USC Dornsife
Center for Economic and Social Research

USC Rossier
School of Education

# Student Pathways from Middle School through High School and into Postsecondary Education: MYP Student Outcomes in a Large United States Public School District

Revised: May 1, 2020

Anna Saavedra, Principal Investigator

Morgan Polikoff, co-Principal Investigator

Shira Korn, Primary Analyst

TABLE OF CONTENTS

# Preface

The International Baccalaureate (IB) Middle Years Programme (MYP), intended as a five-year program for students aged 11 through 16, provides global context for learning through eight subject groups: language and literature, language acquisition, individuals and societies, mathematics, sciences, arts, physical and health education, and design. Developed and supported with funding from the IB Organization (IBO), the MYP learning framework was designed to foster creativity, critical and reflective thinking, communication skills, intercultural understanding, and global engagement. The IB Research Department commissioned researchers at the University of Southern California's Center for Economic and Social Research and Rossier School of Education to conduct a study of IB students' academic progressions from the MYP, to the Diploma Programme, and continuing into postsecondary education, funding this work from July 2018 through January 2020. The IB specified the study research questions and data request specifications, and provided the USC researchers with the data used in this study. This technical report describes the study background, research methods, results, and implications.

# Figures

# Tables

# Acknowledgements

The authors extend our deep appreciation to district staff for voluntarily answering questions about our data and the district context. We thank Olivia Halic and her International Baccalaureate Organization colleagues for their comprehensive support. Finally, with gratitude, we acknowledge financial support for this research from the International Baccalaureate Organization. Any flaws remaining are solely the authors' responsibility. Findings and conclusions expressed in this presentation are those of the authors and do not necessarily reflect the views of the International Baccalaureate Organization.

# Executive Summary

The purpose of this study was to examine International Baccalaureate (IB) students' academic progressions from the Middle Years Program (MYP) to the Diploma Programme (DP), and continuing into postsecondary education, in a large United States school district. We used multilevel modeling and propensity score matching methodologies to examine the pathways of students who attended MYP schools as sixth graders during the 2005-6 or 2006-7 school years, while accounting for their prior academic achievement and demographic characteristics.

Of the 3,147 students enrolled as sixth graders in 2006 or 2007 in district MYP schools, a small fraction, 5 percent, enrolled in the IB Diploma Programme. This group, which we refer to as "MYP+DP," was the primary focus of our analyses. Our comparison group was composed of MYP students who do not enroll in the IB Diploma Programme. We refer to this group as "MYP-non DP." A marginally larger proportion though still low, 15 percent, took one or more IB examinations. One of our seven research questions addressed this group. The MYP+DP students were a more economically advantaged, higher-performing group of students compared to MYP-non DP students. Controlling for all available student-level demographic variables, associated with MYP students' greater probability of enrolling in the IB Diploma were higher middle school achievement scores in mathematics, reading, and science, higher attendance rates, and more average service learning hours.

As expected, middle school achievement scores were strong predictors of high school success for MYP+DP and MYP-non DP students. MYP students with higher middle school achievement scores in mathematics and science tended to perform better on our indicators of high school success (e.g., GPA, ACT scores, SAT scores) regardless of whether they continued to the IB Diploma. Also predictive of high school GPA for both groups was attendance in middle school.

For MYP+ DP students, middle school achievement scores in mathematics, and reading were positively associated with achieving a greater number of IB examinations completed with a score of 4 or higher. Participation in Free and Reduced Price Meals (FARMS) during middle school was negatively associated with the number of IB exams taken and completed with a score of 4 or higher, while Special Education (SPED) classification during middle school was negatively associated with the number of IB exams taken.

Compared with MYP-non DP students, MYP+DP students had higher GPAs, higher total PSAT, SAT and ACT scores, and more service learning hours. These results align with the demographic and prior performance characteristics of students who enrolled in the IB Diploma —they were a more economically advantaged and higher performing group.

Finally, MYP+DP students had higher immediate college enrollment rates (within one year post high school graduation) relative to MYP-non DP students. We found no evidence of significant differences in college enrollment within two years post-graduation.

Because schools participating in the MYP likely differed from those that did not in unmeasured as well as measured ways, and students that enrolled in the IB Diploma also likely differed from those

who did not, results cannot be attributed to either the MYP or the IB Diploma . Further, it is possible that the demographic and prior performance composition of students enrolled in the MYP and IB Diploma in this district in 2020 differs from that of students who enrolled in the MYP in 2006 or 2007. Finally, the IB changed the MYP in several fundamental ways in 2014 that may also limit generalizability of presented results.

# Introduction

The purpose of this study was to examine International Baccalaureate (IB) students' academic progressions from the Middle Years Program (MYP) to the Diploma Programme (DP), and continuing into postsecondary education, in a large United States school district. We used multiple regression modeling and propensity score matching methodologies to examine the pathways of students who attended MYP schools as sixth graders during the 2005-6 or 2006-7 school year, while accounting for their prior academic achievement and demographic characteristics. Results contribute to nascent quantitative research describing relationships between students' MYP participation and their later academic outcomes.

In this chapter, we provide the background motivating the study and present our research questions. In subsequent chapters we describe the study context, research approach, sample, data, and analytic methods. Finally, we present results, and conclude with their discussion.

## International Baccalaureate Program Description

Originally founded in 1968 in Geneva, Switzerland as a private means for diplomats' children to earn an internationally-recognized high school diploma, today's IB programs serve students aged 3 to 19 from a variety of backgrounds across more than 5,000 schools worldwide. In schools that have earned the IB authorization to offer one or more of its four programs, teachers use IB curriculum and pedagogy to teach a range of courses and other activities intended to prepare IB-enrolled students for global citizenship, further academic study, and careers. As of 2019, the Primary Years Programme (PYP) serves students aged 3 to 12 in 1,782 schools across 109 countries. The MYP serves students aged 11 to 16 in 1,358 schools across 108 countries. The IB Diploma Programme (DP), which is the original IB program, serves students aged 16-19 in 2,470 schools across 144 countries. And 214 schools across 23 countries offer the Career-Related Programme (CP), the newest of the four. Currently, 1,839 U.S. schools—89 percent of which are public (n=1,643)—offer at least one program, with 664 offering the MYP (IBO, 2019).

All four programs share a common commitment to developing the same skills and traits in students. The IB Learner Profile describes the objective of developing students to be: inquirers, knowledgeable, thinkers, communicators, principled, open-minded, caring, risk-takers, balanced, and reflective (IBO, 2020).

### Diploma Programme

The Diploma Programme (DP) was the original IB program, founded in 1968 as a means for high school students to earn an internationally-recognized diploma. The DP curriculum includes six subject areas and three interdisciplinary core requirements. The subject areas include language and literature, language acquisition, individuals and societies, sciences, mathematics, and the arts. The core requirements include the Theory of Knowledge epistemology course, a 4,000-word "extended essay," and the Creativity, Activity, and Service program. To earn the IB Diploma, in addition to fulfilling course and core requirements, students also must score above a defined threshold on IB-created and

-administered examinations on all six subjects. Students can also take one or more IB courses without qualifying to earn the IB Diploma. Among U.S. high school graduates in 2013 who participated in the DP, approximately 50 percent took courses but were not qualified as diploma candidates (Pilchin, Caspary & Woodworth, 2019).

### Middle Years Programme

The developers of the MYP designed it to share the DP philosophy and prepare students for the DP (IBO, 2008). Piloting began in 1987, and eight years later the IB officially adopted the MYP (IBO, 2019). As of 2019, the MYP curriculum includes eight subject areas: the six in the DP plus physical and health education, and design. In addition to their subject area courses, MYP students also take part in an interdisciplinary unit involving at least two subject groups, and complete an age-appropriate, long-term service-learning or personal project with practical implications.

In 2014, the IB revised the MYP, such that the version currently described in its materials differs in a few ways compared to the prior version of the program experienced by the students in our study. According to Head of MYP Development Robert Harrison and two IB colleagues, the alterations between the older and newer versions are primarily curricular and structural, including: new personal project requirements; improved documentation of curriculum; greater implementation flexibility and quality assurance; and greater attention to rigorous research of program impact (Harrison, Albright, & Manlove, 2015).

## Background

While a growing body of research addresses the MYP curriculum and its implementation, there is limited rigorous quantitative research describing the impact of MYP on student outcomes. Few studies to date have addressed the impact of MYP on student outcomes, and those that do are challenged by selection bias. Schools choosing to go through the IBO's rigorous accreditation process and earn the right to offer the MYP likely differ from schools that do not in ways, both measured and unmeasured, affecting student outcomes.

To date, three studies have addressed MYP impact. Tan & Bibby (2010) found that MYP students perform better than non-MYP students on standardized measures of mathematics, expository writing, and reading achievement, administered to students in international schools. Gordon and Bergeron (2015) found that students who perform better during the MYP also perform better on IB exams. In research most comparable to our present study, Wade & Wolanin (2015) used data from the same district to show that MYP students participated in AP and IB examinations at higher rates and performed better on those examinations than students from non-MYP schools. However, MYP participation was not related to SAT or ACT test-taking or performance.

Studies of the impact of the IB Diploma on student outcomes are similarly few in number. While all show positive relationships between IB Diploma enrollment and academic measures in secondary and post-secondary schools (e.g., Bergeron 2015; Caspary, 2011; Coca et al., 2012; Gordon, VanderKamp, & Halic, 2015; Halic, 2013; Higher Education Statistics Agency, 2016; Shah et al., 2010), none excepting Saavedra (2014) address selection bias. Saavedra, using rigorous sensitivity analyses to bound

propensity score estimates, makes a causal argument about the impact of IB Diploma enrollment on students' probability of on-time graduation from high school.

The primary contribution of the present study is to build upon Wade and Wolin (2015), following MYP students' outcomes beyond high school to include post-secondary enrollment and persistence.

## Research questions

We grouped our research questions[1] into three categories:

- Student demographic and academic characteristics predicting IB Diploma enrollment
- Relationships between students' IB participation, demographic and academic characteristics, and high school trajectories
- Relationships between students' IB enrollment and their postsecondary school trajectories

### Student demographic and academic characteristics that predict IB Diploma enrollment

Our first two research questions address relationships between students' demographic and academic characteristic and whether they subsequently enroll in the IB Diploma .

*1)* What are the demographic and academic characteristics of the MYP students who continue to the IB Diploma, compared with those of MYP students who do not continue to the IB Diploma ?

*2)* Which academic and non-academic characteristics of the MYP students predict enrollment in the IB Diploma , when demographic and school-level variables are taken into account?

### Relationship between students' IB participation, demographic and academic characteristics, and high school trajectories

Next, we address relationships between MYP participation and academic outcomes in high school. For most questions, we compared MYP students who participated in the IB Diploma relative to those who did not.

*3)* Which MYP student variables predict success in high school, in general, (for students who do not continue to IB Diploma ), and in the IB Diploma (for students who continue)) when demographic and school-level variables are taken into account?

*4)* Which characteristics of the MYP students predict the numbers of IB exams taken?

*5)* How does the academic performance of the MYP students who continue in the IB Diploma compare with the academic performance of MYP students who do not continue in the IB Diploma?

[1] The IB defined the research questions in their request for proposals.

*6)* How do the academic behaviors (e.g., attendance, suspension, student service learning) of MYP students who continue to the IB Diploma compare with the academic behaviors of MYP students who do not continue to the IB Diploma?[2]

### Relationship between students' IB enrollment and their postsecondary school trajectories

Finally, our third set of research questions address the relationship between MYP participation and post-secondary enrollment, again relative to IB Diploma students.

7) How do the postsecondary trajectories of MYP students who continue in the IB Diploma compare with the postsecondary trajectories of the MYP students who do not continue?

8) Does enrollment in the MYP, coupled with enrollment in the IB Diploma, increase the likelihood that students enroll in college immediately after high school graduation and/or within two years of high school graduation?

[2] We could not address the IB's posed research question, "Does enrollment in the MYP coupled with enrollment in the DP increase the probability that students graduate from high school?" because we did not have access to an indicator of whether students graduated from high school.

# District Context

The district under study, located in the mid-Atlantic United States, enrolled more than 150,000 students in the 2018-19 school year and is classified as a large suburban locale (National Council for Education Statistics, 2019). On average, district students score approximately one grade level above the national average. However, there are large performance gaps: Black students in the district score a half-grade level below the national average for all students; economically-disadvantaged students score one grade level below; and Hispanic students score three-quarters of a grade level below (Reardon, 2019). According to the district webpage, the four-year graduation rate is 88 percent, though lower for Hispanic and Black students.

Eight district schools offer the DP, eight offer the whole-school MYP, one offers the whole-school PYP, and two offer the CP. Four high schools offering the IB Diploma also offer the MYP for students in the 9th and 10th grades (Metis, 2016). During 2015-16, only 2 percent of district students (2,837 of 156,819) were enrolled in an IB program. Of those enrolled, 31 percent were White, 24 percent Hispanic, 20 percent Black, and 20 percent Asian. As we show in Figure 1, Hispanic and Black students were underrepresented relative to their district proportions, Asian students were overrepresented, and White student enrollment was approximately proportional (U.S. Department of Education Civil Rights Data, 2015).

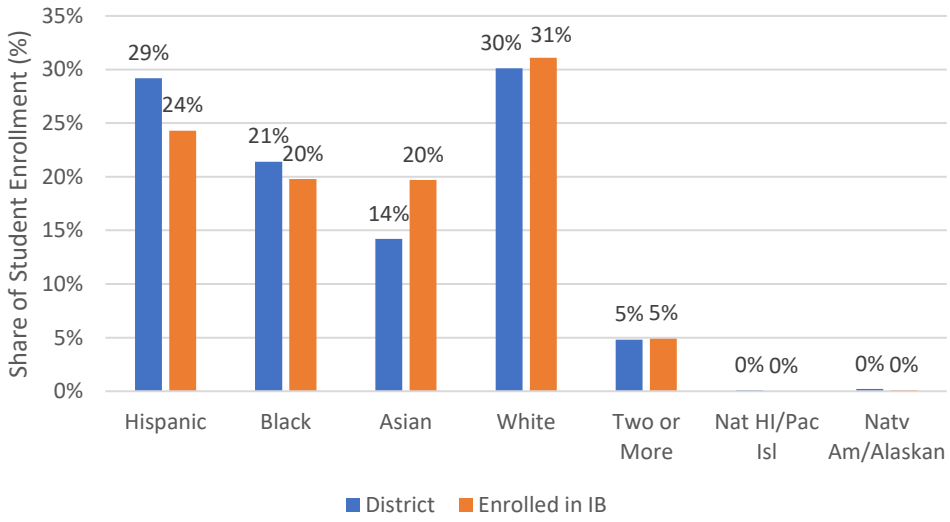


Figure 1: Proportion of students enrolled in IB and the district, by racial/ethnic category.

In 2015, the district commissioned external consultants to audit the district's magnet and choice programs, then provide recommendations. Of relevance to the present study, the audit found, as of the 2015-16 school year, special programs including IB were not equally available to all students and not well integrated when programs were offered within schools (e.g., the DP). One of the report's conclusions noted, "significant racial and socioeconomic disparities in the enrollment and acceptance

rates to academically selective programs, which suggest a need to revise the criteria and process used to selected students for these programs" (Metis, 2016, p. 9).

If the district has made substantial progress addressing these recommendations, some findings in our study, focusing on students who enrolled in the MYP as sixth graders in 2006 or 2007—particularly those related to student demographics—may no longer be valid.

For descriptive purposes and to give context on the schools in this analysis, in Table 1, we provide summary statistics on school-level variables for all middle schools in our sample from the 2006-07 and 2007-08 school years (i.e. the two years in which two cohorts of sample students were in sixth grade), using publicly available data from the National Center of Education Statistics' Common Core of Data. These data include variables related to school-level student demographics, class size, school type, and enrollment information.

Table 1. School-Level Descriptive Statistics by Year

| | 2006–07<br>Mean/SD | 2007–08<br>Mean/SD |
|---|---|---|
| Student-Teacher Ratio | 14.31 | 14.67 |
| | (1.17) | (1.45) |
| Magnet | 0.29 | 0.29 |
| | (0.49) | (0.49) |
| Total Enrollment | 701.43 | 697.29 |
| | (249.11) | (261.36) |
| Percent White | 0.43 | 0.43 |
| | (0.25) | (0.26) |
| Percent Asian | 0.11 | 0.11 |
| | (0.04) | (0.04) |
| Percent Hispanic | 0.22 | 0.24 |
| | (0.14) | (0.16) |
| Percent Black | 0.24 | 0.23 |
| | (0.13) | (0.13) |
| Percent Female | 0.49 | 0.50 |
| | (0.01) | (0.01) |
| Percent FRPL | 0.29 | 0.30 |
| | (0.17) | (0.19) |
| *N* | 7 | 7 |

Of the seven participating schools in the first year of our data (2006-07), the average student enrollment was approximately 701, with a demographic breakdown of: 43 percent White, 11 percent Asian, 22 percent Hispanic, 24 percent Black. The ratio of boys to girls was nearly even (respectively, 51-49). Additionally, schools enrolled an average of 29 percent economically-disadvantaged students, as measured by the share those receiving free and reduced priced meals. The average student-teacher ratio was approximately 14 students per teacher. Two schools (29 percent) in our sample were magnet schools. The 2006-07 values are quite consistent with 2007-08 values.

As detailed in the next chapter, for most of our analytic approaches presented in this report, we either include school fixed effects or use propensity score matching with an exact match on middle school enrollment. Both models allow control for any fixed differences between school contexts without including school-level controls. Therefore, we do not include these variables in the models presented throughout the remainder of this report.

# Research Design and Methodological Approach

The purpose of this study was to explore differences in the academic trajectories of MYP students who did and did not enroll in the IB Diploma, using multilevel modeling (i.e., fixed effects models) and propensity score matching methodologies to account for students' prior academic achievement and demographic characteristics.

## Sample

The dataset is structured around 3,147 students who were in sixth grade in the 2006-07 or 2007-08 school years (i.e., two cohorts), for whom the district provided student-level panel. Our data do not include student records on any students who did not attend an MYP school. It also does not include teacher-level variables; for example, describing experience or certification type.[3]

The students attended one of seven MYP schools within the district. After middle school, the students did one of the following:

1) Participated in the IB Diploma, as identified by taking the core "Theory of Knowledge" (TOK) course. We use the phrase "enrolled in the IB Diploma " to refer to students identified as enrolling in the IB Diploma by virtue of enrolling in TOK. Therefore we use TOK course-taking as proxy for pursuing the IB Diploma.
2) Did not enroll in the IB Diploma but took one or more IB exams.
3) Did not take any IB courses or examinations.

Of the students in our sample, 163 (5.2 percent) subsequently enrolled in the IB Diploma as full Diploma candidates as identified by having a non-missing TOK exam score, and 491 (15.6 percent) took one or more IB examinations, with the 163 IB Diploma students included in this count.[4] All analyses addressing the relationship between IB Diploma enrollment and outcomes use the sample of 163 students. We refer to this group as MYP+DP students and our reference group as MYP-non DP students. We use the sample of 491 students to address research question 4, "Which characteristics of the MYP students predict the numbers of IB exams taken?" In all analyses excepting those addressing research question 4, we included in the comparison sample the 328 students who took one or more IB examinations but did not enroll in the IB Diploma.

Ideally, we would define IB Diploma participation using course enrollment data rather than examination-taking data—as taking examination(s) is an outcome in itself and one that is conditional on enrolling in the course. However, in comparing the course enrollment data to the examination data, we found that the district's IB course enrollment data was too incomplete for us to consider it a valid and reliable indicator of program participation. The 2016 external review had also noted that as of the

[3] The IBO created the data request, worked with the district on fulfillment, and provided the data file to the USC research team.
[4] The data shows two students with TOK scores but taking no DP examinations.

2015-16 school year, the district "did not systematically track participation in or attrition from its choice and special academic programs" (Metis, 2016, p. 163).

In Table 2, we show most students (n=2,656, or 84 percent) did not take any IB examinations, including TOK.

Table 2. Counts of IB Exam and Theory of Knowledge Taking

| | | Frequency | Percent |
|---|---|---|---|
| Number of IB exams with scores | | | |
| | 0 | 2,656 | 84.4 |
| | 1 | 194 | 6.2 |
| | 2 | 88 | 2.8 |
| | 3 | 35 | 1.1 |
| | 4 | 14 | 0.4 |
| | 5 | 61 | 1.9 |
| | 6 | 89 | 2.8 |
| | 7 | 10 | 0.3 |
| Total | | 3,147 | 100 |
| Indicator for whether student has TOK score | | | |
| | 0 | 2,984 | 94.8 |
| | 1 | 163 | 5.2 |
| Total | | 3,147 | 100 |

In addition, we find students only began taking TOK examinations once they were also taking at least four IB examinations. On average, among students who took the TOK examination, students took 5.6 examinations—validating our definition of participating in the IB Diploma as identified by whether students took the TOK examination.

Eight district high schools offered the DP to the 3,147 MYP students. In Table 3, we show that of the seven MYP middle schools, six "fed" most of their students into specific DP-offering high schools, to the near-exclusion of the others. In other words, most students within a given MYP middle school attended a given IB Diploma -offering high school. We identify a student's middle and high school as the school in which they were enrolled in in sixth and ninth grade, respectively. A few students repeated ninth grade after transferring to another school. For these students, we identify their high school as the one they attended during the first year they were enrolled in ninth grade. We also collapse the students who attended non-DP high schools into one "missing" category.

Table 3. Transition Patterns from Middle School to High School

| | | Ninth-Grade School | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | School ID | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | Missing | Total |
| Sixth-Grade School | 10 | 1 | 2 | 264 | 2 | 1 | 2 | 11 | 2 | 289 | 574 |
| | 11 | 3 | 0 | 2 | 2 | 1 | 256 | 27 | 0 | 184 | 475 |
| | 12 | 4 | 0 | 12 | 0 | 1 | 8 | 27 | 2 | 475 | 529 |
| | 13 | 0 | 492 | 2 | 6 | 4 | 0 | 2 | 4 | 162 | 672 |
| | 14 | 401 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 131 | 532 |
| | 40 | 120 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 28 | 150 |
| | 41 | 157 | 1 | 0 | 1 | 1 | 2 | 0 | 0 | 53 | 215 |
| | Total | 686 | 495 | 281 | 12 | 8 | 268 | 67 | 8 | 1322 | 3147 |

## Measures

For our indicator of participation in the IB Diploma (i.e. MYP+ DP students, n=163), we created a dummy variable equal to one if the student has a non-missing value on the TOK score variable; if not, the variable equals zero (i.e. MYP-non DP students).

Student-level predictors include:

- Demographic characteristics: race/ethnicity[5], gender, status of free/reduced-price lunch (FRL/FARMS), ESOL status, special education status
- Academic/prior achievement characteristics: state-wide assessment scores[6]
- Non-academic behaviors: attendance, suspension, student service learning hours

For academic and non-academic variables that change from year to year (i.e., attendance, suspension, student service learning hours), we create a predictor equal to an average of each student's score on these variables over the course of the middle school years (i.e., sixth through eighth grade). For demographic variables that may or may not change over time (i.e., FRL/FARMS status, ESOL status, special education status) we generate variables for whether the student was never, ever, or always classified by the given designation over the course of their middle school and high school years (two variables total for each category: one for middle school and one for high school).

As mentioned above, we do not include school-level predictors, as both our fixed effects models and propensity score matching approach (using exact match on middle school) account for any between-school differences of concern.

For our student-level outcomes, we include:

[5] In 2010, the race codes changed to include a more racial identification categories (e.g., Pacific Islander and Multiracial). For this analysis, we rely on the race codes prior to 2010 because we have fewer missing values on this variable (788 students in our sample are missing new race codes).
[6] State assessment scores serve as outcomes in some models.

- General measures of academic success (GPA, PSAT/ACT/SAT scores); measures of academic success for students who enrolled in the IB Diploma (number of courses taken, exam scores)
- Measures of post-secondary success (whether the student was enrolled in college within one year of high school graduation and within two years past graduation)

For academic achievement variables changing from year to year (i.e., GPA), we created an outcome variable equal to an average of each student's score on these variables over the course of the high school years (i.e., 9th through 12th grade). We also calculated the GPA outcome separately for 9th-10th grade and 11th-12th grade to account for potential differences in grading between DP and non-DP classes, as the DP is only implemented in grades 11 and 12.

While we had originally discussed examining the relationship between IB Diploma enrollment and high school graduation, we could not conduct this analysis as the district did not provide us reliable indicators of high school completion.

In Table 4, we present summary statistics (means, standard deviations, minimum and maximum values) on our continuous student-level predictors. Because we have data for students across multiple years, we present summary statistics on all student-level variables for each student in their sixth-grade year (except on variables where we calculated averages across middle school grades, for which we present summary statistics on the averages).

On average, students in our sample earned an average scaled score of 421.5 in mathematics, 412.4 in reading, and 416.2 in science. These students attended an average of approximately 95 percent of school days, completed about 17 service learning hours annually during their middle school years, and were suspended at very low rates (the average student was suspended 0.07 times per year while in sixth through eighth grade).

Table 4. Summary Statistics on Student-Level Predictors

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| Middle School Math Score | 2,532 | 421.517 | 72.926 | 27.000 | 559.000 |
| Middle School Reading Score | 2,537 | 412.431 | 77.850 | 28.000 | 549.000 |
| Middle School Science Score | 2,521 | 416.179 | 47.359 | 240.000 | 650.000 |
| % Days Attended (Average across middle school grades) | 2,952 | 0.945 | 0.077 | 0.000 | 1.000 |
| Number of Days Suspended (Average across middle school grades) | 3,147 | 0.074 | 0.297 | 0.000 | 4.667 |
| Service Learning Hours (Average across Middle School Grades) | 2,834 | 17.019 | 14.374 | 0.000 | 221.333 |
| *N* | 3,147 | | | | |

For all categorical variables (student race, FRL/FARMS status, ESOL status, and SPED status), we present counts and frequencies in Table 5. Our sample is approximately 51 percent male, 11 percent Asian, 26 percent Black, 34 percent White, and 28 percent Hispanic. About half of our sample (51 percent) were never classified as economically disadvantaged, measured by whether they ever received free and/or reduced-price meals during or prior to middle school; but 23 percent were classified as economically disadvantaged for some of their middle school years or prior years, and 27 percent were economically disadvantaged for all their middle school years. Additionally, most of our sample were never classified as English learners or special education students (respectively, 68 and 81 percent) for any of their middle school years.

Table 5. Counts and Frequencies on Student-Level Categorical Variables

| | Count | Frequency |
|---|---|---|
| Student's gender | | |
| Male | 1,594 | 50.7 |
| Female | 1,553 | 49.3 |
| Total | 3,147 | 100 |
| Student Race | | |
| American Indian | 9 | 0.3 |
| Asian | 358 | 11.4 |
| Black | 825 | 26.2 |
| White | 1,082 | 34.4 |
| Hispanic | 873 | 27.7 |
| Total | 3,147 | 100 |
| FARMS | | |
| Never | 1,596 | 50.7 |
| Ever | 712 | 22.6 |
| Always | 839 | 26.7 |
| Total | 3,147 | 100 |
| English as a Second Language | | |
| Never | 2,146 | 68.2 |
| Ever | 890 | 28.3 |
| Always | 111 | 3.5 |
| Total | 3,147 | 100 |
| Special Education Status | | |
| Never | 2,537 | 80.6 |
| Ever | 288 | 9.2 |
| Always | 322 | 10.2 |
| Total | 3,147 | 100 |

Many students (n=644) are missing values on the predictors used, particularly on achievement variables. Because we employ listwise deletion for the remainder of this analysis, we exclude from our analytic models all students with missing values on covariates.

## Analytic Approach

Below, we explain our analytic approaches for addressing each of our eight research questions. We begin by summarizing our investigation of missing data and potential implications of missing data for interpretations of results.

### Missing Data Treatment

The first step of our data screening process was to assess the extent to which attrition (e.g., due to students enrolled in MYP schools in sixth grade who later moved out of the district, or for whom schooling pathway was unknown) may have contributed to biasing subsequent results. We define our attrition indicator as any student who withdrew from the district for a reason other than graduation.

Our approach was to use student-level covariates to predict student attrition from the district. We present two logistic regression models, appropriate for cases in which the outcome is a binary variable: one without middle school fixed effects (FE) and one including them. Functionally, including middle school fixed effects limits the analysis to comparing only students within the same middle school—thus, eliminating between-school differences that may also contribute to student attrition from the sample.

We present the results of our logistic regression models in Table 6. Coefficients are exponentiated, common practice for logistic regression models, such that coefficients greater than 1 should be interpreted as *increasing* the likelihood of attrition and coefficients less than 1 should be interpreted as *decreasing* the likelihood of attrition. The results show that on average, students who attrited from this study had lower middle school science scores, attendance rates, service learning hours, and a greater number of days suspended. They were also more likely to be Black (relative to White), less likely to be female, and less likely to have always received free and reduced-priced meals and/or SPED services.

Table 6. Logistic Regression Models Predicting Attrition

| | (1) Logistic Regression | (2) Logistic Regression with Middle School FE |
|---|---|---|
| Middle School Math Score | 1.000 | 1.001 |
| | (0.002) | (0.001) |
| Middle School Reading Score | 0.999 | 0.999 |
| | (0.002) | (0.001) |
| Middle School Science Score | 0.996* | 0.996** |
| | (0.002) | (0.002) |
| % Days Attended (Average across Middle School Grades) | 0.000*** | 0.000*** |
| | (0.000) | (0.000) |
| Number of Days Suspended (Average across Middle School Grades) | 1.740** | 1.839** |
| | (0.309) | (0.363) |
| Service Learning Hours (Average across Middle School Grades) | 0.990 | 0.991 |
| | (0.006) | (0.005) |

| | | |
|---|---|---|
| Female | 0.910 | 0.904 |
| | (0.050) | (0.102) |
| American Indian | 1.562 | 1.784 |
| | (1.064) | (1.547) |
| Asian | 0.754 | 0.878 |
| | (0.180) | (0.207) |
| Black | 1.320 | 1.570* |
| | (0.214) | (0.276) |
| Hispanic | 0.908 | 1.029 |
| | (0.096) | (0.205) |
| Ever FARMS | 0.755 | 0.800 |
| | (0.174) | (0.141) |
| Always FARMS | 0.728* | 0.748 |
| | (0.103) | (0.130) |
| Ever ESOL | 0.868 | 0.862 |
| | (0.123) | (0.137) |
| Always ESOL | 1.988 | 1.983* |
| | (1.082) | (0.563) |
| Ever SPED | 1.239 | 1.276 |
| | (0.202) | (0.244) |
| Always SPED | 0.592*** | 0.577* |
| | (0.093) | (0.133) |
| Cohort FE | Yes | Yes |
| School FE | No | Yes |
| *N* | 2503 | 2503 |

Exponentiated coefficients; standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Ultimately, the predictors of attrition are highly correlated with the predictors of participation in the IB Diploma presented in the "Results" section of this report. MYP+DP students had more available data than MYP-non DP students, which may have introduced additional bias into our estimates.

## Student Demographic and Academic Characteristics Predict IB Diploma Enrollment

To address research questions 1 and 2:

1) What are the demographic and academic characteristics of the MYP students who continue to the IB Diploma , compared with those of MYP students who do not continue to the IB Diploma ?

2) Which academic and non-academic characteristics of the MYP students predict enrollment in the IB Diploma , when demographic and school-level variables are taken into account?

we first descriptively analyzed the differences between these two groups of students using t-statistics and Chi-square tests, both which are standardized measures of group difference. These tests are useful for determining "unconditional" differences between two groups, beyond those that can be reasonably expected due to chance. Unconditional differences do not control for the contributions of any other correlated variables beyond the single one tested.

Results (tables shown in Appendix A) describe unconditional differences in descriptive student demographic and academic performance metrics between students who did and did not continue to enroll in the IB Diploma . Results also show some, although fewer, observable differences between cohorts, motivating our decision to include cohort-fixed effects in all subsequent models. The inclusion of cohort fixed effects allows us to estimate results using only within-cohort variation (e.g., differences between MYP+DP and MYP-non DP students within the same cohort of students), thereby eliminating the influence of unmeasured differences between cohorts that might bias estimates.

To answer our second question, we predicted IB Diploma enrollment with student-level characteristics using two models: a logistic regression model, and a multilevel fixed effects model. Given so few schools in our model, the fixed effects model is preferred because it only compares students within the same middle school. Therefore this model provides a more conservative estimate, given the likely non-random differences between schools. In other words, using a middle school fixed effect controls for both observed and unobserved time-invariant differences between schools, while the logit model only accounts for observed differences.

We fit model 1:

(1) $$logit(DP_{is}) = \beta_{0s} + \beta_{1s}X_{is} + \beta_{01}X_s + \gamma_{cohort} + \gamma_s + \epsilon_{is}$$

Where $DP_{is}$ is an indicator of whether or not the MYP student continued onto IB Diploma ; $X_{is}$ is a vector of student-level demographic and academic characteristics; $\gamma_{cohort}$ represents cohort fixed effects; $\gamma_s$ is a school fixed effect; and $\epsilon_{is}$ is the residual error. Functionally, we begin with a model using student-level controls only, and then add school fixed effects subsequently. The school fixed effect lets us to examine differences between students in the same school, thereby controlling for unobserved differences in IB Diploma enrollment across schools—like feeder patterns and school-specific selection requirements—that may be correlated with either student- or school-level variables of interest. In our logit model, we use cluster robust standard errors at the middle school level to account for serial correlation between students attending the same middle school.

We could not control for unmeasured factors that may have been associated with students' decision to enroll in the IB Diploma, for example their own motivation to enroll and/or parental involvement. Since these same unmeasured factors could also affect student academic performance and other outcomes of interest, these models do not provide any causal attribution of the effect of MYP+DP or MYP-non DP on subsequent student outcomes. Instead the coefficients on $X_{is}$ allow us to identify measured differences between MYP students who do and do not continue to the IB Diploma .

### Relationship Between Students' IB Participation and High School Trajectories

To address research question 3,

*3)* Which MYP student variables predict success in high school, in general, (for students who do not continue to IB Diploma ), and in the IB Diploma (for students who continue) when demographic and school-level variables are taken into account?

we used OLS and fixed effects models, similar to the approach used above. We split our sample into two groups—MYP+DP and MYP-non DP students—and fit a multilevel fixed effects model predicting success in high school:

(2) $Y_{is} = \beta_{0s} + \beta_{1s}X_{is} + \gamma_{cohort} + \gamma_s + \epsilon_{is}$

Where $Y_{is}$ is our outcome of interest; $X_{is}$ is a vector of student-level demographic and academic characteristics; $\gamma_{cohort}$ represents cohort-fixed effects; $\gamma_s$ is a school fixed effect (we include in this model fixed effects got both middle school and ninth grade); and $\epsilon_{is}$ is the residual error.

For students enrolled in IB courses (n=489) and/or enrolled in the IB Diploma (n=163), we were also interested in the factors predicting success in IB Diploma (e.g., number of courses taken, exam scores). To explore this question, we used model 2 and limited our sample to students who are enrolled in at least one IB Diploma course. Similar to the above analysis, we interpreted significant $\beta_{1s}$ coefficients as the MYP student variables predicting success in IB Diploma .

To address research question 4,

*4)* Which characteristics of the MYP students predict the numbers of IB exams taken?

we used the sample of MYP students who enrolled in one or more IB courses (including those who did and did not take TOK, n=489), and predicted the number of IB exams taken. We fit a model similar to model 2, except that we used a Poisson regression, appropriate for count outcomes. All other variables are the same as previously described. The coefficients of interest are on the vector of student-level characteristics, $\beta_{is}$, representing variation in number of IB exams taken that can be explained by each unique student-level characteristic. Again, we cannot interpret $\beta_{1s}$ in model 2 as the causal effect of enrolling in the IB Diploma, because many non-random yet unobserved district-school- and student-level characteristics may drive the choice to enroll in the IB Diploma (e.g., MYP-DP feeder pathways, motivation, peer effects).

### Propensity score matching approach to compare MYP students who did and did not continue onto IB Diploma on high school outcomes

To address the following research questions:

*5)* How does the academic performance (i.e., weighted GPA; PSAT scores, ACT scores, and SAT scores) of the MYP students who continue in the IB Diploma compare with the academic performance of MYP students who do not continue in the IB Diploma ?

*6)* How do the academic behaviors (i.e., attendance, suspension, student service learning) of MYP students who continue to the IB Diploma compare with the academic behaviors of MYP students who do not continue to the IB Diploma ?

we used a propensity score matching approach to construct observationally similar samples of treatment (i.e., enrolled in IB Diploma ) and control (i.e., not enrolled in IB Diploma ), and estimate average high school trajectory treatment effects. As the propensity score matching allows different characteristics to compensate one another, and does not guarantee matching on specific variables, we first stratified the sample by school, then conducted matching within each. This type of blocking

design helps to account for the school difference, especially if schools differ on important and unobserved aspects (e.g., feeder patterns, school-based admission requirements, effective leadership, parent involvement).

While propensity score matching may provide slightly better estimates than a standard OLS model, we caution against the interpretation of our estimates as causal evidence of the effect of MYP and/or IB Diploma on student outcomes. To provide causal estimates, the propensity score matching approach would have to assume we could predict treatment (IB Diploma enrollment) with all relevant factors, observed and unobserved. This scenario is highly unlikely given our set of student predictors, which includes limited observed factors and no way to control for unobserved factors like student motivation to enroll in the IB Diploma—as is possible through methodological approaches including randomization and regression discontinuity. Because predictors of IB Diploma enrollment also highly correlate with our outcomes of interest, the approach outlined in this section is used to identify observational differences between MYP-non DP and MYP+DP students rather than provide causal estimates.

For academic performance and behavior outcomes changing from year to year (i.e., GPA, attendance, suspension, student service learning), we created an average of each student's score on that variable across each grade of high school. For GPA, we additionally create an average of 9th-10th grade GPA and 11th-12th grade GPA, given possible differences in grading standards due to the implementation of the IB Diploma in grades 11 and 12.

Functionally, we conducted our propensity score analysis using the *psmatch2* command in Stata (Leuven & Sianesi, 2018). We predicted propensity scores using the same variables used to predict IB Diploma enrollment above, and performed an exact match within middle school cells using the procedure for matching within strata, outlined by Leuven and Sianesi (2018). Because we are unable to achieve the statistics necessary for hypothesis testing using the methods outlined above, we rely on a user-generated workaround that consists of computing propensity scores within strata and matching for the whole sample. Then, we can regress our treatment indicator on our outcome, weighted by propensity scores, to obtain our treatment estimates and the necessary parameters for hypothesis testing. (This method is detailed here: https://www.stata.com/statalist/archive/2012-02/msg01112.html.)

Using model 3, we estimated the average treatment on the treated (ATT) of IB Diploma enrollment for academic performance and behavior outcomes. We used kernel matching estimators (e.g., the Epanechnikov kernel) to construct a match for each student enrolled in IB Diploma using a weighted average of the control students (i.e., those not enrolled in IB Diploma ), and vice versa:

$$(3)\quad \theta_{T,C} = \left(\frac{1}{N^s}\right)\left[\Sigma\{\mathrm{Y_T} - \Sigma\, W(T,C)Y_C\} - \Sigma\{\mathrm{Y_C} - \Sigma\, W(C,T)Y_T\}\right]$$

Where $\theta_{T,C}$ represents the ATT for each outcome; $N^s$ represents the number of students in the treatment and control groups under the common support; $\mathrm{Y_T}$ and $\mathrm{Y_C}$ are, respectively, the average student outcomes for students in the treatment and control conditions; and $W(T,C)$ and $W(C,T)$ are kernel weights for each comparison outcome, which is a function of its distance (in terms of

propensity score) from the observation of interest. We interpreted $\theta_{T,C}$ as the relationship between the IB Diploma and each outcome conditional on the propensity score. Given that many predictors of IB Diploma enrollment (e.g., academic performance and behaviors) are also highly correlated with high school success, we do not interpret any results as causal.

As a sensitivity analysis, we also fit OLS regressions predicting academic achievement and behavioral outcomes with the same set of predictors and fixed effects as those used in our propensity score matching approach. We expected the results of our OLS models would mirror the results of our propensity score, given that both approaches rely on the same set of student covariates.

### Relationship Between Students' IB Enrollment and Postsecondary Trajectories

To understand how IB Diploma enrollment predicts academic trajectories beyond high school, we again harnessed the propensity score match approach. Our approach to questions 7 and 8:

*7)* How do the postsecondary trajectories of MYP students who continue in the IB Diploma compare with the postsecondary trajectories of the MYP students who do not continue in the IB Diploma ?

*8)* Does enrollment in the MYP coupled with enrollment in the IB Diploma increase the likelihood that students enroll in college immediately after high school graduation and/or within 2 years of high school graduation?

mirrors the matching approach described in the section above, comparing students who did and did not continue onto IB Diploma on their high school outcomes. Namely, we used model 1 to predict the probability a student enrolls in the IB Diploma , and kernel matching estimators to construct observationally-similar matched samples. Then we estimated an ATT using model 3, in which outcomes of interest (i.e., $Y_T$ and $Y_C$) are postsecondary school outcomes (i.e., student enrolled in college immediately or within two years of high school graduation) rather than secondary school outcomes.

As we did not have access to any postsecondary outcomes except college enrollment, we collapsed questions 7 and 8, creating two indicators of college enrollment: 1) an indicator equal to 1 if the student reported attending college within one year after graduation; 2) an indicator equal to 1 if the students reported attending college within one or two years after graduating high school (i.e., report of college attendance one or two years post high school graduation). The rest of our analysis follows the same methods described above as applicable to models 1 and 3.

# Results

Here we present the results of our analyses as described above by research question, or set of research questions as applicable. We use the analytic sample of 163 students who participated in the IB Diploma (i.e. MYP+DP) to address all questions except for number four, for which we use the sample of 491 students who took at least one IB exam.

## Student demographic and academic characteristics that predict IB Diploma enrollment

*1)* What are the demographic and academic characteristics of the MYP students who continue to IB Diploma, compared with those of MYP students who do not continue to the IB Diploma?

We present the results of our t-tests and chi-2 tests comparing MYP-non DP and MYP+DP students in Appendix A Tables A1 and A3. Appendix A Tables A2 and A4 present results by cohort. We find severable observed differences by cohort. To account for these differences, we use cohort fixed effects. Cohort fixed effects eliminate bias due to time invariant differences between cohorts. We found that MYP students who continued to participate in the IB Diploma (MYP+DP students) differed from those who did not (i.e. MYP-non DP students) on a variety of observed characteristics. Relative to MYP-non DP students, the probability was considerably greater that MYP+DP students would be female, white, never classified as economically disadvantaged, never English learners, and never special education students. MYP+DP students had slightly higher middle school achievement scores in math, reading, and science, higher rates of attendance, more service learning hours, and fewer suspensions per year. In summary, the MYP students who enrolled in the IB Diploma were a more economically advantaged, higher-performing group compared to those MYP students who did not continue to the IB Diploma .

*2)* Which academic and non-academic characteristics of the MYP students predict enrollment in the IB Diploma , when demographic and school-level variables are taken into account?

We present the results of our logistic regression models predicting enrollment in the IB Diploma in Appendix B Table B1. The first column shows models without school-level fixed effects and the second column adds fixed effects. The fixed effect specification, eliminating unobserved time-invariant differences between schools that may bias estimated relationships, is preferred. Coefficients are exponentiated as is typical when presenting the results of logistic regression models. Significant coefficients greater than 1 thus should be interpreted as a positive relationship with the outcome, while coefficients less than 1 should be interpreted as a negative relationship with the outcome.

While coefficients are largely similar between the two models, there are several subtle differences, likely explained by differences between schools (e.g., average attendance rates). The fixed effects model, comparing students within the same school, provides a more plausible comparison group for our IB Diploma students. Controlling for all student-level demographic variables, we found that higher middle school achievement scores in mathematics, reading, and science, as well as more average service learning hours, were associated with a greater probability of continuing onto IB Diploma

coursework. For example, each additional point on a MYP student's middle school math score is associated with a 1.3 percent (1.3=100*(1.013-1) increase in the odds of enrolling in the IB Diploma . For interpretation, a student scoring at the mean on the middle school math test has a 4 percent probability of enrolling in IB Diploma , compared to 9 percent for a student who scores one standard deviation above the mean, holding all other predictors at their means.

## Relationship between students' IB participation, demographic and academic characteristics, and high school trajectories

*3)* Which MYP student variables predict success in high school, in general, (for students who do not continue to IB Diploma ), and in the IB Diploma (for students who continue) when demographic and school-level variables are taken into account?

*4)* Which characteristics of the MYP students predict the numbers of IB exams taken?

We present, in Appendix C Table C1 and C2, the estimated coefficients models predicting measures of academic success in high school (i.e., GPA, state-wide assessment scores, PSAT/ACT/SAT scores, high school graduation) for MYP students who did and did not continue to the IB Diploma. Controlling for student- and school-level variables, we found that middle school math and science scores were associated with higher scores on almost all measures of academic success (i.e., weighted GPA, PSAT score, ACT score, and SAT score) for both MYP-non DP and MYP+DP students. For example, a 1-point increase in middle school math score was associated with a 1.91-point increase in total SAT score for MYP-non DP students and a 2.60-point increase for MYP+DP students on the same outcome. Additionally, a 1-point increase in middle school science score was associated with a 2.3- and 3.8-point increase in total SAT score for MYP-non DP and MYP+DP students, respectively. Higher middle school reading scores were associated with higher GPA and PSAT scores for MYP+DP students, and higher SAT scores for MYP-non DP students. Namely, a 1-point increase in middle school reading score was associated with a 0.003-point increase in weighted GPA and a 0.118 point increase in PSAT score for MYP+DP students. For MYP-non DP students, a 1 point increase in middle school reading score was associated with a 0.883-point increase in SAT score.

We also estimated significant relationships for both MYP-non DP and MYP+DP students between middle school academic behavior variables and high school outcomes. Higher attendance rates in the MYP were associated with higher GPAs for both groups and higher PSAT scores for the MYP+DP sample. MYP-non DP students who completed more service learning hours in middle school also tended to achieve higher high school GPAs, though the magnitude of the coefficient is quite small; a one-hour difference in average service learning hours was associated with a 0.004-point increase in high school GPA. We estimate relationships between academic behavior variables with three definitions of GPA (i.e., total high school GPA, $9^{th}$-$10^{th}$ grade GPA, and $11^{th}$-$12^{th}$ grade GPA), motivated by the different grading rigor between DP and non-DP tracks. Nonetheless, our coefficients are relatively consistent across models.

The relationship between middle school suspension and high school performance was significant, though its direction was unexpected for MYP+DP students. The interpretation of the coefficient is that more days of suspensions were associated with lower high school GPAs for MYP-non DP

students, which is intuitive, but with higher GPAs for MYP+DP students, a non-intuitive result. This unlikely result may be explained by the relatively low number of students who were suspended during middle school, particularly in the group continuing to the IB Diploma. Only two MYP+DP students were suspended during their middle school years, both of whom were reasonably high-performing with weighted GPAs greater than 4.4. Thus, while these coefficients are statistically significant, we do not interpret them as reliable estimates of the relationship between suspensions and high school achievement, given the minimal variation on this predictor present in the MYP+DP group.

We note that in Appendix C, Table C1, some variables (e.g., suspensions in the ACT [MYP+DP] model; American Indian and Always ESOL in all MYP+DP models) show missing or zero coefficients. This is the result of having no students in the category of interest on which to estimate the coefficient (e.g., American Indian students who continue onto IB Diploma ).

We present, in Appendix D Table D1, the estimated coefficients from the models the number of IB exams taken and number of IB exams taken with a score of 4 or higher. Middle school achievement scores in mathematics and reading were positively associated with a greater number of IB examinations completed with a score of 4 or higher. For example, holding all other covariates at their means, we predicted students scoring at the mean on their middle school math test would complete 2.3 IB tests with a score of 4 or higher. In comparison, we predicted students scoring one standard deviation above the mean would complete 2.5 IB tests with a score of 4 or higher. None of the middle school achievement variables were significant predictors of the number of IB exams taken.

Socioeconomic status (ever or always FARMS) was negatively associated with both the number of IB exams taken and number of IB exams with a score of 4 or higher. SPED classification was negatively associated with number of IB exams taken.

## Relationship between students' IB enrollment, and their high school trajectories

*5)* How does the academic performance of the MYP students who continue in the IB Diploma compare with the academic performance of MYP students who do not continue in the IB Diploma ?

*6)* How do the academic behaviors (e.g., attendance, suspension, student service learning) of MYP students who continue to IB Diploma compare with the academic behaviors of MYP students who do not continue to the IB Diploma ?

We present the results of both our OLS (i.e linear probability model, LPM) and propensity score models predicting high school success in Appendix E: Table E1 presents the results from OLS models, and Table E2 presents propensity score matching estimates. The first column of Table E2 shows the coefficient describing the estimated relationship between IB Diploma enrollment and outcomes, using the standard method for stratified matching in Stata (i.e., the Leuven and Sianesi estimate). This method does not permit estimation of standard errors and p-values. Thus, the second column shows the estimated IB Diploma coefficient of a user-generated method, which allowed us to estimate hypothesis testing parameters (i.e., standard errors and p-values), shown in the third and fourth columns.

Estimates across the OLS, Leuven and Sianesi, and user-generated methods are quite similar, largely driven by using the same student covariates in all three models, accounting for the same observed variation. Similarly, all three approaches are subject to omitted variable bias, and none permit causal attribution of results to the MYP or IB Diploma.

Table E1 shows IB Diploma enrollment was associated with higher GPAs, higher total PSAT scores, higher total SAT and ACT scores, and more service learning hours. For example, for MYP students, enrolling in the IB Diploma was associated with a 0.31-point increase in weighted GPA, and an 85.87-point increase in SAT total score. Additionally, enrolling in IB Diploma was associated with a 9.3-point increase in PSAT total score, a 1.2-point increase in ACT score, and a 13.7-hour increase in annual service learning hours. More service learning hours among IB Diploma students is expected, given the IB Diploma's Creativity, Action, and Service program requirements. There is no evidence of any significant differences in high school attendance or suspension outcomes between MYP students who did and did not continue onto IB Diploma .

The propensity score matching estimates shown in Table E2 tell a similar story, that IB Diploma enrollment was positively associated with weighted GPA, total PSAT score, and service learning hours.[7] The coefficients on the IB Diploma treatment are of approximately similar magnitudes to the OLS results using both the Leuven and Sianesi and mean comparison methods, as are the p-values estimated using the mean-comparison method. For example, the coefficient on GPA is 0.27 and 0.28, respectively, for the Leuven and Sianesi and mean comparison methods, compared to the coefficient of 0.31 from the OLS model. The estimated p-value from the mean comparison method suggests that this difference is highly significant ($p<0.001$), again mirroring the OLS prediction.

## Relationship between students' IB enrollment, and their postsecondary school trajectories

7) Does enrollment in the MYP coupled with enrollment in the IB Diploma increase the likelihood that students enroll in college immediately, within one year of school graduation and within two years of high school graduation?

Appendix F Tables F1 and F2 present the OLS (i.e. linear probability model) and propensity score estimates of the relationship between IB Diploma enrollment (MYP+DP) and whether students attended college one year after graduating high school, and within two years of graduating from high school. The tables are organized the same way as those presented in Appendix E.

From our propensity score model, we estimated that MYP+DP enrollment was associated in a 9 percent increase in the probability of college enrollment within one year of high school graduation. There was no evidence of significant differences between MYP+DP and MYP-non DP students in college enrollment within two years post-graduation. Given relatively small sample sizes, it was not

[7] The user-generated mean comparison method was unable to estimate the relationship between DP enrollment and SAT outcome.

possible to determine whether the null results for college enrollment within two year was the result of no relationship, or lack of sufficient statistical power to detect the effect.

# Discussion

Of the 3,147 students enrolled as sixth graders in 2006 or 2007 in district MYP schools, a small fraction, 5 percent, enrolled in the IB Diploma. A marginally larger proportion though still low, 15 percent, took one or more IB examinations. The MYP students who chose to enroll in the IB Diploma (MYP+DP) were a more economically advantaged, higher-performing group of students compared to MYP students who did not (MYP-non DP). Controlling for all available student-level demographic variables, higher middle school achievement scores in mathematics, reading, and science, as well as higher attendance rates, and more average service learning hours were associated with MYP students' greater probability of enrolling in the IB Diploma .

As expected, middle school achievement scores, particularly in science and mathematics, were strong predictors of high school success for both MYP-non DP and MYP+DP students. Middle school attendance was predictive of GPA for both groups as well.

For MYP students who participated in the DP (MYP+DP), middle school achievement scores in mathematics, reading, and science were positively associated with a greater number of IB examinations scoring 4 or higher, and middle school achievement scores in reading were positively associated with the number of IB exams taken. Socioeconomic status (ever or always FARMS) and SPED classification during middle school were negatively associated with both the number of IB exams taken and IB exams scored 4 or higher.

We also find that MYP+DP enrollment was associated with higher GPAs, higher total PSAT scores, higher total SAT and ACT scores, and more service learning hours. These results align with the demographic and prior performance characteristics of MYP students who enrolled in the IB Diploma—they were a more economically advantaged and higher-performing group.

Finally, we found that MYP+DP enrollment was associated with an 9 percentage point increase in the probability of college enrollment within one year post-graduation. There was no evidence of significant differences between MYP+DP and MYP-non DP students in college enrollment within two years post-graduation (pooled across the two years).

As noted in the introduction, there is limited prior rigorous quantitative research describing the impact of MYP on student outcomes, and existing studies are challenged by selection bias because schools offering the MYP likely differ from those that do not in both measured and unmeasured ways that also affect student outcomes. All three studies addressing MYP impact (Gordon & Bergeron, 2015; Tan & Bibby, 2010; Wade & Wolanin, 2015) found that MYP students perform better than non-MYP students on various academic outcomes measured through high school. Given the designs of the three studies, it is not possible to attribute this improved performance to the MYP itself rather than to the unmeasured characteristics of schools offering the MYP.

Studies of the IB Diploma's impact on student outcomes also show positive relationships between IB Diploma enrollment and secondary and post-secondary academic measures (e.g., Bergeron, 2015; Caspary, 2011; Coca et al., 2012; Gordon, VanderKamp, & Halic, 2015; Halic, 2013; Higher Education

Statistics Agency, 2016; Shah et al., 2010), though none except Coca et al. (2012) and Saavedra (2014) use statistical methods to address selection bias. Similar to the extant MYP research excepting the latter, it is not possible to attribute this improved performance to the IB Diploma itself rather than to the characteristics of students enrolling in the IB Diploma.

The present study faces the same challenge, such that results cannot be attributed to either the MYP or the IB Diploma. Reinforcing this concern are the very low proportions of all sample MYP students enrolling in the IB Diploma (5 percent) and taking any IB examinations (15 percent). The contribution of the present study was to build primarily upon Wade and Wolin (2015), following MYP students' outcomes beyond high school to include post-secondary enrollment and persistence. Results demonstrate a positive relationship between MYP+DP enrollment and post-secondary enrollment within one year post high school graduation. However, the relationship becomes statistically indistinguishable from zero two years post high school graduation.

## Limitations and areas of further research

As referenced above, a limitation of this study was the lack of a mechanism through which to causally attribute estimated relationships between IB Diploma enrollment and student outcomes. There would be value in conducting studies of the causal impact of the MYP and/or DP on student outcomes that harness exogenous variation such as randomized treatment assignment or regression discontinuity. A randomized controlled trial design could, for example, randomize programmatic roll-out among schools that have completed IB's candidacy process to offer the MYP and/or IB Diploma such that half implement one year and the other half implement the subsequent year. A regression discontinuity design, for example, might harness random variation around district- or school-imposed test score requirements for admission into the IB Diploma. This approach would work only with large student sample sizes and would require a high proportion of enrollment among eligible students.

Sample size was another constraint in the present study, as the respective proportions of students who enrolled in the DP and took IB examinations were low relative to the full sample. It is possible that with larger sample sizes, the relationship between DP enrollment and post-secondary enrollment within two years could have been significant. There would be value in future studies harnessing larger sample sizes.

A third limitation was the treatment definition itself. As the district's 2016 external review found and our data reflected, the district was unable to provide reliable IB course enrollment data, which we verified through attempting to align course enrollment and IB examination data. Therefore, we had to use IB examination-taking to create our DP and IB course-taking treatment indicators. It is not ideal to create treatment indicators using variables that also serve as outcomes. Ideally, future studies can access reliable IB course enrollment data, and so can create treatment indicators.

## Implications for policy and/or practice

We recommend caution in inferring implications of this study for policy and practice. One reason relates to the study limitations. Another reason is generalizability of results from 2005 and 2006 cohorts to students who may enroll in MYP schools and/or DP courses in 2020 and beyond. In 2015, the district's external review resulted in the recommendation of increasing the diversity of students

enrolled in the MYP and DP, as well as all other district choice and magnet programs. Though we do not have access to data that would confirm whether the district has taken steps leading to greater diversification as recommended, if they have, the results presented in this report are not likely to apply to current cohorts.

Another reason for which these results may not generalize between the specific study setting, including cohort timing, is that as referenced in the introduction section of this report, the IBO revised the MYP in 2014, such that the version currently implemented in schools worldwide includes new personal project requirements, improved documentation of curriculum, and greater implementation flexibility and quality assurance (Harrison, Albright, & Manlove, 2015). It is possible that results based on cohorts that participated in the pre-2014 version of the MYP might have different results than those participating after schools implemented the program changes.

# References


Australian Council for Educational Research. (2010). Primary Years Programme and Middle Years Programme student performance on the International Schools' Assessment. Retrieved from https://research.acer.edu.au/monitoring_learning/14/

Bergeron, L. (2015). Diploma Programme students' enrollment and outcomes at US postsecondary institutions 2008–2014. Bethesda, MD, USA. International Baccalaureate Organization. http://www.ibo.org/globalassets/publications/ib-research/dp/dp-student-enrollment-outcomes-us-en.pdf

Bunnell, T. (2010). The International Baccalaureate Middle Years Programme after 30 years: A critical inquiry. Journal of Research in International Education, 10 (3), 261-274.

Caspary, K. (2011). Postsecondary enrollment patterns of IB certificate and diploma candidates from U.S. high schools. Research Brief. Menlo Park, CA: SRI International. http://www.ibo.org/contentassets/d1c0accb5b804676ae9e782b78c8bc1c/postsecondaryenrollmentpatternsofibcertificateanddiplomacandidates-usdomestic-2011.pdf

Coca, V., Johnson, D., Kelley-Kemple, T., Roderick, M., Moeller, E., Williams, N., and Moragne, K. (2012). "Working to My Potential: Experiences of CPS Students in the International Baccalaureate Diploma Programme." Chicago, Illinois, USA. The Consortium on Chicago School Research.

Gordon, M. & Bergeron, L. (2015). The use of multilevel modeling and the level two residual file to explore the relationship between Middle Years Programme student performance and Diploma Programme student performance. Social Science Research, 50, 147-163.

Halic, O. (2013). "Postsecondary Educational Attainment of IB Diploma Programme Candidates from US High Schools." Geneva, Switzerland. International Baccalaureate Organization. http://www.ibo.org/contentassets/d1c0accb5b804676ae9e782b78c8bc1c/researchbrief-postsec-3eng-1.pdf

Harrison, R. Albright, E, & Manlove, S. (2015). Evolving the IB Middle Years Programme. International Schools Journal 35 (1).

International Baccalaureate Organization (2020). The International Baccalaureate Learner Profile. Retrieved from https://ibo.org/benefits/learner-profile/

International Baccalaureate Organization (2019). The International Baccalaureate: Education for a Better World. Retrieved from https://www.ibo.org

International Baccalaureate Organization (2008). MYP: From Principles into Practice. Cardiff, Wales: International Baccalaureate.

Leuven, E., & Sianesi, B. (2018). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s432001.html

Metis Associates. (2016). [District name redacted]: Study of Choice and Special Academic Programs: Report of Findings and Recommendations. Metis Associates: New York NY.

National Council for Education Statistics (2019). Common Core of Data. Retrieved from https://nces.ed.gov/ccd/

Pilchin, A., Caspary, K. & Woodworth, K. (2019). Postsecondary outcomes of IB Diploma Programme graduates in the U.S. Retrieved from https://www.ibo.org/contentassets/d1c0accb5b804676ae9e782b78c8bc1c/us-postsecondary-outcomes-final-report.pdf

Reardon, S. (2019). The Educational Opportunity Project at Stanford University. Retrieved from https://edopportunity.org

Saavedra, A. (2014). The Academic Impact of Enrollment in International Baccalaureate Diploma Programs: A Case Study of Chicago Public Schools. *Teachers College Record* 116 (5)

Shah. S, Dean, M, and Chen, YC. (2010). Academic performance of IB students entering the University of California System from 2000-2002. Geneva, Switzerland. International Baccalaureate Organization.

Tan, L. & Bibby, Y. (2011). Performance comparison between IB school students and non-IB school students on the International Schools' Assessment (ISA) and on the Social and Emotional Wellbeing Questionnaire. Report prepared for the International Baccalaureate Organization by the Australian Council for Educational Research.

U.S. Department of Education (2015). Civils Rights Data Collection. Retrieved from https://ocrdata.ed.gov/Page?t=d&eid=26931&syk=8&pid=2373

Wade, J.H. & Wolanin N.L. (2015). A comparison of MYP and non-MYP students' participation and performance in high school. Bethesda, MD, USA. International Baccalaureate Organization. http://www.ibo.org/globalassets/publications/ib-research/myp/myp-participation-and-performance-full-report.pdf

# Appendix A: Group Comparisons by Diploma Programme Status and Cohort

Table A1. T-Tests on Student- and School-Level Variables by Diploma Programme Status

| | Mean (MYP+DP) | Mean (MYP-non DP) | Diff. | Std. Error | Obs. |
|---|---|---|---|---|---|
| Middle School Math Score | 468.713 | 418.397 | -50.317*** | 5.927 | 2532 |
| Middle School Reading Score | 458.904 | 409.365 | -49.539*** | 6.340 | 2537 |
| Middle School Science Score | 466.200 | 412.902 | -53.298*** | 3.781 | 2521 |
| % Days Attended (Average across Middle School Grades) | 0.970 | 0.944 | -0.026*** | 0.006 | 2952 |
| Number of Number of Days Suspended (Average across Middle School Grades) | 0.004 | 0.078 | 0.074** | 0.024 | 3147 |
| Service Learning Hours (Average across Middle School Grades) | 25.891 | 16.489 | -9.402*** | 1.157 | 2834 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2. T-Tests on Student- and School-Level Variables by Cohort

| | Mean (Cohort=1) | Mean (Cohort=2) | Diff. | Std. Error | Obs. |
|---|---|---|---|---|---|
| Middle School Math Score | 420.679 | 422.370 | -1.691 | 2.899 | 2532 |
| Middle School Reading Score | 408.543 | 416.396 | -7.854* | 3.088 | 2537 |
| Middle School Science Score | 412.678 | 419.767 | -7.089*** | 1.882 | 2521 |
| % Days Attended (Average across Middle School Grades) | 0.950 | 0.940 | 0.010*** | 0.003 | 2952 |
| Number of Number of Days Suspended (Average across Middle School Grades) | 0.072 | 0.077 | -0.005 | 0.011 | 3147 |
| Service Learning Hours (Average across Middle School Grades) | 16.199 | 17.862 | -1.663** | 0.539 | 2834 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3. Chi-2 Tests on Categorical Variables by Diploma Programme Status

| | MYP-non DP | MYP+DP | Total |
|---|---|---|---|
| Student gender | | | |
| Male | 1,534 | 60 | 1,594 |
| Female | 1,450 | 103 | 1,553 |
| Total | 2,984 | 163 | 3,147 |
| *Pearson chi2(1) = 13.176Pr = 0.000* | | | |
| Student Race | | | |
| American Indian | 9 | 0 | 9 |
| Asian | 309 | 49 | 358 |
| Black | 806 | 19 | 825 |
| White | 1,003 | 79 | 1,082 |
| Hispanic | 857 | 16 | 873 |
| Total | 2,984 | 163 | 3,147 |
| *Pearson chi2(4) = 96.979Pr = 0.000* | | | |
| FARMS | | | |
| never | 1,464 | 132 | 1,596 |
| ever | 692 | 20 | 712 |
| always | 828 | 11 | 839 |
| Total | 2,984 | 163 | 3,147 |
| *Pearson chi2(2) = 64.763 Pr = 0.000* | | | |
| ESOL | | | |
| never | 2,018 | 128 | 2,146 |
| ever | 855 | 35 | 890 |
| always | 111 | 0 | 111 |
| Total | 2,984 | 163 | 3,147 |
| *Pearson chi2(2) = 11.574 Pr = 0.003* | | | |
| SPED | | | |
| never | 2,383 | 154 | 2,537 |
| ever | 282 | 6 | 288 |
| always | 319 | 3 | 322 |
| Total | 2,984 | 163 | 3,147 |
| *Pearson chi2(2) = 21.550 Pr = 0.000* | | | |

Table A4. Chi-2 Tests on Categorical Variables by Cohort

| | Cohort 1 | Cohort 2 | Total |
|---|---|---|---|
| Student gender | | | |
| Male | 806 | 788 | 1,594 |
| Female | 801 | 752 | 1,553 |
| Total | 1,607 | 1,540 | 3,147 |
| *Pearson chi2(1) = 0.323 Pr = 0.570* | | | |
| Student Race | | | |
| American Indian | 6 | 3 | 9 |
| Asian | 189 | 169 | 358 |
| Black | 435 | 390 | 825 |
| White | 558 | 524 | 1,082.00 |
| Hispanic | 419 | 454 | 873 |
| Total | 1,607 | 1,540 | 3,147 |
| *Pearson chi2(4) = 5.620 Pr = 0.229* | | | |
| FARMS | | | |
| never | 825 | 771 | 1,596 |
| ever | 379 | 333 | 712 |
| always | 403 | 436 | 839 |
| Total | 1,607 | 1,540 | 3,147 |
| *Pearson chi2(2) = 4.673 Pr = 0.097* | | | |
| ESOL | | | |
| never | 1,101 | 1,045 | 2,146 |
| ever | 442 | 448 | 890 |
| always | 64 | 47 | 111 |
| Total | 1,607 | 1,540 | 3,147 |
| *Pearson chi2(2) = 2.680 Pr = 0.262* | | | |
| SPED | | | |
| never | 1,296 | 1,241 | 2,537 |
| ever | 145 | 143 | 288 |
| always | 166 | 156 | 322 |
| Total | 1,607 | 1,540 | 3,147 |
| *Pearson chi2(2) = 0.090 Pr = 0.956* | | | |

# Appendix B: Student Demographic and Academic Characteristics Predicting IB Diploma Enrollment

Table B1. Models Predicting Enrollment in the IB Diploma

| | (1) | (2) |
|---|---|---|
| Middle School Math Score | 1.014* | 1.013** |
| | (0.006) | (0.005) |
| Middle School Reading Score | 1.012* | 1.012** |
| | (0.005) | (0.004) |
| Middle School Science Score | 1.012* | 1.013** |
| | (0.005) | (0.004) |
| % Days Attended (Average across MS Grades) | 410.078 | 2951.627 |
| | (2229.854) | (12260.106) |
| Number of Days Suspended (Average across MS Grades) | 0.159 | 0.107 |
| | (0.252) | (0.212) |
| Service Learning Hours (Average across MS Grades) | 1.012* | 1.011* |
| | (0.005) | (0.004) |
| Female | 1.338 | 1.377 |
| | (0.232) | (0.277) |
| Asian | 2.847* | 2.644*** |
| | (1.447) | (0.705) |
| Black | 1.434 | 1.377 |
| | (0.530) | (0.456) |
| Hispanic | 1.718 | 1.890 |
| | (0.816) | (0.668) |
| Ever FARMS | 0.573 | 0.594 |
| | (0.257) | (0.190) |
| Always FARMS | 0.336*** | 0.369** |
| | (0.109) | (0.142) |
| Ever ESOL | 1.289 | 1.243 |
| | (0.550) | (0.343) |
| Ever SPED | 0.537 | 0.537 |
| | (0.241) | (0.244) |
| Always SPED | 0.570 | 0.676 |
| | (0.449) | (0.501) |
| Cohort FE | Yes | Yes |
| School FE | No | Yes |
| *N* | 2413 | 2413 |

Exponentiated coefficients; Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Appendix C: Relationship Between Academic Characteristics and High School Trajectories for MYP Students Who Did and Did Not Proceed to the IB Diploma

Table C1. Predictors of High School Success for MYP-non DP and MYP+DP Students

| | (1) Weighted GPA (MYP-non DP) | (2) Weighted GPA (MYP+DP) | (3) 9-10 Weighted GPA (MYP-non DP) | (4) 9-10 Weighted GPA (MYP+DP) | (5) 11-12 Weighted GPA (MYP-non DP) | (6) 11-12 GPA (MYP+DP) |
|---|---|---|---|---|---|---|
| MS Math Score | 0.001*** | 0.005*** | 0.001*** | 0.005** | 0.002*** | 0.005*** |
| | (0.000) | (0.001) | (0.000) | (0.001) | (0.000) | (0.001) |
| MS Reading Score | 0.000 | 0.003* | 0.000 | 0.003* | 0.000 | 0.002 |
| | (0.000) | (0.001) | (0.000) | (0.001) | (0.000) | (0.001) |
| Middle School Science Score | 0.011*** | 0.003* | 0.011*** | 0.003* | 0.010*** | 0.002* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| % Days Attended (Average across MS Grades) | 5.460*** | 4.316*** | 5.415*** | 4.296*** | 4.375*** | 4.306*** |
| | (0.455) | (1.182) | (0.456) | (1.241) | (0.507) | (1.174) |
| Number of Days Suspended (Average across MS Grades) | -0.313*** | 1.878** | -0.299*** | 2.009** | -0.373*** | 1.750* |
| | (0.067) | (0.691) | (0.067) | (0.725) | (0.082) | (0.686) |
| Service Learning Hours (Average across MS Grades) | 0.004*** | 0.001 | 0.004*** | 0.001 | 0.004*** | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Female | 0.340*** | 0.119* | 0.347*** | 0.134* | 0.344*** | 0.105 |
| | (0.032) | (0.056) | (0.032) | (0.058) | (0.032) | (0.055) |
| American Indian | -0.262 | | -0.227 | | -0.245 | |
| | (0.279) | | (0.279) | | (0.285) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Asian | 0.008 | -0.084 | 0.002 | -0.063 | 0.009 | -0.104 |
| | (0.061) | (0.071) | (0.061) | (0.075) | (0.058) | (0.071) |
| Black | -0.305*** | -0.443*** | -0.287*** | -0.409** | -0.260*** | -0.475*** |
| | (0.051) | (0.096) | (0.051) | (0.100) | (0.051) | (0.095) |
| Hispanic | -0.272*** | -0.380*** | -0.256*** | -0.379** | -0.241*** | -0.380*** |
| | (0.055) | (0.105) | (0.055) | (0.111) | (0.053) | (0.105) |
| Ever FARMS | -0.183*** | -0.097 | -0.178*** | -0.119 | -0.185*** | -0.077 |
| | (0.049) | (0.091) | (0.049) | (0.096) | (0.048) | (0.091) |
| Always FARMS | -0.168** | 0.080 | -0.163** | 0.036 | -0.170*** | 0.122 |
| | (0.052) | (0.125) | (0.052) | (0.131) | (0.051) | (0.124) |
| Ever ESOL | 0.066 | -0.001 | 0.061 | -0.002 | 0.053 | 0.000 |
| | (0.044) | (0.070) | (0.044) | (0.074) | (0.044) | (0.070) |
| Always ESOL | 0.006 | | 0.062 | | 0.020 | |
| | (0.101) | | (0.102) | | (0.102) | |
| Ever SPED | 0.044 | 0.034 | 0.029 | -0.019 | 0.085 | 0.087 |
| | (0.058) | (0.129) | (0.058) | (0.136) | (0.057) | (0.129) |
| Always SPED | -0.042 | -0.673* | -0.039 | -0.491 | -0.088 | -0.854** |
| | (0.061) | (0.316) | (0.061) | (0.332) | (0.061) | (0.314) |
| Cohort FE | Yes | Yes | Yes | Yes | Yes | Yes |
| School FE | Yes | Yes | Yes | Yes | Yes | Yes |
| *N* | 1588 | 153 | 1588 | 153 | 1390 | 153 |

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C2. Predictors of High School Success for MYP-non DP and MYP+DP Students

| | (1) PSAT Total (MYP-non DP) | (2) PSAT Total (MYP+DP) | (3) ACT (MYP-non DP) | (4) ACT (MYP+DP) | (5) SAT Total (MYP-non DP) | (6) SAT Total (MYP+DP) |
|---|---|---|---|---|---|---|
| MS Math Score | 0.048*** | 0.190** | 0.027*** | 0.084*** | 1.910*** | 2.599*** |
| | (0.011) | (0.068) | (0.005) | (0.020) | (0.198) | (0.653) |
| MS Reading Score | -0.006 | 0.118* | -0.003 | 0.012 | 0.883*** | 0.365 |
| | (0.010) | (0.059) | (0.005) | (0.017) | (0.189) | (0.560) |
| Middle School Science Score | 0.425*** | 0.242*** | 0.067*** | 0.014 | 3.808*** | 2.321*** |
| | (0.015) | (0.050) | (0.006) | (0.016) | (0.220) | (0.471) |
| % Days Attended (Average across MS Grades) | 23.278 | 120.795* | 10.335 | 23.898 | 203.457 | 645.253 |
| | (14.597) | (60.525) | (6.679) | (15.263) | (195.312) | (594.742) |
| Number of Days Suspended (Average across MS Grades) | 1.754 | -10.091 | -2.515 | 0.000 | -72.700 | 257.517 |
| | (2.621) | (34.438) | (1.449) | (.) | (46.152) | (324.221) |
| Service Learning Hours (Average across MS Grades) | 0.029 | 0.054 | 0.011 | 0.019 | 0.009 | 0.236 |
| | (0.027) | (0.063) | (0.009) | (0.015) | (0.355) | (0.629) |
| Female | 2.476** | -0.955 | -0.224 | -1.633* | 13.215 | -57.912* |
| | (0.889) | (2.822) | (0.326) | (0.630) | (10.960) | (26.926) |
| American Indian | 1.944 | | 1.614 | | 12.155 | |
| | (7.765) | | (3.115) | | (111.241) | |
| Asian | -3.283* | -4.765 | -1.760** | -2.350** | -16.852 | -62.697 |
| | (1.615) | (3.554) | (0.645) | (0.831) | (18.263) | (33.557) |
| Black | -7.665*** | -9.378 | -3.120*** | -2.236 | -86.537*** | -148.463** |
| | (1.416) | (5.044) | (0.500) | (1.415) | (17.082) | (47.936) |
| Hispanic | -6.288*** | -5.955 | -2.487*** | -4.957** | -83.555*** | -119.978* |
| | (1.506) | (5.407) | (0.600) | (1.790) | (18.196) | (49.595) |
| Ever FARMS | -4.123** | -7.133 | -0.608 | 3.707* | -72.827*** | 1.219 |
| | (1.363) | (4.624) | (0.605) | (1.375) | (16.565) | (42.590) |
| Always FARMS | -5.169*** | -1.734 | -1.252* | 1.413 | -73.633*** | -8.844 |
| | (1.452) | (6.323) | (0.586) | (1.457) | (17.859) | (60.202) |
| Ever ESOL | -1.378 | -5.897 | -0.063 | -1.830 | -13.538 | -13.622 |
| | (1.250) | (3.554) | (0.540) | (1.069) | (15.279) | (32.763) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Always ESOL | -10.785*** | | -3.797** | | -159.823*** | |
| | (2.998) | | (1.376) | | (47.315) | |
| Ever SPED | -2.152 | -0.739 | -1.026 | -2.050 | -4.153 | 90.114 |
| | (1.578) | (6.447) | (0.590) | (1.387) | (19.301) | (59.997) |
| Always SPED | -9.802*** | -18.169 | -1.662* | -7.964** | -77.066** | -419.753** |
| | (1.727) | (15.773) | (0.701) | (2.461) | (26.696) | (150.015) |
| Cohort FE | Yes | Yes | Yes | Yes | Yes | Yes |
| School FE | Yes | Yes | Yes | Yes | Yes | Yes |
| *N* | 1309 | 149 | 407 | 61 | 927 | 140 |

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Appendix D: Relationship Between MYP+DP Students' Academic Characteristics and IB Diploma Success

Table D1. Predictors of DP Success for MYP+DP Students

| | (1) Number of IB Exams | (2) Number of IB Exams with 4+ |
|---|---|---|
| Middle School Math Score | 0.002 | 0.004** |
| | (0.001) | (0.001) |
| Middle School Reading Score | 0.002 | 0.004** |
| | (0.001) | (0.001) |
| Middle School Science Score | 0.001 | 0.002 |
| | (0.001) | (0.001) |
| % Days Attended (Average across MS Grades) | 0.703 | -0.206 |
| | (1.239) | (1.302) |
| Number of Days Suspended (Average across MS Grades) | -0.346 | -1.035 |
| | (0.471) | (0.733) |
| Service Learning Hours (Average across MS Grades) | -0.001 | -0.000 |
| | (0.001) | (0.001) |
| Female | 0.058 | 0.121 |
| | (0.063) | (0.068) |
| American Indian | -0.180 | 0.014 |
| | (0.713) | (0.715) |
| Asian | 0.059 | 0.010 |
| | (0.087) | (0.092) |
| Black | 0.042 | -0.119 |
| | (0.104) | (0.121) |
| Hispanic | 0.153 | 0.030 |
| | (0.113) | (0.129) |
| Ever FARMS | -0.261* | -0.352** |
| | (0.110) | (0.123) |
| Always FARMS | -0.278* | -0.262 |
| | (0.137) | (0.156) |
| Ever ESOL | -0.016 | 0.024 |
| | (0.093) | (0.102) |
| Always ESOL | -0.804 | -0.329 |
| | (0.587) | (0.589) |
| Ever SPED | -0.261* | -0.235 |
| | (0.118) | (0.131) |
| Always SPED | -0.137 | -0.224 |
| | (0.202) | (0.248) |
| Cohort FE | Yes | Yes |
| School FE | Yes | Yes |
| *N* | 466 | 466 |

Standard errors in parentheses; * p < 0.05, ** p < 0.01, *** p < 0.001

# Appendix E: Relationship Between Students' IB Diploma Enrollment and High School Trajectories

Table E1. OLS Covariate Adjusted Relationships Between IB Diploma Enrollment Status and High School Outcomes

| | (1) Weighted GPA (HS avg) | (2) Weighted GPA (9-10) | (3) Weighted GPA (11-12) | (4) PSAT Total | (5) ACT Total | (6) SAT Total | (7) % Attendance (HS avg.) | (8) Suspensions (HS avg.) |
|---|---|---|---|---|---|---|---|---|
| MYP+DP enrollment | 0.306*** | 0.296*** | 0.292*** | 9.334*** | 1.200** | 85.868*** | 0.008 | 0.015 |
| | (0.054) | (0.055) | (0.050) | (1.436) | (0.445) | (15.321) | (0.005) | (0.017) |
| Middle School Math Score | 0.002*** | 0.002*** | 0.002*** | 0.050*** | 0.028*** | 2.042*** | 0.000 | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.011) | (0.005) | (0.188) | (0.000) | (0.000) |
| Middle School Reading Score | 0.000 | 0.000 | 0.000 | -0.001 | -0.003 | 0.930*** | -0.000 | -0.000 |
| | (0.000) | (0.000) | (0.000) | (0.010) | (0.005) | (0.177) | (0.000) | (0.000) |
| Middle School Science Score | 0.010*** | 0.010*** | 0.009*** | 0.405*** | 0.066*** | 3.433*** | 0.000** | -0.001*** |
| | (0.000) | (0.000) | (0.000) | (0.014) | (0.006) | (0.196) | (0.000) | (0.000) |
| % Days Attended (MS avg.) | 5.588*** | 5.548*** | 4.516*** | 30.937* | 10.192 | 232.171 | 0.958*** | -0.178 |
| | (0.434) | (0.435) | (0.479) | (14.168) | (6.219) | (185.206) | (0.041) | (0.139) |
| Number of Days Suspended (MS avg.) | -0.330*** | -0.316*** | -0.385*** | 1.366 | -2.304 | -76.289 | -0.039*** | 0.334*** |
| | (0.065) | (0.065) | (0.080) | (2.607) | (1.413) | (45.309) | (0.006) | (0.021) |
| Service Learning Hours (MS avg.) | 0.004*** | 0.004*** | 0.003*** | 0.036 | 0.011 | 0.048 | 0.000 | -0.000 |
| | (0.001) | (0.001) | (0.001) | (0.025) | (0.008) | (0.315) | (0.000) | (0.000) |
| Female | 0.325*** | 0.332*** | 0.325*** | 2.351** | -0.263 | 3.061 | -0.010*** | -0.016 |
| | (0.030) | (0.030) | (0.029) | (0.845) | (0.298) | (10.165) | (0.003) | (0.010) |
| American Indian | -0.294 | -0.258 | -0.266 | 1.409 | 1.752 | 6.257 | 0.023 | 0.044 |
| | (0.271) | (0.272) | (0.277) | (7.767) | (3.051) | (110.568) | (0.025) | (0.087) |
| Asian | -0.032 | -0.035 | -0.030 | -3.787** | -1.841*** | -25.602 | 0.009 | -0.009 |
| | (0.053) | (0.053) | (0.051) | (1.446) | (0.528) | (15.918) | (0.005) | (0.017) |
| Black | -0.337*** | -0.319*** | -0.298*** | -8.420*** | -3.143*** | -95.666*** | 0.002 | 0.029 |
| | (0.048) | (0.048) | (0.047) | (1.349) | (0.471) | (15.916) | (0.005) | (0.015) |
| Hispanic | -0.301*** | -0.285*** | -0.271*** | -6.907*** | -2.496*** | -90.024*** | 0.001 | -0.020 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (0.051) | (0.051) | (0.049) | (1.426) | (0.553) | (16.869) | (0.005) | (0.016) |
| Ever FARMS | -0.173*** | -0.169*** | -0.172*** | -3.902** | -0.390 | -62.548*** | -0.002 | 0.012 |
| | (0.046) | (0.046) | (0.045) | (1.305) | (0.555) | (15.478) | (0.004) | (0.015) |
| Always FARMS | -0.157** | -0.153** | -0.155** | -5.023*** | -0.993 | -71.047*** | -0.002 | 0.040* |
| | (0.049) | (0.049) | (0.048) | (1.407) | (0.548) | (16.998) | (0.005) | (0.016) |
| Ever ESOL | 0.059 | 0.055 | 0.046 | -1.723 | -0.201 | -15.043 | -0.006 | -0.032* |
| | (0.041) | (0.041) | (0.040) | (1.180) | (0.491) | (13.999) | (0.004) | (0.013) |
| Always ESOL | -0.025 | 0.032 | -0.016 | -11.280*** | -3.944** | -171.399*** | 0.003 | -0.047 |
| | (0.098) | (0.099) | (0.098) | (2.981) | (1.339) | (46.834) | (0.009) | (0.032) |
| Ever SPED | 0.038 | 0.024 | 0.077 | -2.098 | -0.936 | -3.013 | 0.007 | -0.001 |
| | (0.055) | (0.055) | (0.054) | (1.531) | (0.548) | (18.382) | (0.005) | (0.018) |
| Always SPED | -0.053 | -0.047 | -0.099 | -9.391*** | -1.422* | -79.628** | 0.000 | 0.016 |
| | (0.059) | (0.059) | (0.058) | (1.699) | (0.656) | (25.659) | (0.006) | (0.019) |
| Cohort FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| MS FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 9th grade school FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| *N* | 1741 | 1741 | 1543 | 1458 | 468 | 1067 | 1686 | 1741 |

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table E2. Propensity Score Matching Estimates of the Relationship Between IB Diploma Enrollment and High School Outcomes

| | Leuven and Sianesi estimate | Mean comparison (ATT) | Mean comparison (SE) | Mean comparison (p) |
|---|---|---|---|---|
| Weighted GPA | 0.270 | 0.284*** | 0.053 | 0.000 |
| Weighted GPA (9-10) | 0.260 | 0.272*** | 0.053 | 0.000 |
| Weighted GPA (11-12) | 0.258 | 0.269*** | 0.054 | 0.000 |
| PSAT Total | 5.868 | 8.327** | 2.575 | 0.001 |
| SAT Total | 113.712 | 113.738*** | 25.456 | 0.000 |
| % attendance (HS avg.) | 0.006 | 0.008* | 0.003 | 0.029 |
| Number of Suspensions (HS avg.) | -0.011 | -0.005 | 0.003 | 0.103 |
| Service Learning Hours (HS avg.) | 14.367 | 14.806** | 5.261 | 0.005 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Appendix F: Relationship Between Students' IB Diploma Enrollment and Postsecondary School Trajectories

Table F1. OLS Covariate-Adjusted Relationship Between DP Enrollment and College Enrollment

| | (1) Immediate College Enrollment (Within 1 Year) | (2) College Enrollment (Within 2 Years) |
|---|---|---|
| IB Diploma enrollment | 0.098* | 0.063 |
| | (0.042) | (0.040) |
| Middle School Math Score | 0.000 | 0.000 |
| | (0.000) | (0.000) |
| Middle School Reading Score | 0.000 | 0.000 |
| | (0.000) | (0.000) |
| Middle School Science Score | 0.001* | 0.001* |
| | (0.000) | (0.000) |
| % Days Attended (Average across Middle School Grades) | 1.989*** | 2.178*** |
| | (0.331) | (0.321) |
| Number of Days Suspended (Average across Middle School Grades) | -0.175*** | -0.159*** |
| | (0.050) | (0.048) |
| Service Learning Hours (Average across Middle School Grades) | 0.002* | 0.001 |
| | (0.001) | (0.001) |
| Female | 0.013 | 0.029 |
| | (0.023) | (0.022) |
| American Indian | -0.358 | -0.217 |
| | (0.207) | (0.201) |
| Asian | 0.006 | 0.012 |
| | (0.041) | (0.039) |
| Black | -0.009 | -0.020 |
| | (0.037) | (0.035) |
| Hispanic | -0.087* | -0.087* |
| | (0.039) | (0.038) |
| Ever FARMS | -0.016 | -0.020 |
| | (0.035) | (0.034) |
| Always FARMS | -0.050 | -0.067 |
| | (0.038) | (0.036) |
| Ever ESOL | 0.032 | 0.007 |
| | (0.031) | (0.030) |
| Always ESOL | -0.087 | -0.136 |
| | (0.075) | (0.073) |
| Ever SPED | 0.107* | 0.079 |
| | (0.042) | (0.040) |

| | | |
|---|---|---|
| Always SPED | -0.029 | -0.052 |
| | (0.045) | (0.044) |
| Cohort FE , MS FE, 9th grade school FE | Yes | Yes |
| *N* | 1741 | 1741 |

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table F2. Propensity Score Relationship Between IB Diploma Enrollment and Probability of Enrollment in College Immediately and Within Two Years Post-Graduation from High School

| | Leuven and Sianesi estimate | Mean comparison (Difference) | Mean comparison (Standard error) | Mean comparison (p-value) |
|---|---|---|---|---|
| Immediate College Enrollment (Within 1 Year) | 0.158 | 0.090* | 0.042 | 0.034 |
| College Enrollment (Within 2 Years) | 0.081 | 0.061 | 0.039 | 0.123 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$