THE IMPACT OF HIGH SCHOOL ADVANCED PLACEMENT COURSE PARTICIPATION ON COLLEGE ENROLLMENT AMONG WOULD-BE FIRST-GENERATION COLLEGE STUDENTS

A Thesis
submitted to the
Faculty of the Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of
the requirements for the
degree of Master of Public Policy
in Public Policy

By Chelsea R. Goodly, M.Ed.

Washington, DC
April 12, 2018
THE IMPACT OF HIGH SCHOOL ADVANCED PLACEMENT COURSE PARTICIPATION ON COLLEGE ENROLLMENT AMONG WOULD-BE FIRST-GENERATION COLLEGE STUDENTS

Chelsea R. Goodly, M.Ed.

Thesis Advisor: Donna Morrison, Ph.D.

ABSTRACT

As the number of would-be first-generation college students increases in high schools across the United States, policymakers must look for more ways to increase rates of postsecondary enrollment for this often vulnerable group of students. Among the many strategies that have been explored to increase postsecondary access is the expansion of Advanced Placement programs in high schools. These programs expose high school students to college level course work and provide students the opportunity to potentially earn college credit. Because of their rigorous academic nature and favorably among highly-selective colleges, participation in these courses are thought to enhanced students’ affinity and preparation for college. While research that affirms this belief has been mixed, the majority of these studies have focused on student samples in which would-be first-generation college students have been largely absent or underrepresented. This study seeks to fill this gap in the literature by exploring how and if participation in Advanced Placement courses increases the likelihood of college enrollment among would-be first-generation college students. The analysis suggests that AP course participation exerts a significant degree of influence over students’ college going decisions independent of other key factors, consistent with the primary hypothesis of the study.
I would like to express my gratitude to my thesis advisor, Dr. Donna Morrison, for her invaluable support, guidance, and encouragement.

The research and writing of this thesis is dedicated to students everywhere with the courage to dream big.

Thank you,

Chelsea Goodly
# Table of Contents

- **Introduction** ................................................................................................. 1
- **Background** .................................................................................................. 3
- **Review of Literature** .................................................................................. 4
  - Advanced Placement Courses ..................................................................... 4
  - College Enrollment ..................................................................................... 5
- **Data and Methods** ....................................................................................... 9
  - Analysis Sample .......................................................................................... 10
  - Data Limitations ......................................................................................... 11
  - Variables ..................................................................................................... 12
  - Empirical Model and Estimation Strategy .................................................. 14
  - Methodological Limitations ....................................................................... 15
- **Results** ......................................................................................................... 18
  - Are students who take AP courses different than those who do not take AP courses? ..... 17
  - Are students who enroll in college different than those who do not enroll in college? ..... 19
  - What percentage of students in AP course go on to enroll in college compared to those who do not take AP courses? .................................................................. 20
  - Regression Analysis ................................................................................... 20
    - How does academic motivation influence college enrollment? .............. 21
    - How do academic expectations influence college enrollment? ............... 22
    - Does the influence of educational expectations persist when controlling for academic motivation, and vice versa? ............................................................... 23
List of Figures

Figure 1: Conceptual Framework .........................................................30

List of Tables

Table 1: Means and Standard Deviations for Variables Included in Probit Model ............... 31
Table 2: Percent Distribution of Independent Variable Characteristics for Would-Be First Generation College Students, by AP Course Participation ........................................... 32
Table 3: Percent Distribution of Independent Variable Characteristics for Would-Be First Generation College Students, by College Enrollment ............................................ 33
Table 4: Percent Distribution College Enrollment for Would-Be First-generation College Student, by AP Course Participation .............................................................. 34
Table 5: Regression Output from Probit Model ............................................ 35
INTRODUCTION

Equity and access are top issues within education. Across the field, researchers, policymakers, and education leaders are devoted to finding ways to enhance equity and increase access for vulnerable student populations. In particular, as the importance of higher education attainment on lifetime earnings and outcomes has increased, the focus placed on expanding traditionally underserved students’ access to postsecondary education has intensified. Among the potential strategies that have been discussed and studied is the expansion of Advanced Placement (AP) programs in high schools. AP courses are designed to increase college readiness by providing high school students with the opportunity to engage with college-level content and curriculum across a broad range of subject areas. While research around the efficacy of these programs has been mixed, the fact that schools serving the highest number of minority and low-income students have the lowest rates of AP course availability is of much concern given studies that show AP course taking is a key factor in the admission criterion of highly selective colleges and universities (Burdman, 2000).

The majority of research on AP course taking has focused on its impact on high school students at large, using samples in which minority and other vulnerable student populations are underrepresented. With the exception of several notable studies on specific subgroups of students, this shortage has limited the field’s understanding of how AP course taking impacts the likelihood of college enrollment for these groups. One particularly important, but understudied population of students is would-be first-generation college students – those students whose parents both have had no postsecondary education experience (Ishitani, 2006; Chen, 2005). Lower parental education attainment is associated with a lower likelihood of college enrollment in the next generation. Available data reveal that while 46% of students who have a parent with
at least a college degree go on to receive a college degree or more, only 17% of students who
have parents with no college education go on to do the same (Lauff & Ingels, 2013). Among the
barriers to postsecondary education for these students are the prohibitive cost of tuition and a
lack of familiarity with navigating the college-application and admission process. Given these
unique obstacles, understanding the role that AP course taking can have on college enrollment
for would be first-generation college students can yield potentially significant insights to
improving outcomes for this group of students.

The present analysis seeks to contribute to the body of research on the impact of AP
course taking on student outcomes by examining the relationship between AP course
participation and college enrollment for would-be first-generation college students. This paper
uses data from the Educational Longitudinal Study of 2002 (ELS: 2002), which tracks a
nationally representative sample of high school sophomores through their transitions into the
workforce and postsecondary institutions. Given the need to estimate the occurrence likelihood
of the binary dependent variable, college enrollment, this paper utilizes probit modeling to assess
the impact of AP course taking while also accounting for observable school, family, and student
level characteristics.
BACKGROUND

Higher education attainment is associated with positive outcomes such as full-time employment and increased lifetime salary. (Kena et. al, 2015) For this reason, determining the school, family, and student-level factors that are associated with increased college enrollment is of great interest to researchers. However, college access and attainment are unequally distributed between students by factors like race and socioeconomic status. (Lauff & Ingles, 2013) Students from low-income backgrounds and students of color face increased obstacles to college enrollment including lower levels of academic preparation, institutional barriers, and financial limitations. These underlying divisions are exacerbated by the predictive power of parental education on children’s academic attainment. For this reason, increasing the number of first-generation college students in 2- and 4-year universities is vital to shifting the cycle of generational education attainment and improving the long term financial and economic outcomes for these students.

Increasing the number of first-generation college students is an important step towards improving the long-term economic and financial well-being for these students. By understanding how AP course participation influences the likelihood of would-be first-generation college students to attend college, resources and investments can be better allocated at the federal, state, and local level to drive improvements and boost outcomes.
REVIEW OF LITERATURE

While the literature examining the impact of AP coursework on students’ predisposition towards college enrollment is rather dense, studies of how this relationship translates for would-be first-generation college students is incomplete and inconclusive. For this reason, understanding the current research on the impact of AP coursework and the factors associated with college enrollment can help to provide a foundation to further explore and estimate this relationship on would-be first-generation college students.

Advanced Placement Courses

The prevalence of Advanced Placement (AP) courses throughout American high schools has grown significantly in recent years. Not only are AP courses used by many colleges as an indicator of academic readiness, but they are also viewed by many as a tool to prepare and expose high school students to college academics. (College Board, 2005; Geiser & Santelices, 2004) Determining the school and student-level factors that are associated with increased college enrollment is of great interest to researchers. Some have suggested that participation in Advanced Placement courses is one such factor. Gollub, et al. (2002) find that there is a direct correlation between access to rigorous academic programs, including high school AP courses and college acceptance. Additionally, this work is affirmed by research that lists AP course taking as a key factor in the admission criterion of highly selective colleges and universities (Burdman, 2000; College Board, 2005). Despite these findings, some research has called into question the connection between participation in AP courses and long-term college success. (Geiser & Santelices, 2004). This finding and others like it rest in part on the assumption that the quality of AP course instruction varies greatly between high schools where often the quality of
AP courses in low-income, high minority schools is not as high. (Hallett & Vennegas, 2011). Despite the tension between these findings, the relationship of AP coursework on at-risk student populations has still yet to be fully developed to make a clear determination as to the efficacy of AP participation in impacting future college enrollment for these groups in particular.

While there is an established literature on the impact of AP coursework on long-term outcomes for students at large, little work has been done to assess this impact on would-be first-generation college students, a particularly vulnerable group of students. This research can help add to existing literature by further exploring the impact of AP course work specifically on this group, and determine additional school and student-level characteristics that are most associated with higher education attainment.

College Enrollment

A long-standing body of literature suggests that college enrollment is shaped by a confluence of factors. These multidimensional factors fit within 3 primary domains: school-level, family-level, and student-level.

At the school level, factors like course rigor, school resources, and teacher quality have been shown to be positively correlated with increased student achievement, and as result, impact students’ college enrollment. The key conceptual frame underlying this relationship is the assumption that strong academic preparation increases a student’s desire to continue their formal education after high school by enrolling in college. For example, research shows that even after controlling for poverty and students’ language status, teacher preparation and certification are by far the strongest correlates of student achievement in reading and mathematics (Darling-Hammond, 2000). This suggests that exposure to high quality teachers would positively
influence would-be first-generation college students to enroll in college. Additional research that adds another dimension to this relationship has shown that at students’ perception of school climate and quality are far better predictors of student achievement than actual school climate as measured by school climate surveys (Maxwell, et.al., 2017). Affirming this point is the fact that after accounting for individual factors and socioeconomic status, variance in student achievement has consistently been explained by students’ school climate ratings (Hoy & Hannum, 1997; Collins & Parson, 2010).

Family characteristics like family income, parental education attainment, and parental expectations have all been shown to influence children’s long- and short-term educational outcomes.

Socioeconomic status (SES) is largely a key characteristic at the family level. A wealth of literature has found a positive correlation between a student’s SES and their educational attainment and family SES influences the location and quality of schools a student attends (Gooding, 2001). This relationship has huge ramifications for the opportunities and outcomes possible for students with low-SES and speaks to the influence that family characteristics can have on student outcomes.

Parental expectations are another family-level characteristic that has been found to bear much weight on student performance. Conclusive research shows that when parents hold higher expectations of students, students perform at higher levels and persist in school longer compared to students whose parents hold lower expectations. (Yamamoto & Holloway, 2010; Gooding, 2001). Important to understanding this relationship is research showing that parental expectations do not directly influence students’ desire to attend college, but influence students' perceptions of subjective norms around what they believe is expected of them (Carpenter and Fleishman, 1987).
Thus parent expectations influence the expectations students have for themselves and, to the extent that students’ self-expectations motivate their behavior, parental expectations are strongly related to postsecondary enrollment. These findings are relevant to this analysis. Parental expectations appear to be an important factor in influencing student’s perceptions of themselves and therefore weight heavily on students’ postsecondary aspirations and goals.

Also of particular relevance to this study is the nature of parental education attainment. Current literature finds that the impact of parent education surpasses the impact of both SES and other family characteristics (Gooding, 2001). A key reality is that parents’ education attainment has serious and long standing implications for their children. Higher education attainment is associated with positive outcomes such as full-time employment and increased lifetime salary that can enable parents to provide positive education experiences and opportunities for their child. (Kena et. al, 2015) Conversely, lack of college education increases the likelihood of low wage work and poverty. In turn this limits the types of resources and financial and social capital parents can provide their child, impacting their child’s long term education outcomes. This includes limits to extracurricular activities, tutoring, and SAT prep course that could boost college acceptance. It also creates non-economic barriers for children when parents lack familiarity with the college application process, essay writing, and navigating the financial aid process. Affirming this is a study conducted on university freshman found that students whose parents had higher levels of education have higher levels of achievement than students whose parents had lower educational levels (Gooding, 2001).

Both common sense and research show a positive correlation between student ability and student achievement. Numerous studies point not only to a clear link between ability and achievement, but also to ability and college enrollment (Manski and Wise, 1983). However,
much more nuanced is research that examines the relationship between student expectations and motivation and their long-term achievement. How a student perceives themselves as a learner and their orientation towards academics shapes their intrinsic motivation and in turn likelihood of college enrollment. Various studies have found a positive correlation between self-reported postsecondary aspirations in high school and eventual enrollment in postsecondary institutions. (Carpenter & Fleishman 1987; Hossler & Stage, 1992).
DATA AND METHODS

This study utilizes data from the Educational Longitudinal Study of 2002 (ELS: 2002) sponsored by the National Center for Education Statistics (NCES). The ELS: 2002 is a longitudinal study of high school sophomores meant to track students’ education and life outcomes as they progress beyond traditional school and into postsecondary institutions and the workforce. The study is a part of the NCES’s high school longitudinal studies program which was established to study the personal and professional development of students at varying stages through their lives. Taken together the data sets within the study describe the educational and life experiences of high school students across four decades – the 1970s, 1980s, 1990s, and 2000s.

ELS: 2002 is based on data taken in the base year 2002, first follow up in 2004, second follow up in 2006. The ELS: 2002 includes survey data of approximately 17,000 high school sophomores in 750 high schools administered during the spring term of each selected year. Survey data was also collected from every students’ parent, one math and English teacher per school, each school principal, and each school’s head librarian/media center specialist. This multilevel survey design provides researchers with a fuller picture of student experiences both in school, at home, and in their community. With this information, researchers are empowered with more information to answer questions about how students’ environment, family background, and other contextual factors impacted their transitions out of school and into postsecondary institutions and the workforce. This data better positions researchers and policymakers to better understand the full array of factors that shape students’ long-term personal and professional outcomes.

Schools and students are the study’s basic units of analysis and were determined in a two stage selection process. First schools were selected based on probability proportional to size.
1,221 eligible public, Catholic, and private schools from a pool of approximately 27,000 schools were contacted and 752 ultimately participated. In the next stage, approximately 26 sophomores were randomly selected from the enrollment lists provided by each school. To ensure comparability on subgroups, Asian and Hispanic students, as well as students in private school were oversampled. Each survey round the same students were surveyed and transfer students were separately followed as well. In addition to the survey administered during the base year, the first and second follow ups include additional information regarding postsecondary plans, drop outs, and other time variant characteristics. In the public use data file which is utilized here a cross-sectional first follow-up weight is used for sample members who completed a questionnaire in the first follow-up and a first follow-up panel weight for sample members who completed a questionnaire in both the base year and first follow-up, including those with base-year imputed data.

**Analysis Sample**

The analysis here is based on student, parent, and school-level ELS: 2002 data taken in the base year in 2002, first follow up in 2004, second follow up in 2006. While the ELS: 2002 data set includes all high school students, this study is restricted to the academic, postsecondary, and workforce transitions of would-be first-generation college students. Consistent with relevant research, would-be first-generation college students are defined as students whose parents both have had no postsecondary education experience. (Ishitani 2006; Chen 2005). To restrict the sample appropriately, parent-level data is used to establish the college generation status as well as other primary family characteristics. Student-level data is used to assess AP course taking, academic achievement, expectations, motivation, and post high school transitions. Because
transcript data is restricted in the public use data file, self-reported information about receipt of an academic honor is used as a proxy for academic achievement. School-level data is used to determine access to AP courses which serves as a proxy indicator of school quality.

Data Limitations

Despite the large size and thoughtful collection of the ELS:2002 data, the specific nature of the present study and the researcher’s lack of access to restricted-level data limited the scope of the present data analysis sample.

First, the specific nature of the present research required that many of the over 17,000 cases in the ELS:2002 data set be dropped. The present data analysis sample is restricted to would-be first-generation college students who attend a school where AP courses are offered. This was done to account for selection bias that would otherwise be present due to the inability to compare students who opt into AP courses to students who would opt in to AP course but do not have the option to take AP courses at their school. This significantly reduced the size of the data sample which creates challenges around variance and statistical significance in interpreting regression results.

Additionally, estimating the relationship between AP course participation and college enrollment can be challenging due to the cross directionality of the bias. While it is true that AP courses can be positively correlated with increased rates of college enrollment, it is unclear if AP course participation leads student to enroll in college or if the desire to enroll in college leads students to participate in AP courses. Because of this ambiguity, a certain degree of bias will be present in the regression model. Future efforts to examine this issue should take steps to account for endogeneity bias.
Next, the self-reported survey design of the ELS:2002 data creates certain challenges. Because students, parents, and school administrators were asked to complete these surveys themselves, this introduces the probably of human error, misreporting, and incomplete data. In every case there is no way to know if a survey respondent felt pressure to misrepresent information that was embarrassing or would rather not reveal. Further, the survey data contains many incomplete and ‘do not know’ responses. These gaps in information can harm the accuracy of the models generated in this study.

Lastly, due to lack of access to restricted data, the present study utilizes only ELS: 2002 public use data. For this reason, variables like student academic information, family SES, and AP course exam grades, were unable to be used. Instead the present study makes use of proxy variables to approximate these values in place of the actual restricted data. While using proxy data can help account for the absence more precise indicators, this data is in many cases only a weak indicator of the actual restricted variable. For example, family SES or income is a restricted variable in the ELS: 2002 dataset. Instead of using this information in the regression model, family SES is represented by a proxy series of control variables generated from a series of survey questions that asks students if certain items, appliances, and technology are present in their home. While is it true that these questions pick up on a degree of SES information, it is not a perfect substitute for real quantitative information about family income.

Variables

The present study seeks to explore the relationship between AP course participation and college enrollment among would-be first-generation college students. The key dependent variable is college enrollment as measured by ELS: 2002 second follow-up student survey data.
The ELS:2002 second follow-up survey contains a series of questions inquiring about students’ academic and social outcomes, one of which asks student respondents whether they have ever attended postsecondary school. For the purpose of this analysis, the key dependent variable is a dummy variable coded to 1 if the responded answered ‘yes’ to this question and 0 the respondent answered ‘no’. Observations with missing values were dropped from the sample.

The mean and standard deviation of independent variables used in this analysis are reported in Table 1 below. As is the case for the dependent variables, missing values were dropped from the sample. While imputation can be a viable option, depending on the method, there can be disadvantages, such as a failure to reflect random error or variance and to deflate standard errors. (Saunders et.al., 2006)

The independent variables are all situated within one of two levels: the student-level or the family-level.

Beginning at the student-level, the chief independent variable is AP course participation, a dummy variable coded based on student’s self-reported response to a second follow up survey question that asks whether they have ever participated in an AP course while enrolled in high school. Another key independent variable is student achievement, a dummy variable that utilizes self-reported survey data about whether a student has received an academic honor as a proxy for academic achievement. Also included is a series of categorical dummies that measure student motivation. These variables were generated using two survey questions that asks students to rate their feelings of interest in and fulfill with school on a Likert scale. This information was use to generate two separate dummy variables to indicate whether a student answered agree or disagree. Lastly, the final key explanatory variables at the student-level are a series of dummies that reflect whether a student desires to attend college as well as if they perceive their father and
peers expect them to attend college as well. These are all based on self-reported survey data that asks student to respond ‘yes’ or ‘no’ to these questions directly. At the student-level is also a dummy control variable to account for student gender.

At the family level, are a series of control variables that measure family socioeconomic status. These variables were selected from series of survey questions that ask students if certain items, appliances, and technology are present in their home. While there is no publically available data that directly measures family SES, the presence or absence of these conditions can serve as a proxy for family SES status. Based on an analysis of this data, computer ownership and in-home internet access showed the most variation across the sample and were selected to serve as proxy for family socioeconomic status.

While there are no independent variables at the school level, school level factors have been accounted for in the sample selection criteria. To avoid selection bias and ensure the students within the sample are comparable in nature, the sample contains only those students who attend schools that offer AP course. In this way AP course availability serves as a proxy for school quality and ensures that students who do not participate in AP course are not doing so for reasons of lack of access.

*Empirical Model and Estimation Strategy*

Given that the dependent variable is binary, OLS is not an appropriate estimation strategy. Its usefulness is limited by that fact that it assumes the relationship between the independent and dependent variable is linear. When applied to models with binary dependent variables that seek to estimate probability, OLS can return values outside of the 0-1 probability range and often fails to produce the most precise line of fit. This paper employs probit modeling
to examine whether participation in Advanced Placement courses at the high school level increases the probability of college enrollment among would-be first-generation college students. The relationship between AP course participation and college enrollment is represented by the probit model in which the (conditional) probability of a “successful” outcome is being modelled, shown by: 

\[ P[Y_i=1| X_{i1}, \ldots, X_{Ki}; \beta_0, \ldots, \beta_K] = \Phi(\beta_0 + \sum_{k=1}^{K} \beta_k X_{ki}) \]

In this model, \( Y \) represents the probability of college enrollment and \( \Phi \) is the cumulative distribution function of the standard normal distribution. The regressors include a dummy variable that represents whether a student participated in at least one AP course, a vector of observable student level characteristics like achievement and motivation, a vector of observable family level characteristics like SES and parental expectations, a vector of observable school level characteristics like teacher quality and school type, and the unobservable determinants of college enrollment outcomes. The primary treatment of interest is participation in an AP course and the primary counterfactual is that a would-be first-generation college student does not participate in an AP course at the high school level.

**Methodological Limitations**

The primary limitation of the present study is endogeneity related to the cross directionality of the bias between the key independent and dependent variable. While the experience of participating in an AP course can increase students’ future likelihood of enrolling in college, it could also be the case that students who plan to enroll in college are more likely to choose to participate in an AP course. The ambiguity in the relationship between AP course
participation and college enrollment creates inherent bias in the estimate produced from OLS or probit regression modeling.

One potential remedy for endogeneity is the use of propensity matching which is used in many cases throughout the literature. While useful in certain cases, propensity matching has been found to be unreliable and inaccurate when compared to actual experimental impacts and thus was ruled out for the purpose of this study (Agodini & Dynarski, 2001).

Another potential remedy to address this limitation is the inclusion of an instrumental variable. Including an instrumental variable in the model that is correlated with AP participation but not with college enrollment can account for and nullify the inherent endogeneity. While this would be a viable option, identifying a relevant and exogenous instrumental variable among the publically available survey data was not possible.

Omitted variable bias presents another limitation within this study. In order to obtain an unbiased estimate from any model, it must be the case that there are no variables omitted from the model that are correlated with both dependent and independent variables. Given that this study relies on observable, self-reported survey information, there are many observable and unobservable factors that are not accounted for in the model which meet this criteria thus creating the potential for OVB. For example, race, student ability, and family cohesion are all factors not accounted for in the present study which are reasonably correlated with both AP participation and likelihood of college enrollment.

Lastly, the potential for correlated error terms also presents certain limitation within this study. In order to obtain an unbiased estimate from a regression model, it must be the case that there are no variables omitted from the model that are correlated with \( X_{t-1} \) and correlated with \( Y \) and that \( u_t \) is random for all observations. However, given the nature of the ELS:2002 data set
this is not the case. As is common for education research, within the ELS:2002 data set of over 17,000 students, there are clusters of students belonging to the same high school. What this creates is a correlated error term among these clustered set of students due the unobservable characteristics that exist at the school level. Thus, the available data make up two separate levels, level 1 which refers to information about individual students and level 2 which refers to information about each school represented in the sample. Using a single-level estimation strategy like OLS or probit forces the statistical software to only select one level as the basis for analysis, increasing the risk of incorrect estimates (Quandt-Fath, 2014). While using an estimation strategy like hierarchical linear modeling could help to account for the multi-leveled nature of the data, the limitation of the ELS: 2002 public use data set prevented this possibility.
**RESULTS**

**Are Students Who Take AP Courses Different Than Those Who Do Not Take AP Courses?**

The first question of interest to the present study is whether would-be first-generation college students who take AP courses are systematically different than those who do not take AP courses in observable ways. A bivariate analysis of the data indicates that the answer is yes. Table 2 compares the percent distribution of students in AP courses against those not in AP courses classified by the independent variable characteristics.

At the student level, the population of students who participate in AP courses skews slightly female by about 5%-points. As expected, a significantly higher portion of students in AP courses report having received an academic honor while in school compared to those students not in AP courses. Additionally, there is a roughly 10%-point difference, favoring AP students (49%), in the share who report high levels of academic motivation compared to non-enrollees (38%). Given the increased academic rigor of AP courses and the academic esteem associated with them, the greater degree of motivation of students in AP courses is not surprising.

The highest levels of divergence between AP versus non-AP students is found in their expectation for themselves and, especially, in other’s expectations of them. While roughly 63% of students in AP courses expect to go to college, this is true of only 58% of non-AP students. An even bigger differential – 15%-points – is observed across the two groups in terms of their perceptions of the educational aspirations their fathers hold for them. Eighty-five percent of students doing AP coursework report that their fathers expect them to go to college compared to 70 percent of those not in AP course. Additionally, students in AP courses are more likely than non-AP students to report that their peers expect for them to go to college – 70% and 59%, respectively. These findings point to the clear relevance of parental and peer influence as
differentiating factors between those students in AP versus those not in AP. Lastly family SES as measured by computer ownership and internet access does not seem to be a strong differentiating factor between AP versus non-AP students. While the differential for computer ownership is small, access to internet is about 5%-points more prevalent among students enrolled in AP courses.

Are Students Who Enroll in College Different Than Those Who Do Not Enroll in College?

The next question of interest is whether would-be first-generation college students who enroll in college are systematically different than those who do not enroll in college. Table 3 shows the result of a bivariate analysis that compares the percent distribution of students enrolled in college against those not enrolled in college classified by the independent variable characteristics.

At the student level, the population of students who enroll in college skews slightly female by about 7%-points. The direction of this skew consistent with and slightly larger than the female skew accosted with the AP student population. Similar to the results in Table 2, a significantly higher portion of students enrolled in college report having received an academic honor compared to their counterparts not in college. Additionally, the share of students who report being highly motivated is markedly greater among college enrollees (44%) versus non-college students (31%); amounting to a more than 10%-point difference.

This, and several additional differentials between college-attendees versus non-attendees mirror the differences observed between AP-enrollees and non-enrollees. For example, when they were high school juniors, nearly 80% of students who ultimately enrolled in college, compared to just over half (52%) of their non-college-attending counterparts, reported that their
fathers expected them to get a university education. The comparable figures for peers’ expectations is 67 percent among college-attendees and 45 percent for students who were not enrolled as of 2006. These findings point to the clear relevance of parental and peer influence as differentiating factors between those students who enroll in college versus those who do not.

**What Percentage of Students in AP Course Go onto Enroll in College Compared to Those Who Do Not Take AP Courses?**

Given the many similarities of the distributions found in Table 2 and Table 3, a third question of relevance to this study is whether students who participate in AP courses go on to enroll in college at higher rates than those students who do not participate in AP courses.

Table 4 shows the percent distribution of college enrollment for would-be first-generation college students by AP course participation. Here we see a higher percent of students (76%) who take AP courses go on to enroll in college versus those who do not participate in AP courses (61%). This 15%-point difference is statistically significant and gives credence to the hypothesis that AP course participation is positively associated with college enrollment among would-be first-generation college students.

**Regression Analysis**

Having established a bivariate association between AP course-taking and college enrollment, the next step in this study is to explore whether this relationship persists net of controls for family socioeconomics, ability, and gender that are associated with enrollment in AP courses. Table 5 presents marginal effects from estimating a series of probit models predicting college enrollment. The baseline model includes the SES proxy variables that measure computer
and internet ownership, the ability proxy variable that indicates whether a student has been recognized for an academic honor, and a dummy variable for gender. The findings are consistent with existing literature.

Students who have been recognized for their academic performance are 13%-points more likely to enroll in college than students who have not received an academic award (p<.00). Males are 5%-points less likely than their female counterparts to transition into postsecondary education. The marginal effect of computer ownership is not statistically significant (p =.15), while internet access has both a positive and statistically significant marginal effect (p<.05) on the conditional probability of enrolling in college. To the extent that having in-home internet access is a proxy for higher levels of family SES, it is not surprising that it would be associated with an increased likelihood of college enrollment. Specifically, all else being equal, students who report having internet access in their home are 16%-points more likely to enroll in college compared to students who do not have in-home internet access.

How does academic motivation influence college enrollment?

While family socioeconomics and ability exert a strong influence on students’ long term academic outcomes, research also shows that when students are motivated and excited by academics they perform better. For this reason, the motivation proxy variables that measure students’ reported interest and fulfillment with school are added to the baseline model to determine how and if these factors might be correlated with ability and SES.

Interestingly interest in school does not have a statistically significant effect (p=.19), which is inconsistent with literature that identifies this form of motivation as an important factor in shaping student outcomes. While the impact of fulfillment by school is in fact positive and
significant, the marginal effect is relatively small as students who report being fulfilled by school are only 6%-points more likely to enroll in college that their unfilled peers all else being equal.

However, what is important to note in Model 2 is the decrease in the marginal effect of achievement. With the addition of the motivation proxy variable, the effect of having received an academic honor drops slightly, to 12%-points. While small, this change suggests that the variables are correlated and that at least part of the effect initially attributed to student achievement is transmitted through the degree of motivation that a student feels towards school. Knowing that one mechanism of achievement is in fact student motivation highlights that student achievement, to the extent that it is captured by this measure, is not a monolithic factor of ability; students of similar achievement levels may have different levels of interest in school in ways that matter for long term outcomes.

How do academic expectations influence college enrollment?

While motivation is a subject of key interest to researchers, so too is the influence that students’ expectations of themselves and the expectations they perceive others have of them play in shaping students’ long term outcomes. For this reason, Model 3 replaces the student motivation proxy variables with the variables that measure students’ expectations to determine if and how these variables are correlated with college enrollment as well as family SES and ability.

As shown in Table 5, all of the expectations variables are statistically significant and strongly positively correlated with likelihood of college enrollment. First, students who report that they expect to go to college are 36%-points more likely to enroll in college compared to their peers without college expectations, a marginal effect greater than any other variable in the model. This finding is consistent with academic research that suggests students who see
themselves as college material are more likely to attend. The next largest marginal effect is associated with paternal expectations. Students who perceive that their fathers expect them to attend college are 14%-points more likely to attend college than their peers, when controlling for all other factors in the model, including students’ expectations of themselves. This is particularly important given that the students who make up this sample all have parents with no college experience. Thus, despite their parents’ lack of college education, their parents still hold a strong degree of influence as student make their own college decisions. Lastly, students who perceive that their peers expect them to go to college are 8%-points more likely to enroll in college than students whose peers do not expect them to attend college. This is not surprising to the degree that the beliefs and expectations of students’ social circles influence their own decisions and behaviors.

Taken together these findings suggest that beyond ability and SES, students’ perceptions around college and their beliefs in their own ability exert an overwhelming influence on students’ academic outcomes. Of particular interest to the study is the fact that paternal expectations exert a stronger influence on students’ college enrollment patterns than their peers, which is inconsistent with accepted literature that suggest students’ peers have a dominating influence on students’ academic and social development (Harris, 1998).

Does the influence of educational expectations persist when controlling for academic motivation, and vice versa?

After determining the unique influence that expectations and motivations exert on their own, understanding how these important features interact together is the subject of Model 4 in Table 5.
Surprisingly when taken together, the proxy variables for student motivation decrease drastically and are no longer significant. Conversely the marginal effects on student expectations remain relatively unchanged and all are still statistically significant. This finding has large implications on how researchers understand the processes that shape students’ college going patterns for would-be first-generation college students. This model suggests that above and beyond students’ interest in school, student expectations have a far greater explanatory power in their likelihood of college enrollment when controlling for the most important observable factors like SES and achievement.

How does AP course enrollment influence college enrollment?

After accounting for the range of student and family-level characteristics, Model 5 in Table 5 explores the central questions of this study – does AP participation increase the probability that a student will attend college?

Model 5 shows that AP course participation is positively and statistically significantly correlated with the likelihood of college enrollment. Specifically, would-be first-generation college students who participate in an AP course are 8%-points more likely to enroll in college than their peers who did not participate in an AP course, controlling for gender, achievement, family socioeconomics, student expectations, and student motivation. This finding shows that AP course participation exerts a significant degree of influence over students’ college going decisions independent of other key factors, consistent with the primary hypothesis of the study.

Of key interest is the fact that the size of the marginal effect of AP participation is of similar magnitude to the influence exerted from peer expectations and being recognized for an academic honor, all around 8%-points and all statistically significant. Additionally, of note is that
even after accounting for the influence of AP course participation, students’ expectations of themselves and perceptions of the other’s expectations of them remain the most important factor in shaping students’ likelihood of future college enrollment.
DISCUSSION

As the number of would-be first-generation college students increases in high schools across the United States, policymakers must look for more ways to increase rates of postsecondary enrollment for this often vulnerable group of students. Among the many strategies that have been explored to increase postsecondary access is the expansion of Advanced Placement programs in high schools. These programs expose high school students to college level course work and provide students the opportunity to potentially earn college credit. Because of their rigorous academic nature and favorably among highly-selective colleges, participation in these courses are thought to enhanced students’ affinity and preparation for college. While research that affirms this belief has been mixed, the majority of these studies have focused on student samples in which would-be first-generation college students have been largely absent or underrepresented. This study seeks to fill this gap in the literature by exploring how and if participation in Advanced Placement courses increases the likelihood of college enrollment among would-be first-generation college students.

The analysis suggests that AP course participation increases the likelihood of college enrollment among would-be first-generation college students. Specifically, students who participate in these courses are 8%-points more likely to enroll in college than their peers who do not participate when holding key school, family, and student level characteristics constant. These findings speak to the strength and positive role AP course participation plays in shaping students’ academic pathways. The estimated marginal effect size also has important policy relevance. While additional research it is needed, these preliminary findings suggest that by expanding access to AP courses among would-be first-generation college students, policymakers
can help to strengthen high school to postsecondary transitions for this group of students. Additionally, the high marginal effect attributed to students’ expectations of themselves and their perception of other’s expectations of them is also has important policy relevance. These findings speak to the significance of the ways students’ identities as learners and what they believe about themselves is shaped by those around them, the school culture, and the rigor of course work they are exposed to. It is possible to imagine that exposure to more Advanced Placement courses is one strategy that can be used to influence students’, peers’, and familial expectations and thus impact college enrollment.

**Limitations**

Despite care to implement rigorous standards in the creation of the analysis file and regression design, the model estimates are likely biased due both to inherent endogeneity between the key independent and dependent variable as well as omitted variable bias. Because students who plan to attend college may self-select into AP courses it was not possible within the constraints of this analysis to fully account for this effect. For this reason, the estimate produced in the model could account for both the effect of AP course participation in shaping college enrollment decisions, and conversely for college enrollment decisions shaping AP course participation. Additionally, because of this study’s reliance on proxy variables as measures of family SES and student academic ability, the imprecision of these estimates is another likely source of bias. Lastly, because of the limits of the ELS:2002 public data use file, there are likely many more variables correlated with AP participation and college enrollment that are not accounted for in the model.
In addition to data limitations, the constraints placed on the dataset led to a significant reduction in sample size that weakened the predictive power of the model. In an effort to account for selection bias, the sample was restricted to only would-be first-generation college students who have reported education information on both parents and who attend high schools that offer AP courses. Because of these qualifying criteria, the ultimate regression model has an $n = 1414$ which is smaller than the ideal size to ensure statistical significance.

Future Research

While this study affirms the hypothesis that AP courses are positively associated with college enrollment, further research is needed to fully explain the nature of this relationship and the factors at play in shaping this positive correlation. Because the present study shows only a positive association due to participation in an AP course, future studies should look to better understand what aspects of AP courses are most associated with long term student achievement. For example, factors such as subject area, grade-level at which the course is offered, teacher quality, and class characteristics could each have varying roles of influence. Another question for future research is whether the number of AP courses a student takes is associated with a higher likelihood of college enrollment.

Additionally, because this study is limited to would-be first-generation college students who have data reported for both parents and who attend schools that offer AP courses, future studies should look to explore whether this relationship manifests for a wider array of would-be first-generation students. For example, propensity matching analysis could explore the relative benefits of AP course participation between students who do and do not attend schools where AP
course are offered. Additionally, future work should explore the academic pathways for would-be first-generation students who come from single-parent homes.

Lastly, because the motivating feature of this study and others like it is to increase postsecondary attainment for would-be first-generation college students, future studies should determine if the positive benefits of AP course participation follow students to and through college enrollment and translate into increased postsecondary persistence and academic achievement.
APPENDIX: FIGURES AND TABLES

Parental Education

Family Level Characteristics (SES, Expectations)

School Level Characteristics (AP course availability)

AP Course Participation

College Enrollment

Figure 1. Conceptual Framework
Table 1. Means and Standard Deviations for Variables Included in Probit Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Explanatory Variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In AP</td>
<td>0.07</td>
<td>0.26</td>
</tr>
<tr>
<td>Ability- Received Honor</td>
<td>0.25</td>
<td>0.44</td>
</tr>
<tr>
<td>School – Interesting</td>
<td>0.38</td>
<td>0.48</td>
</tr>
<tr>
<td>School Fulfilling</td>
<td>0.40</td>
<td>0.49</td>
</tr>
<tr>
<td>Self - Wants College</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td>Father- Wants College</td>
<td>0.71</td>
<td>0.45</td>
</tr>
<tr>
<td>Peers- Wants College</td>
<td>0.60</td>
<td>0.49</td>
</tr>
<tr>
<td><strong>Explanatory/Control Variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.47</td>
<td>0.50</td>
</tr>
<tr>
<td>Computer</td>
<td>0.82</td>
<td>0.38</td>
</tr>
<tr>
<td>Internet</td>
<td>0.73</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>Dependent Variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrolled in College</td>
<td>0.62</td>
<td>0.49</td>
</tr>
</tbody>
</table>
Table 2. Percent Distribution of Independent Variable Characteristics for Would-Be First-generation College Students, by AP Course Participation

<table>
<thead>
<tr>
<th>Variables</th>
<th>AP</th>
<th>No AP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student Level Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.41</td>
<td>0.47</td>
</tr>
<tr>
<td><strong>Ability</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received Honor</td>
<td>0.52</td>
<td>0.23</td>
</tr>
<tr>
<td><strong>Motivation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School – Interesting</td>
<td>0.49</td>
<td>0.38</td>
</tr>
<tr>
<td>School – Fulfilling</td>
<td>0.26</td>
<td>0.42</td>
</tr>
<tr>
<td><strong>Expectations for College</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self - Wants College</td>
<td>0.63</td>
<td>0.58</td>
</tr>
<tr>
<td>Father- Wants College</td>
<td>0.85</td>
<td>0.70</td>
</tr>
<tr>
<td>Peers- Wants College</td>
<td>0.70</td>
<td>0.59</td>
</tr>
<tr>
<td><strong>Family Level Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer</td>
<td>0.88</td>
<td>0.82</td>
</tr>
<tr>
<td>Internet</td>
<td>0.77</td>
<td>0.73</td>
</tr>
</tbody>
</table>
Table 3. Percent Distribution of Independent Variable Characteristics for Would-Be First-generation College Students, by College Enrollment

<table>
<thead>
<tr>
<th>Variables</th>
<th>Enrolled in College</th>
<th>Not Enrolled</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student Level Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.43</td>
<td>0.50</td>
</tr>
<tr>
<td>Ability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received Honor</td>
<td>0.33</td>
<td>0.16</td>
</tr>
<tr>
<td>Motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School – Interesting</td>
<td>0.44</td>
<td>0.31</td>
</tr>
<tr>
<td>School – Fulfilling</td>
<td>0.41</td>
<td>0.43</td>
</tr>
<tr>
<td>Expectations for College</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self- Wants College</td>
<td>0.63</td>
<td>0.58</td>
</tr>
<tr>
<td>Father- Wants College</td>
<td>0.79</td>
<td>0.52</td>
</tr>
<tr>
<td>Peers- Wants College</td>
<td>0.67</td>
<td>0.45</td>
</tr>
<tr>
<td><strong>Family Level Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer</td>
<td>0.89</td>
<td>0.74</td>
</tr>
<tr>
<td>Internet</td>
<td>0.81</td>
<td>0.62</td>
</tr>
</tbody>
</table>
Table 4. Percent Distribution College Enrollment for Would-Be First-Generation College Student, by AP Course Participation

<table>
<thead>
<tr>
<th>Variables</th>
<th>Enrolled in College</th>
</tr>
</thead>
<tbody>
<tr>
<td>In AP</td>
<td>0.76</td>
</tr>
<tr>
<td>Not in AP</td>
<td>0.61</td>
</tr>
</tbody>
</table>
Table 5. Regression Output from Probit Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1 dF/dx</th>
<th>Model 2 dF/dx</th>
<th>Model 3 dF/dx</th>
<th>Model 4 dF/dx</th>
<th>Model 5 dF/dx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, Male</td>
<td>-0.05*</td>
<td>-0.05*</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>Received Honor</td>
<td>0.13***</td>
<td>0.13***</td>
<td>0.09***</td>
<td>0.09***</td>
<td>0.09***</td>
</tr>
<tr>
<td>Computer Ownership</td>
<td>0.07</td>
<td>0.07</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Internet Access</td>
<td>0.16***</td>
<td>0.17***</td>
<td>0.13***</td>
<td>0.14***</td>
<td>0.14***</td>
</tr>
<tr>
<td>Self, College</td>
<td></td>
<td></td>
<td>0.35***</td>
<td>0.35***</td>
<td>0.35***</td>
</tr>
<tr>
<td>Father, college</td>
<td></td>
<td></td>
<td>0.14***</td>
<td>0.14***</td>
<td>0.14***</td>
</tr>
<tr>
<td>Peers, college</td>
<td></td>
<td></td>
<td>0.08**</td>
<td>0.08**</td>
<td>0.08**</td>
</tr>
<tr>
<td>Interested</td>
<td>0.04</td>
<td>.00</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fulfilling</td>
<td>0.06*</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP Participation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.08†</td>
</tr>
</tbody>
</table>

† P ≤ 0.1
* P ≤ 0.05
** P ≤ 0.01
*** P ≤ 0.001
BIBLIOGRAPHY


