons of groups experiencing different changes to the treatment, including comparing groups whose treatment (e.g., number of AP courses available) change more relative to those who change less. If the lower-treated group has a larger per-unit treatment effect, such comparisons can result in a negative effect, even if the effect of the treatment is positive for both groups.

de Chaisemartin and D'Haultfœuille (2020) propose an alternative estimator, which they call DID_M, which eliminates "forbidden comparisons" and averages, across groups and time, all comparisons of groups whose treatment changes to groups whose treatment doesn't change. More specifically, DID_M is the "weighted average of DID terms comparing the evolution of the outcome in groups whose treatment went from d to d' between $t-1$ and t and in groups with a treatment of d at both dates, across all possible values of d, d', and t" (de Chaisemartin and D'Haultfœuille 2020, p. 2981). For example,"switcher" schools that went from offering 5 to 6 AP courses from 2007 to 2008 would be compared to schools that offered 5 APs in both years ("stayers"). In this approach, switchers (i.e., schools going from d to d' AP offerings between a given $t-1$ and t) without a corresponding stayer (a school with d APs in both $t-1$ and t) are not used in estimation (and vice versa). Fewer comparisons means that DID_M estimates tend to be less precise than TWFE ones.

I re-estimate the effects in Tables 2 through 5 (effects of course availability on course- and exam-taking) and Tables 6 through 9 (effects of course availability on college outcomes) using the DID_M approach, implemented with the did_multiplegt Stata command (de Chaisemartin, D'Haultfœuille, and Guyonvarch 2019). The results of this alternative approach are included as Appendix Tables A20 through A23 (effects on course-taking) and Appendix Tables A24 through A27 (effects on college outcomes). Because they leverage fewer comparisons, the DID_M estimates are substantially noisier than the TWFE estimates. However, the magnitudes lead to similar conclusions.

Although all of the first stage results using the approach of de Chaisemartin and D'Haultfœuille (2020) (Appendix Tables A20 through A23) are noisier than the TWFE estimates, the magnitudes are similar and again suggest that non economically disadvantaged, non-underrepresented minority, and higher-achieving students are more likely to take advantage of newly offered AP courses. The strongest and only statistically significant DID_M estimate is the effect on the top 25 percent by middle school test score (in Appendix Table A23), who increase their number of AP courses by 0.077 when an additional one becomes available. This is similar to the TWFE estimate of 0.074 reported in Table 5.

The estimates of DID_M for college outcome effects are consistent with the TWFE estimates, but ultimately too imprecise to provide additional evidence that additional AP courses improve outcomes overall or for specific subgroups. Like the majority of the TWFE estimates, none of the DID_M effects in Appendix Tables A24 through A27 are statistically significant. The positive effects found with TWFE estimates for the highest achieving students—increases to four-year and competitive college enrollment and on-time BA receipt—are the same sign but no longer significant using DID_M. The TWFE estimates on four-year enrollment, competitive enrollment, and four-year BA receipt were 0.5, 0.5, and 0.7 percentage points, respectively (Table 9); the DID_M estimates are 0.9, 1.1, and 0.3 percentage points (Appendix Table A27). However, the confidence intervals around the DID_M estimates all contain the equivalent TWFE estimates, so I cannot reject that they are the same.

7 Conclusion

Using administrative data from the state of Michigan and exploiting within-school, across-time variation in AP course offerings, I have shown that introducing more AP courses fails to close gaps in access and outcomes. When schools increase the number of AP courses available, a small proportion (one percent) of students of all backgrounds cross the extensive margin of AP participation. However, the more advantaged students—higher achieving, non-URM, and those from higher income families—increase their already higher average AP courseload by more than their disadvantaged peers. This finding is consistent with work by historians, sociologists, and education researchers arguing that the AP program, like many other examples of educational resources, benefits already privileged students and systematically excludes the already marginalized, thus perpetuating inequities (Solorzano and Ornelas 2002; Schneider 2009; Lewis and Diamond 2015; Rodriguez and McGuire 2019). These studies, as well as the current analysis, suggest that without a concerted effort to ensure equal access for all students, expanding AP offerings will most likely only worsen educational inequality.

Even if students were granted truly equal access to AP courses, it is not obvious that college outcome gaps would close. I find very limited evidence that access to additional advanced courses improves college enrollment, college quality, or postsecondary attainment. Although my primary results suggest that the most academically prepared students may benefit from AP in terms of quality of initial college enrollment and on-time BA receipt, this finding doesn't hold under all alternative estimation approaches. Even if there is a benefit for high-achieving students, it would only serve to widen existing gaps.

Despite a push by some policymakers to use AP courses as a tool for combating inequality and improving college readiness, the current study complements recent research (Conger, Long, and McGhee 2023) showing that expanding AP access is unlikely to do so, at least not without additional incentives or supports. In both my setting—Michigan schools making year-to-year adjustments in AP offerings—and that of Conger, Long, and McGhee (2023)—a national set of schools that had never offered AP science adding it to the curriculum—the program largely failed to deliver the outcomes its proponents espouse.

The causal evidence on the AP program is not universally negative. Jackson (2010, 2014) found positive achievement and college completion effects of paying students to pass AP exams; Smith, Hurwitz, and Avery (2017) found that passing an AP exam improved on-time college graduation. However, the positive effects in these cases are tied to students taking and passing exams, not simply enrolling in AP courses. This is consistent with the positive effects on four-year graduation that I find for the highest-achieving students only, who are also much more likely to take AP exams. (Smith, Hurwitz, and Avery (2017) also find no heterogeneity in graduation effects by income or race, which is consistent with my own results.) In settings where a school or student on the margin of offering or taking AP is relatively disadvantaged, most students are unlikely to benefit without an additional push to take and succeed on the exams.

The policy implications from the current as well as previous work are similar: putting financial and legal resources towards expanding AP access is, by itself, unlikely to achieve the goal of closing gaps in educational outcomes. If educators and policymakers strive to address educational inequality, additional resources focused on AP exams are likely necessary, and may be best targeted more explicitly at disadvantaged students. In the longer term, policies to address earlier differences in academic achievement may also allow more students to benefit from AP curricula.

References

Avery, Christopher, and Joshua Goodman. 2022. Ability signals and rigorous coursework: evidence from AP Calculus participation. *Economics of Education Review* 88:102237.

Avery, Christopher, Oded Gurantz, Michael Hurwitz, and Jonathan Smith. 2018. Shifting college majors in response to Advanced Placement exam scores. *Journal of Human Resources* 53 (4): 918–956.

Baum, Christopher F, Mark E Schaffer, and Steven Stillman. 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *The Stata Journal* 7 (4): 465–506.

Beard, Jonathan J, Julian Hsu, Maureen Ewing, and Kelly E Godfrey. 2019. Studying the relationships between the number of APs, AP performance, and college outcomes. *Educational Measurement: Issues and Practice* 38 (4): 42–54.

Betts, Julian R. 2011. The economics of tracking in education. In *Handbook of the economics of education,* 3:341–381. Elsevier.

Burnett, Alyson, and Paul Burkander. 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.

Chatterji, Roby, Neil Campbell, and Abby Quirk. 2021. Closing advanced coursework equity gaps for all students. Center for American Progress.

College Board. 2003. A brief history of the Advanced Placement program. Retrieved from https://ca01000794.schoolwires.net/cms/lib/CA01000794/Centricity/Domain/295/APUSH/AP_History_history.pdf.

———. 2021a. Annual AP program participation 1956-2021. Retrieved from https://reports.collegeboard.org/media/pdf/2021-annual-ap-program-participation_1.pdf.

———. 2021b. Program summary report. Retrieved from https://reports.collegeboard.org/media/pdf/2021-ap-program-summary-report_1.pdf.

Conger, Dylan, Alec I Kennedy, Mark C Long, and Raymond McGhee. 2021. The effect of Advanced Placement science on students' skills, confidence, and stress. *Journal of Human Resources* 56 (1): 93–124.

Conger, Dylan, Mark C Long, and Patrice Iatarola. 2009. Explaining race, poverty, and gender disparities in advanced course-taking. *Journal of Policy Analysis and Management* 28 (4): 555–576.

Conger, Dylan, Mark C Long, and Raymond McGhee Jr. 2023. Advanced Placement and initial college enrollment: evidence from an experiment. *Education Finance and Policy* 18 (1): 52–73.

Darolia, Rajeev, Cory Koedel, Joyce B Main, J Felix Ndashimye, and Junpeng Yan. 2020. High school course access and postsecondary STEM enrollment and attainment. *Educational Evaluation and Policy Analysis* 42 (1): 22–45.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review* 110 (9): 2964–96.

———. 2023. Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey. *The Econometrics Journal* 26 (3): C1–C30.

de Chaisemartin, Clément, Xavier D'Haultfœuille, and Yannick Guyonvarch. 2019. Did_multiplegt: DID estimation with multiple groups and periods in Stata. *URL: https://ideas. repec. org/c/boc/bocode/s458643. html.*

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in Kenya. *American Economic Review* 101 (5): 1739–1774.

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 (1_suppl): 53S–79S.

Fazlul, Ishtiaque, Todd Jones, and Jonathan Smith. 2021. College credit on the table? Advanced Placement course and exam taking. *Economics of Education Review* 84:102155.

Gonzalez, Naihobe. 2017. How learning about one's ability affects educational investments: evidence from the Advanced Placement program. Mathematica Policy Research Working Paper No. 52.

Goodman, Sarena. 2016. Learning from the test: raising selective college enrollment by providing information. *Review of Economics and Statistics* 98 (4): 671–684.

Gurantz, Oded. 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.

Hochman, William R. 1970. Advanced Placement: can it change with the times? *College Board Review,* no. 77, 16–19.

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.

———. 2014. Do college-preparatory programs improve long-term outcomes? *Economic Inquiry* 52 (1): 72–99.

Jacob, Brian A, and Tamara Wilder. 2010. Educational expectations and attainment. NBER Working Paper No. 15683.

Klopfenstein, Kristin. 2010. Does the Advanced Placement program save taxpayers money? The effect of AP participation on time to college graduation. In *AP: a critical examination of the Advanced Placement program,* edited by Philip M Sadler, Gerhard Sonnert, Robert H Tai, and Kristin Klopfenstein, 189–218. Harvard Education Press.

Klopfenstein, Kristin, and Kit Lively. 2012. Dual enrollment in the broader context of college-level high school programs. *New Directions for Higher Education* 2012 (158): 59–68.

Klopfenstein, Kristin, and M Kathleen Thomas. 2010. Advanced Placement participation: evaluating the policies of states and colleges. In *AP: a critical examination of the Advanced Placement program,* edited by Philip M Sadler, Gerhard Sonnert, Robert H Tai, and Kristin Klopfenstein, 167–188. Harvard Education Press.

Lewis, Amanda E, and John B Diamond. 2015. *Despite the best intentions: how racial inequality thrives in good schools.* Oxford University Press.

Malkus, Nat. 2016. The AP peak: public schools offering Advanced Placement, 2000-12. American Enterprise Institute.

Marascuilo, Leonard A, and Maryellen McSweeney. 1972. Tracking and minority student attitudes and performance. *Urban Education* 6 (4): 303–319.

National Student Clearinghouse. 2023. StudentTracker for educational organizations. URL: https://www.studentclearinghouse.org/educational-organizations/studenttracker-for-educational-organizations/Accessed June 20, 2023.

Quinton, Sophie. 2015. Closing the achievement gap, one school at a time. *The Atlantic* and *National Journal*, 30 January. https://www.theatlantic.com/politics/archive/2015/01/closing-the-achievement-gap-one-school-at-a-time/431826/.

Rodriguez, Awilda, and Keon M McGuire. 2019. More classes, more access? Understanding the effects of course offerings on Black-White gaps in Advanced Placement course-taking. *The Review of Higher Education* 42 (2): 641–679.

Rodriguez, Awilda, Mary EM McKillip, and Sunny X Niu. 2013. The earlier the better? Taking the AP in 10th grade. College Board Research Report No. 2012-10.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2023. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics* 235 (2): 2218–2244.

Rothschild, Eric. 1999. Four decades of the Advanced Placement program. *The History Teacher* 32 (2): 175–206.

Schneider, Jack. 2009. Privilege, equity, and the Advanced Placement program: tug of war. *Journal of Curriculum Studies* 41 (6): 813–831.

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.

Solorzano, Daniel G, and Armida Ornelas. 2002. A critical race analysis of Advanced Placement classes: a case of educational inequality. *Journal of Latinos and Education* 1 (4): 215–229.

Staiger, Douglas, and James H Stock. 1997. Instrumental variables regression with weak instruments. *Econometrica* 65 (3): 557–586.

Stock, James H, and Motohiro Yogo. 2005. Testing for weak instruments in linear IV regression. In *Identification and inference for econometric models: essays in honor of Thomas Rothenberg,* edited by Donald W.K. Andrews and James H. Stock, 80–108. Cambridge University Press.

Theokas, Christina, and Reid Saaris. 2013. Finding America's missing AP and IB students. The Education Trust Shattering Expectations Series. Retrieved from https://edtrust.org/wp-content/uploads/2013/10/Missing_Students.pdf.

Tugend, Alina. 2017. Who benefits from the expansion of AP classes? *The New York Times Magazine*, 7 September. https://www.nytimes.com/2017/09/07/magazine/who-benefits-from-the-expansion-of-ap-classes.html.

U.S. Department of Education. 2016. U.S. Education Department awards 41 states and the District of Columbia $28.4 million in grants to help students from low-income families take Advanced Placement tests. Press release. Retrieved from https://www.ed.gov/news/press-releases/us-education-department-awards-41-states-and-district-columbia-284-million-grants-help-students-low-income-families-take-advanced-placement-tests.

Table 1: Sample Descriptive Statistics

	Analysis Sample			Michigan	
	Mean	Std dev	N non-missing	Mean	N non-missing
A. Student level characteristics					
Female	0.51	0.50	173,077	0.50	973,383
White	0.75	0.44	173,151	0.76	973,910
Black	0.17	0.38	173,151	0.17	973,910
Asian	0.04	0.20	173,151	0.02	973,910
Hispanic	0.03	0.18	173,151	0.04	973,910
Native	0.01	0.09	173,151	0.01	973,910
Eligible for free or reduced-price lunch	0.24	0.43	173,151	0.29	971,772
Middle school math test score (std.)	0.27	1.01	154,549	0.10	841,906
AP courses available junior & senior year	9.79	4.57	173,151	-	-
AP courses taken junior & senior year	0.79	1.38	173,151	-	-
AP tests taken	0.74	1.59	135,272	-	-
B. School-cohort level characteristics					
Average middle school math test score	0.08	0.43	687	-0.18	6,144
School enrollment	1377	500	689	685	5,780
Town or rural location	0.23	0.42	689	0.48	5,798
Pupil-to-teacher ratio	21.41	10.97	686	21.58	5,566
Per pupil instructional spending	6360	1783	686	6433	5,357
Local unemployment rate	8.87	4.61	689	9.50	5,798
AP courses available year t and $t-1$	8.56	4.64	689	-	-

Notes: In Panel A, the unit of observation is a single student. In Panel B, the unit of observation is a school-by-cohort. "Native" includes American Indian, Alaska Native, Native Hawaiian, and other Pacific Islander students. Middle school math test score is measured as a standardized scale score, standardized on the full population of test takers within year, grade, and subject. I use eighth grade test score if available and seventh grade score if not. School-year characteristics are all measured in year $t-2$, except for AP course availability. AP course availability is measured as the number of unique AP subjects offered over two years; if a subject is offered both years, it is counted once. The full population of Michigan seniors is based on administrative enrollment data. The full population of public Michigan high schools is based on the Common Core of Data.

Table 2: Effect of AP Course Availability on AP Course- and Exam-Taking

	(1) Any AP courses taken	(2) # AP courses taken	(3) Any AP exams taken	(4) # AP exams taken
# of AP courses available at school	0.011***	0.032***	0.004**	0.011
in junior and senior year	(0.003)	(0.009)	(0.002)	(0.008)
Mean of outcome variable	[0.351]	[0.786]	[0.274]	[0.736]
Kleibergen-Paap Wald F statistic	15.48	11.65	3.99	2.18
Observations	173,151	173,151	135,272	135,272
Cohorts	2005-2012	2005-2012	2007-2012	2007-2012

Notes: Table reports estimate of α_1 in Equation 1, which regresses student-level course- and exam-taking on AP course availability. AP availability is measured at the school-by-cohort level, and counts the number of AP courses available to a high school cohort in their junior and senior year. Regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 3: Effect of AP Course Availability on AP Course- and Exam-Taking, by Family Income

	(1) Any AP courses taken	(2) # AP courses taken	(3) Any AP exams taken	(4) # AP exams taken
Effect of # available AP courses for:				
Students from low-income families	0.009**	0.010	0.000	-0.008
	(0.004)	(0.008)	(0.003)	(0.007)
Group mean	[0.196]	[0.364]	[0.139]	[0.297]
Observations	41,974	41,974	35,378	35,378
Students from non-low-income families	0.011***	0.038***	0.005**	0.018**
	(0.003)	(0.011)	(0.002)	(0.009)
Group mean	[0.401]	[0.921]	[0.322]	[0.892]
Observations	131,177	131,177	99,894	99,894
Cohorts	2005-2012	2005-2012	2007-2012	2007-2012

Notes: Low-income status is proxied by eligibility for free or reduced-price lunch (FRPL). Effects by income are estimated with separate estimations of Equation 1 by FRPL status. AP availability is measured at the school-by-cohort level, and counts the number of AP courses available to a high school cohort in their junior and senior year. Regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. Group-level means of the course- and exam-taking variables are in brackets. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 4: Effect of AP Course Availability on AP Course- and Exam-Taking, by Race and Ethnicity

	(1) Any AP courses taken	(2) # AP courses taken	(3) Any AP exams taken	(4) # AP exams taken
Effect of # available AP courses for:				
Black, Hispanic, & Native students	0.012**	0.012	0.002	0.003
	(0.005)	(0.011)	(0.003)	(0.008)
Group mean	[0.202]	[0.362]	[0.125]	[0.248]
Observations	37,018	37,018	29,041	29,041
White & Asian students	0.010***	0.036***	0.004**	0.014
	(0.003)	(0.011)	(0.002)	(0.009)
Group mean	[0.392]	[0.901]	[0.315]	[0.870]
Observations	136,133	136,133	106,231	106,231
Cohorts	2005-2012	2005-2012	2007-2012	2007-2012

Notes: Effects by race are estimated with separate estimations of Equation 1 by underrepresented minority status. Underrepresented minority includes students who identify as Black, Hispanic, American Indian, Native Hawaiian, or Pacific Islander. AP availability is measured at the school-by-cohort level, and counts the number of AP courses available to a high school cohort in their junior and senior year. Regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. Group-level means of the course- and exam-taking variables are in brackets. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 5: Effect of AP Course Availability on AP Course- and Exam-Taking, by Prior Achievement

	(1) Any AP courses taken	(2) # AP courses taken	(3) Any AP exams taken	(4) # AP exams taken
A. Linear Test Score Interaction				
# AP courses available at school	0.010***	0.023**	0.002	-0.006
in junior & senior year	(0.003)	(0.011)	(0.002)	(0.009)
Middle school math test score	0.166***	0.222***	0.124***	0.051
	(0.017)	(0.055)	(0.016)	(0.069)
# of AP courses available * math score	0.005***	0.046***	0.008***	0.070***
	(0.001)	(0.005)	(0.001)	(0.007)
Observations	154,549	154,549	123,003	123,003
Cohorts	2005-2012	2005-2012	2007-2012	2007-2012
B. Top 25 vs. Bottom 75 Percent of Test Scores				
Effect of # available AP courses for:				
Bottom 75% of test score distribution	0.012***	0.023**	0.003*	0.002
	(0.003)	(0.009)	(0.002)	(0.005)
Group mean	[0.250]	[0.442]	[0.170]	[0.330]
Observations	116,319	116,319	92,564	92,564
Top 25%	0.012***	0.074***	0.008*	0.048*
	(0.004)	(0.022)	(0.005)	(0.026)
Group mean	[0.706]	[1.941]	[0.625]	[2.062]
Observations	38,230	38,230	30,439	30,439
Cohorts	2005-2012	2005-2012	2007-2012	2007-2012

Notes: Middle school math test score is measured as a standardized scale score, standardized on the full population of test takers within year, grade, and subject. I use eighth grade test score if available and seventh grade score if not. Students missing a test score are not included in this analysis. In Panel A, effects by academic preparation are estimated using a single equation (Equation 2), where course availability is interacted with the continuous measure of test score. In Panel B, effects by academic preparation are estimated with separate estimations of Equation 1 by test score group. All regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. Group-level means of the course- and exam-taking variables are in brackets. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: Effect of AP Course Availability on College Outcomes

	(1) Enrolled any college	(2) Enrolled 2-year college	(3) Enrolled 4-year college	(4) Enrolled compet.+ college	(5) Earned BA in 4 years	(6) Earned BA in 6 years
# AP courses available junior and senior year	0.002	-0.001	0.002	0.002	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Mean of outcome variable	[0.700]	[0.288]	[0.430]	[0.394]	[0.158]	[0.320]
Observations	173,151					
Cohorts	2005-2012					

Notes: Table reports estimates of β_1 from Equation 3, which regresses student-level college outcomes on AP course availability. AP availability is measured at the school-by-cohort level, and counts the number of AP courses available to a high school cohort in their junior and senior year. Regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. All college outcomes are unconditional on initial enrollment. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 7: Effect of AP Course Availability on College Outcomes, by Family Income

	(1) Enrolled any college	(2) Enrolled 2-year college	(3) Enrolled 4-year college	(4) Enrolled compet.+ college	(5) Earned BA in 4 years	(6) Earned BA in 6 years
Effect of # available AP courses for:						
Students from low-income families	0.005	0.002	0.002	0.002	0.000	-0.000
	(0.004)	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)
	[0.558]	[0.300]	[0.271]	[0.229]	[0.045]	[0.132]
Observations	41,974					
Students from non-low-income families	0.000	-0.002	0.001	0.001	0.001	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
	[0.745]	[0.284]	[0.480]	[0.446]	[0.194]	[0.380]
Observations	131,177					

Notes: Low-income status is proxied by eligibility for free or reduced-price lunch (FRPL). Effects by income are estimated with separate estimations of Equation 3 by FRPL status. AP availability is measured at the school-by-cohort level, and counts the number of AP courses available to a high school cohort in their junior and senior year. Regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. Group-level means of the outcome variables are in brackets. All college outcomes are unconditional on initial enrollment. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 8: Effect of AP Course Availability on College Outcomes, by Race and Ethnicity

	(1) Enrolled any college	(2) Enrolled 2-year college	(3) Enrolled 4-year college	(4) Enrolled compet.+ college	(5) Earned BA in 4 years	(6) Earned BA in 6 years
Effect of # available AP courses for:						
Black, Hispanic, & Native students	0.006	0.003	0.004	0.005	0.002	0.004
	(0.004)	(0.003)	(0.004)	(0.004)	(0.002)	(0.003)
	[0.589]	[0.274]	[0.330]	[0.284]	[0.054]	[0.147]
Observations			37,018			
White & Asian students	0.000	-0.002	0.001	0.001	0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
	[0.730]	[0.292]	[0.457]	[0.423]	[0.186]	[0.367]
Observations			136,133			

Notes: Effects by race are estimated with separate estimations of Equation 3 by underrepresented minority status. Underrepresented minority includes students who identify as Black, Hispanic, American Indian, Native Hawaiian, or Pacific Islander. AP availability is measured at the school-by-cohort level, and counts the number of AP courses available to a high school cohort in their junior and senior year. Regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. Group-level means of the outcome variables are in brackets. All college outcomes are unconditional on initial enrollment. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 9: Effect of AP Course Availability on College Outcomes, by Prior Achievement

	(1) Enrolled any college	(2) Enrolled 2-year college	(3) Enrolled 4-year college	(4) Enrolled compet.+ college	(5) Earned BA in 4 years	(6) Earned BA in 6 years
Effect of # available AP courses for:						
Bottom 75% of test score distribution	0.002	-0.001	0.002	0.002	-0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
	[0.678]	[0.339]	[0.358]	[0.318]	[0.096]	[0.245]
Observations			116,319			
Top 25% of test score distribution	0.003	-0.002	0.005**	0.005*	0.007***	0.004
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
	[0.857]	[0.159]	[0.717]	[0.686]	[0.367]	[0.601]
Observations			38,230			

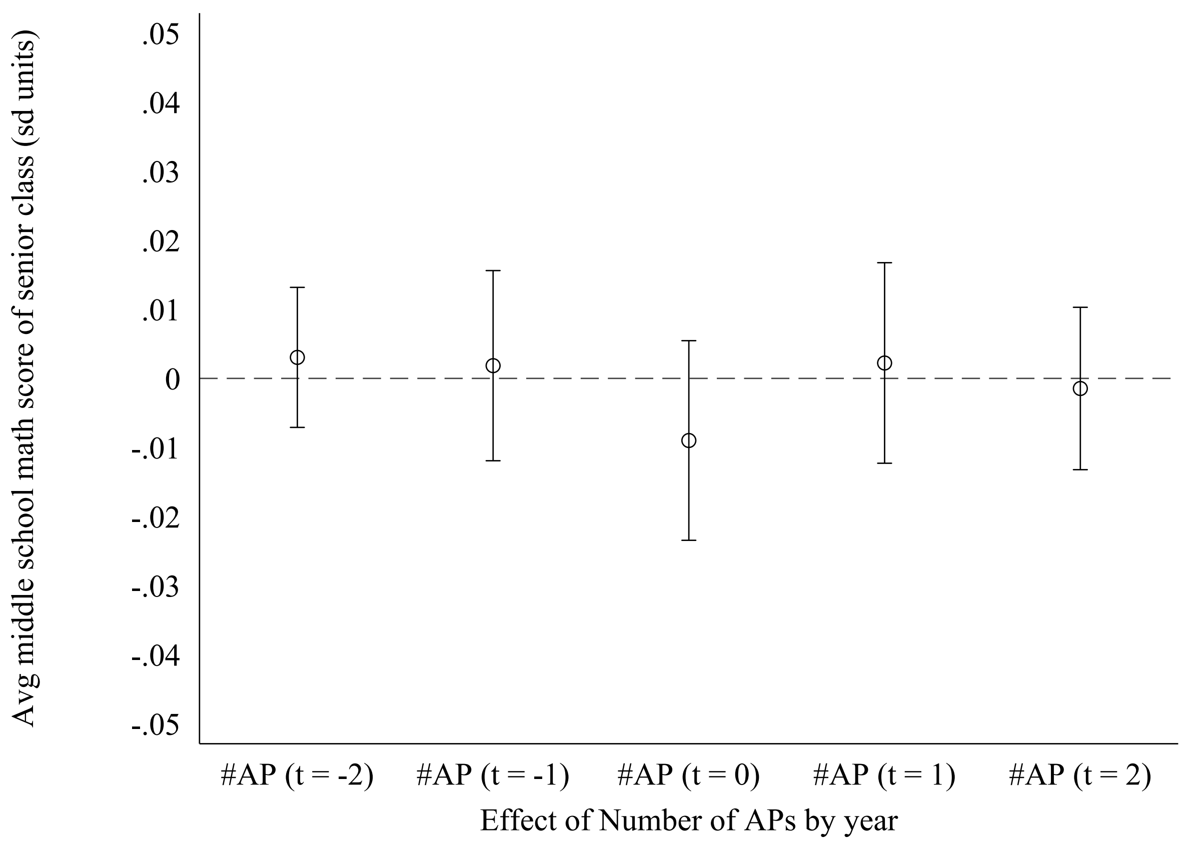
Notes: Middle school math test score is measured as a standardized scale score, standardized on the full population of test takers within year, grade, and subject. I use eighth grade test score if available and seventh grade score if not. Students missing a test score are not included in this analysis. Effects by academic preparation are estimated with separate estimations of Equation 3 by test score group. Regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. Group-level means of the outcome variables are in brackets. All college outcomes are unconditional on initial enrollment. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 10: First Stage, Reduced Form, and 2SLS Effects of AP Course Availability on Student Outcomes, by Alternative Cuts of Prior Achievement

		Part of middle school test score distribution			
	All	Top 50%	Top 25%	Top 10%	Top 5%
A. First Stage: Effect of # AP courses available on:					
# AP courses taken	0.032***	0.048***	0.074***	0.117***	0.108**
	(0.009)	(0.017)	(0.022)	(0.032)	(0.046)
Kleibergen-Paap Wald F statistic	11.65	7.89	11.41	13.14	5.53
B. Reduced Form: Effect of # AP courses available on:					
Enrolled in any college	0.002	0.002	0.003	0.001	0.005
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Enrolled in 2-year college	-0.001	-0.000	-0.002	0.001	0.001
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Enrolled in 4-year college	0.002	0.001	0.005**	0.001	0.005
	(0.002)	(0.002)	(0.002)	(0.003)	(0.005)
Enrolled in competitive+ college	0.002	0.001	0.005*	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.003)	(0.005)
Earned BA in 4 years	0.001	0.001	0.007***	0.007	0.007
	(0.001)	(0.002)	(0.003)	(0.005)	(0.006)
Earned BA in 6 years	0.001	-0.000	0.004	0.006*	0.008
	(0.002)	(0.002)	(0.003)	(0.004)	(0.005)
C. IV Analysis: LATE of taking one additional AP course on:					
Enrolled in any college	0.049	0.038	0.046	0.010	0.044
	(0.067)	(0.046)	(0.034)	(0.026)	(0.038)
Enrolled in 2-year college	-0.034	-0.002	-0.023	0.006	0.010
	(0.058)	(0.048)	(0.036)	(0.023)	(0.029)
Enrolled in 4-year college	0.057	0.014	0.069**	0.010	0.042
	(0.054)	(0.040)	(0.031)	(0.026)	(0.038)
Enrolled in competitive+ college	0.060	0.013	0.064**	0.016	0.023
	(0.056)	(0.044)	(0.032)	(0.025)	(0.041)
Earned BA in 4 years	0.035	0.018	0.101**	0.062*	0.066
	(0.030)	(0.040)	(0.040)	(0.037)	(0.052)
Earned BA in 6 years	0.017	-0.007	0.059	0.056*	0.076
	(0.045)	(0.046)	(0.036)	(0.031)	(0.049)
N	173,151	76,888	38,230	15,120	7,656

Notes: Middle school math test score is measured as a standardized scale score. I use eighth grade test score if available and seventh grade score if not. Students missing a test score are not included in any of the top X% columns. First stage effects come from estimating Equation 1 on students in the top X% of the test score distribution, where X is indicated in the column headers. Reduced form effects come from estimating Equation 3. LATE effects are estimated using 2SLS. All regressions include school fixed effects, year fixed effects, and school-specific linear time trends. Robust standard errors clustered at the school level are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Figure 1: Test for Selection: Effect of Number of AP Courses Available on Average Middle School Math Test Scores of Senior Class



Notes: Figure shows estimated coefficients and 95 percent confidence intervals for the β_k's in Equation 5, which is a school-year-level regression of average standardized middle school math test score of the high school's senior class on the number of AP courses offered every year. Regressions control for school and year fixed effects and school-specific linear time trends.