"A Little Now for a Lot Later: A Look at a Texas Advanced Placement Incentive Program"

C. Kirabo Jackson, Dec 7, 2007
Cornell University
e-mail: ckj5@cornell.edu

Abstract: I analyze a program that was implemented in schools serving underprivileged populations in Texas that pays both students and teachers for passing grades on advanced placement examinations. Exploiting the fact that different schools adopted the program at different times, I use a difference-in-differences strategy. I compare the changes in aggregate student outcomes, before and after adoption, for schools adopting the AP incentive program to the changes experienced over the same time period for carefully selected groups of comparison schools. Adoption of the AP incentive program is associated with a 30 percent increase in the number of students scoring above 1100 on the SAT or 24 on the ACT, and an 8 percent increase in the number of students who matriculate in college in Texas. The per-student costs of the program are very small relative to reasonable estimates of the implied lifetime benefits that accrue to affected students such that the APIP may ameliorate sub-optimal educational investments. Empirical evidence suggests that teachers and students were not simply aiming to maximize their rewards. Anecdotal evidence suggests that the increases in AP participation were due to better access to AP courses, changes in teacher and peer norms towards AP courses, and better student information.

JEL: I20 J24 J00

I am very grateful for the guidance received from Roland Fryer, Caroline Hoxby, and Larry Katz. I gratefully acknowledge helpful comments from Dale Jorgenson and Claudia Goldin as well as participants at the Harvard Labor lunch and the Harvard Econometrics lunch. I am also grateful for the many helpful conversations I had with Walter Dewar of AP Strategies. I also thank Elias Bruegmann, Li Han, Gauri Kartini Shastry, Katharine Emans Sims, Erin Strumpf, and Daniel Tortorice. All errors are, of course, my own.

I. Introduction

The Advanced Placement Incentive Program (APIP) is a novel program that includes cash incentives for both teachers and students for each passing score earned on an Advanced Placement (AP) exam. The APIP was first implemented in ten Dallas schools in 1996, and has been expanded to over forty schools in Texas. The program is targeted primarily to low-income, minority-majority school districts with a view towards improving college readiness. Due to the perceived success of this program, a similar scheme has been adopted in New Mexico and several other US districts are following suit.[1] Cash incentives for students is an understudied phenomena, thus given their growing popularity, it is important to understand (1) whether student incentive programs work, (2) how students of different gender and race may respond to such programs (3) the possible deleterious effects of such programs and (4) the mechanisms through which these incentives may affect student outcomes. To this end, I analyze the APIP, estimate its effect on outcomes such as high-school graduation, SAT/ACT performance, and college matriculation and then I investigate possible mechanisms through which the APIP may affect school outcomes.

Across the United States, college matriculation and completion rates for low-income and minority students are much lower than those for non-poor whites.[2] Much of the gaps in degree attainment across groups occur for those who enter college but do not graduate. Researchers have investigated a variety of hypotheses that could explain differences in college going rates across groups. Hauser (1993) and Kane (1994) find that differences in college-going rates across the different ethnic groups are related to differences in the ability to pay for college and differences in parental education. Heckman and Cameron (2001) find that the long-run factors associated with family environment such as parental education account for most of the differences in college going across ethnic and racial groups. They suggest that policies that improve scholastic ability in high-school and earlier may be far more effective at reducing the

[1] Arkansas [John Lyon (2007)] and New York City [Jennifer Medina (2007)] will adopt similar programs in 2008. Schools in Washington D.C., Virginia and Maryland have announced interest in adopting similar programs [Jay Mathews (2007)].

[2] Using the August 2006 Current Population Survey, I find that 71 percent of white high-school graduates or GED holders between the ages of 25 and 29 ever enrolled in some college program. The corresponding figures are 60 and 52 percent for blacks and Hispanics respectively. The implied two or four-year college completion rates for these same groups are 68% for whites, 51% for blacks and 53% for Hispanics. Similar college completion patterns exist in Texas. For the cohort of students who entered any college in Texas in the fall of 1997, the six year graduation rate is 41 percent for whites compared to only 24 percent for both blacks and Hispanics.

educational differentials than policies that reduce the financial burdens of attending college.

AP courses and exams have received much national attention as a means to make tertiary education more accessible to low-income and minority populations.[3] The rationale behind this push to increase AP participation is the observation that students who take AP courses and examinations are much more likely to enroll and be successful in college than their peers, as measured by college GPA and graduation rates [Eimers (2003); Dougherty et al. (2006); Geiser and Santelices (2004); Morgan and Ramist (1998); Dodd et al. (2002)]. It has also been documented that students who take more rigorous math and science courses in high-school, such as AP courses, have significantly higher SAT scores [College Board (2003)].

There is a self-selection problem in comparing AP-takers to non-AP takers, since observationally similar students who take AP courses may differ in unobservable dimensions such as motivation. Although some studies attempt to account for this self-selection, none are able to isolate exogenous variation in AP taking to uncover a causal relationship. The APIP, however, pays students and teachers for each passing score on an AP exam[4] which produces exogenous variation in AP taking that is unrelated to students' intrinsic motivation. At the first ten Dallas high-schools to adopt the program in 1996, the number of students taking AP exams in math, English and science increased from 269 in 1995 to 729 in 1996. Figure 1 shows that by 2002, those schools had 132 passing scores per 1,000 juniors and seniors taking math, science and English, compared to 86 in Texas and 80 in the U.S. [AP Strategies (2006)].

Dallas independent school district is a low-income, minority majority district. These populations have much lower participation rates in AP courses [Klopfenstein (2004)]. There may be large lifetime benefits to AP-taking that would be much larger than the monetary rewards paid out. I present a rational choice model of student AP participation that suggests that the large increases in AP participation rates were likely not simply a response to the cash rewards, but may have been due to inefficiencies that were ameliorated by the APIP, such as imperfect

[3] In his 2004 State of the Union Address, President Bush announced a plan in which he proposed an increase from $24 to $52 million annually for the AP program authorized in the No Child Left Behind Act to support state and local efforts to increase access to AP classes and tests (and other challenging curricular end-of-course examinations) for students in low-income schools.[3] Several states have implemented programs with the same objective. A good example is the Western Consortium for Accelerated Learning Opportunities (WCALO) consisting of Arizona, Colorado, Hawaii, Idaho, Montana, New Mexico, Oregon, South Dakota, and Utah. [http://www.whitehouse.gov/news/releases/2004/01/20040120-7.html]

[4] See section II for a detailed description of the financial incentives.

student information[5], student myopia[6], suboptimal teacher effort, or suboptimal track placement.

The effect of the APIP reflects a combination of the effect of AP course taking, and other aspects of the program such as increased teacher training and student and teacher incentives. Research indicates that students and teachers respond to incentives for the tasks for which the incentives are provided.[7] In the most closely related of these, Angrist and Lavy (2002) find that students in schools that were eligible for cash rewards for passing the national matriculation examinations in Israel were approximately 7 percent more likely to pass the examinations.[8] Some studies suggest, however, that providing incentives for students without the cooperation of teachers may be ineffective [Kremer, Miguel and Thornton (2004), Angrist and Lavy (2002)]. In addition, an agency model with multiple tasks [Holmstrom and Milgrom (1991)] implies that incentives to perform on one dimension may cause agents to withdraw effort from other dimensions of performance. For instance, teachers may spend less time on untested material when they are rewarded *only* for students' test performance and students may withdraw from difficult courses to maintain high grades when they are rewarded for having a high grade-point average [Glewwe, Ilias and Kremer (2003); Cornwell et al (2005); Binder and Ganderton (2002)]. In fact, in certain situations, a poorly designed incentive system could have deleterious effects on the overall set of outcomes that it was intended to improve. This underscores the importance of looking at the effect of the APIP on both AP and non-AP outcomes.

I use school-level data from the Texas Education Agency to evaluate the effect of the APIP on outcomes.[9] I identify the effects of the program by comparing the change in outcomes of cohorts within the same school, before and after adoption of the program, to the change in outcomes for comparison schools over the same time period. By comparing cohorts within the same school, I am able to eliminate self-selection within a cohort - that is, the self selection that ordinarily makes one student enroll in AP courses and another not to do so. Because the program

[5] According to Card (1995), educational attainment is the result of a lifetime utility maximization problem based on information available to the student at the time. In this framework, new information could change students' educational choices.

[6] See Frederick, Loewenstein, and O'Donoghue (2002) for a discussion of time discounting and time preference.

[7] This list includes Angrist, Bettinger, Bloom, King and Kremer (2002), Kremer, Miguel and Thornton (2004), Lavy (2002), Lavy (2004), Hollenbeck and Stone (2002), Atkinson et al. (2004). See Angrist, Oreopoulos and Lang (2007) for an overview of this literature. There exists a psychology literature documenting that external incentives for children can replace intrinsic motivation such that effort and performance may be worse after the incentives are removed than if they had never been introduced. See Alfie Kohn (1993) for an overview of this research.

[8] In a recent revision of this paper, Angrist and Lavy (2007) they find this effect only for girls and not boys.

[9] While individual student-level data would be optimal for this analysis, such data are not made available due to Texas's interpretation of the Family Educational Rights and Privacy Act (FERPA).

was not randomly adopted across schools, the remaining endogeneity concern is that the schools that adopted the APIP were somehow different than other schools. I eliminate this second form of self-selection by exploiting the fact that administrators could not roll out the program in all interested schools at once. I use as my main control group, the schools that had already decided to adopt the APIP but had not yet had the opportunity to implement it. My control group also helps me account for the effect of potentially confounding policies such as the Texas 10% rule and the statewide Texas Advanced Placement incentive program.[10] My identification strategy is valid so long as the schools have the same incoming distribution of students in their pre-program and post-program cohorts. Because this is an important restriction, I test it empirically and am able to eliminate most plausible scenarios of changing student selection into program schools.

The results show that cohorts in schools affected by the program have more students with high SAT/ACT scores, and more students who matriculate at a college in Texas. While there are no differences by gender, some specifications suggest that the relative improvements in SAT/ACT performance may be largest for minority students. I find no evidence that the APIP schools diverted resources away from other students towards those taking AP courses. I present several pieces of evidence that, when taken in its entirety, suggest that the response was not simply due to students and teachers maximizing their cash rewards. In particular; (1) the increased AP participation does not reflect a substitution away from other advanced courses, (2) the effect of the APIP was no stronger in schools with higher cash rewards (3) AP course enrollment went up for all AP courses even if rewards were only given for certain subjects. This evidence is consistent with the theoretical framework and corroborates claims by guidance counselor s that the increased AP participation is due to increased encouragement, better student information and changes in teacher and peer norms.

The structure of the remainder of this paper is as follows. Section II describes the APIP. Section III describes the data and highlights the differences between the schools selected for the APIP and other high-schools in Texas. Section IV lays out the theoretical framework within which to think about the APIP's effect on AP participation and the other outcomes. Section V motivates and describes the empirical strategy. Section VI analyzes the results. Section VII concludes.

[10] For a detailed description of these policies see the appendix note A1.

II. Description of the AP incentive program

Advanced placement courses are typically taken by students in 11th or 12th grade. The courses are intended to be "college level" and most colleges allow successful AP exam takers to use them to offset degree requirements.[11] Although it is unclear whether AP courses are actually equivalent to courses at all colleges, the fact that selective colleges pay considerable attention to a student's AP scores in the admissions process demonstrates that the exams are considered to be revealing about a student's likely preparation for and achievement in college. The AP program has 35 courses and examinations across 20 subject areas. The length of a course varies from one to two semesters, depending on the pace chosen by the teacher and the scope of the subject taken.[12] The cost per examination is $82 and a fee reduction of $22 is granted to those students with demonstrated financial need. AP exams are administered by the College Board making the type of cheating documented in Jacob and Levitt (2003) unlikely. The exams are graded from 1 through 5, with five being the highest and 3 generally regarded as a passing grade.

The APIP is run by AP Strategies, a non-profit organization based in Dallas. The APIP is entirely voluntary for schools, teachers, and students. The heart of the program is a set of financial incentives for teachers and students based on AP examination performance. The APIP also includes teacher training conducted by the College Board, and a curriculum that prepares students for AP courses in earlier grades. The APIP uses "vertical teams" of teachers. At the top of a vertical team is a lead teacher who not only teaches students, but also spends time providing training for other AP teachers. Vertical teams also include teachers whose grade precedes those in which AP courses are offered. For example a vertical team might create a math curriculum designed to prepare students for AP calculus in 12th grade. This curriculum might start as early as the seventh grade.[13] This aspect of the APIP suggests that some of the benefits to the program may be felt several years after it is first introduced at a school.

The APIP's monetary incentives are intended to encourage both participation and induce effort in AP courses. Lead teachers receive between $3,000 and $10,000 annual salary bonus, and a further $2,000 to $5,000 bonus opportunity based on results. Pre-AP teachers receive an annual supplement of between $500 and $1,000 per year for extra work. AP teachers receive

[11] While this is true in general, the most selective colleges often limit the number of AP credits a student can use. Some selective colleges only allow students to use AP credits to pass out of prerequisites, but *not* towards actual graduation credit.
[12] *Source*: College Board website. http://www.collegeboard.com/student/testing/ap/about.html.
[13] The College Board publishes recommended curricula for the earlier grades.

between \$100 and \$500 for each AP score of 3 or over earned by an 11th or 12th grader enrolled in their course. In addition, AP teachers can receive bonuses of up to \$1,000 based on results. In each school, these bonuses are discretionary. While the amount paid per passing AP score and the salary supplements are well defined, there is variation across schools in the amounts paid. Overall, the APIP can deliver a considerable increase in compensation for teachers.[14]

Students in 11th and 12th grade also receive monetary incentives for performance. The program pays half of each student's examination fees so that students on free or reduced lunch would pay 15 dollars (instead of 30) while those who are not would pay 30 dollars (instead of 60) per exam. Students receive between \$100 and \$500 for each score of 3 or above in an eligible subject for which they took the course. The amount paid per exam is well defined in each school, but there is variation across schools in the amount paid per passing AP exam. This could amount to several hundred dollars for a student who takes and passes several AP examinations during their 11th and 12th grades. For example, one student earned \$700 in his junior and senior years for passing scores in AP examinations [Mathews (2004)]. Since the students must attend the course and pass the AP exams to receive the rewards, student who did not take the AP courses would not take the exams in an attempt to earn the cash rewards. This aspect of the incentives makes them relatively hard to game, and likely to increase overall student learning.

AP courses are taught during regular class time, and generally substitute for another course in the same subject (AP Chemistry instead of 11th grade science for example), for another elective course, or a free period. While AP courses count towards a student's high-school GPA, they are above and beyond what is required for high-school graduation. As a rule, an AP course substitutes for some activity that is less demanding.[15] In addition to the AP courses taught at school, there may be extra time dedicated to AP training. For example, the APIP in Dallas includes special "prep sessions" for students, where up to 800 students gather at a single high-school to take seminars from AP teachers as they prepare for their AP exams [Hudgins (2003)].

Adoption of the APIP works as follows. When a private donor approaches AP Strategies, he or she is asked to select a school district in which they would like to fund a program. The total

---

[14] One AP English teacher in Dallas had 6 students out of 11 score a 3 or higher on the AP examination in 1995, the year before the APIP was adopted. In 2003, she earned an extra \$7,350 when 49 of her 110 students received a 3 or above on the AP exams. In addition, she received a \$500 bonus for improving on her previous year's results, \$1,200 for participating in a training program, and another \$2,500 for extra time she spent tutoring. In sum, she earned \$11,550 for participating in the program; a substantial increase in annual earnings. **[**Mathews (2004).]

[15]**Source:** Executive Vice President AP Strategies, and counselor s at several Dallas high-schools. For a detailed description of the high-school graduation requirements see Appendix notes A2.

cost of the program ranges between $100,000 and $200,000 per school per year, depending on the size of the school and its students' propensity to take AP courses. The average cost per student in an AP class ranges between $100 to $300. Private donors pay for between 60% to 75% of the total costs of the program, and the district covers the remainder. Districts usually pay for teacher training and corresponding travel and lodging, teacher release time, and some of the supplies and equipment costs. The donors fund the bonuses to students and teachers that are associated with passing AP scores, stipends to teachers for attending team meetings, bonuses to teachers and administrators for passing AP scores, and some of the supplies and equipment costs. Today, the districts may be able to fund its contribution to the APIP using earmarked funds from the statewide AP incentive program and No Child Left Behind. However, in the first years of the program such funds were not available.

The donors choose the subjects that will be rewarded, and ultimately determine the size of the financial rewards. While there are some differences across schools, in most schools, English, mathematics and sciences are rewarded. Taking the donors preferences as given, AP Strategies selects schools, based on the school's willingness and ability to participate. There is variation in the timing of the introduction of the program in certain schools that I exploit in order to identify the effect of the program. As illustrated in Figure 2, in total there are 41 schools that have adopted the APIP to date and 61 schools that will have adopted the program by 2008, so the number of treated units is relatively small. The exact timing of which schools got the program is essentially random. Donor availability and donor preferences are the primary reason for variation in program implementation. To quote the Executive Vice President of AP Strategies, "Many districts are interested in the program but there are no donors. So there is always a shortage of donors". In most cases the donor wants a specific district.[16] Where there are several districts that are competing for the same donor, the donor's preference determines the district or the schools within the district that will adopt the program. Once a willing group of schools has been accepted by the donor the program is implemented the following calendar year. The seven schools to

[16] To give some examples: ST Microelectronics is located in the Carrolton-Farmers community and funded this districts schools; Amarillo and Pflugerville are both funded by anonymous donors who requested those specific districts; Houston, and Austin programs are funded primarily by the Michael Dell foundation who wanted to fund districts with a high proportion of minority students. The first ten Dallas schools were chosen based on proximity to AP strategies. The next few Dallas schools were funded by LAMPS because the schools met the criteria of the types of schools they wanted to support. The last few Dallas schools were funded by the O'Donell foundation to complete the funding of Dallas ISD. Burkburnett and City View programs were funded by The Priddy Foundation who wanted to fund schools in the Wichita Falls area.

adopt the APIP in 2008, however, decided to have the pre-AP preparation portion of the program in place for at least a year before the incentive scheme was implemented.

III. Data

The data on school demographics, high-school graduation rates, and college entrance examinations are from the publicly available Academic Educational Indicator System (AEIS) on the Texas Education Agency website. These high-school level data span the years 1994 through 2004. Urbanicity data come from the 2002 Common Core Data. College enrollment data come from the Texas Higher Education Coordinating Board website and are available for the years 2002 through 2005. The final dataset combines these publicly available data with a listing of program schools by year obtained from AP Strategies. The total number of schools in the sample to have adopted the program is 40, while 59 are scheduled to have adopted the program by 2008.

The summary statistics in Tables 1 and 2 are broken up by the eligible group, and the non-eligible group. Eligible schools are schools that will have adopted the program by 2008. Restricting to schools for which there are SAT/ACT data reduces the sample to 1,438 schools from the universe of all public schools in Texas. The unit of observation is a school year, therefore a two-year period would include two observations per school. Some variables are not available for some years, and some schools do not exist during all years, thus sample sizes and composition may vary. Standard errors of the sample means are in parenthesis. The demographic variables are summarized in Table 1.

Schools that were selected for the APIP look quite different from schools that have not yet been selected and may never be selected for the APIP. The APIP schools have much larger black and Hispanic enrollment shares, and lower white enrollment shares. The APIP schools are also much larger than non-eligible schools as evidenced by the larger 10^{th} 11^{th} and 12^{th} grade enrollments. The APIP schools have relatively more Limited English Proficient (LEP) students and students who are classified as economically disadvantaged. About 90% of the schools selected for the APIP are located in a large or medium sized city compared to less than 40% of the non-selected schools. In 2002, two new rural schools adopted the APIP. As one looks across the columns in Table 1, one notices that the schools selected for the APIP have a larger increase in Hispanic enrollment shares and a larger decrease in black enrollment shares than the non-selected schools. The APIP schools have similar proportions of economically disadvantaged students in the 1993-95 period, but by the 2003-2005 period, the APIP schools have much higher economically disadvantaged enrollment shares than other schools in Texas.

The outcome variables summarized in Table 2 are: the number of high-school graduates, the proportion of graduates taking college-entrance examinations (SAT or ACT), the number of students scoring above 1100 in the SAT or 24 on the ACT examinations (scoring above criterion)[17], and the number of students who are enrolled in college in Texas the fall following their graduation year (2002 through 2005 only). The proportion of 11th and 12th graders who take an AP or International Baccalaureate (IB) examination, and the percentage of graduates who score above criterion are broken up by student ethnicity. IB courses are college level courses taken in high-school that are comparable to AP courses. The percentage of 9th through 12th graders who take a dual enrollment course (a college course taken while in high-school) and the percentage of high-school graduates who meet equivalency standards on the exit level TAAS/TASP are also included. A student who achieves TAAS/TASP equivalency has done well enough on the exit-level TAAS to have a 75% likelihood of passing the Texas Academic Skills Program (TASP) test, now called the Texas Higher Education Assessment (THEA).[18]

Looking at the ratio of graduates to tenth graders, one sees that the share of students who do not graduate at that high-school is much higher in APIP schools than the control schools. One also notices, by looking at the college numbers in Table 2 and the graduate numbers in Table 1, that the ratio of college attendees to graduates is 62% for the non-APIP schools and only 42% for the APIP schools. The proportion of 12th graders who score above criterion on the SAT/ACT is roughly 24% for the control schools and only about 11% for the APIP schools. The APIP schools, however, do have higher rates of AP/IB participation and higher dual enrollment rates. This difference is due largely to the fact that rural schools, which make up almost half of the control sample tend to have much fewer AP and IB course offerings than urban schools. The percentage of white, black and Hispanic non-special education graduates who score above criterion is also included in Table 2.[19] These variables have the number of non-special education graduates in the

[17] Readers should note that the SAT examination was re-centered in 1995 year such that above-criterion variables are unaffected in the sample period.
[18] Passing the THEA and the TASP is a bare minimum requirement for entry into any public higher education institution in Texas. See http://www.thea.nesinc.com/.
[19] This is the form in which the data are available. The numbers of non-special education graduates is available for the campus but are not publicly available broken down by ethnicity. As such, the number of students scoring above criterion for the campus is available while the same statistics are not available for students of different ethnicities.

denominator, and as such are based on a self-selected sample. For this reason, it would be misguided to compare the values of these variables for the APIP schools to those of the other high-schools since the other high-schools have a larger share of 10th graders who graduate. There is no strong discernable difference in the changes in the outcomes over time between the APIP schools and the other Texas high-schools.

IV. Theoretical Background

This section analyzes the APIP through the lens of two theoretical models; the multitask principal agent problem and the Becker-Rosen schooling decision model. These models enable us to predict the effect the APIP would have on the outcomes in question.

A The Education Production Function

One can anchor the teacher's problem within a principal agent framework. $Y_{AP} = f_1(e_{s1}, e_{t1})$ is number of passing AP exams. $Y_{other} = f_2(e_{s2}, e_{t2})$ is student output in other areas. Overall student output is a function of AP knowledge and other knowledge such that $Y_{total} = g(Y_{AP}, Y_{other})$. Teacher effort in APs and other tasks is e_{t1} and e_{t2} respectively. Student effort in APs and other tasks is e_{s1} and e_{s2} respectively. Teachers pay a cost $\tau(e_{t1})$ to exert effort in APs and cost $\tau(e_{t2})$ to exert effort in other areas. Students pay a cost $\tau(e_{s1})$ to exert effort in APs and cost $\tau(e_{s2})$ to exert effort in other areas. Both Y_{AP} and Y_{other} are increasing in their arguments. School administrators would like to maximize Y_{total}, while teachers maximize their utility $U_t = h_t(Y_{total}) - \tau(e_{t1}) - \tau(e_{t2})$ where $h_{Y_{total}} \geq 0$. Teachers take student effort as given.

With the introduction of the APIP, teacher pay is more closely tied to the AP output of their students. The principal agent multitask model predicts that where good AP performance is more likely with higher teacher effort, teachers will exert more effort to improve students' AP output [Holmstrom and Milgrom (1991)]. If the increased teacher effort on AP output comes at the expense of teacher effort on other tasks, then Y_{other} will decrease. Whether the increase in Y_{AP} and the possible resulting decrease in Y_{other} leads to an overall increase or reduction in Y_{total} depends on the complementarities between the two types of knowledge. Thus the model predicts that the APIP will lead teachers to improve AP performance, while the prediction on teachers' impact on other outcomes is ambiguous.

The gains to a student of taking APs are unambiguously greater under the APIP and therefore student AP effort and AP participation will increase at the margin under the APIP. Students gain some knowledge and skills through taking AP courses and potentially better AP instruction, which would increase AP output Y_{AP}. If student effort in other areas is unchanged, total output Y_{total} will also increase. Consider SAT/ACT performance a measure of Y_{total}. All else equal, those students who have taken AP courses as a result of the program will score at least as high or higher on the SAT/ACT. As such, *ceteris paribus*, one would expect the number of students scoring above criterion on the SAT/ACT to increase as a result of the APIP. The above discussion has assumed that student effort in other tasks e_{s2} is unchanged after APIP adoption. Using exactly the same argument as teachers, the increased student effort in APs could lead to a decrease in effort on other tasks, and lead to lower overall output Y_{total}. Whether effort on other tasks increase or decreases, and whether total output increases or decreases, depends on the complementarities between Y_{AP} and Y_{other}. The *a priori* predictions of the direction of the effort effect is unclear. The effect of the APIP on college going decisions is best understood with a schooling decision model.

B The Schooling Decision Model

Consider a simplified Becker-Rosen model of schooling in which the log of earnings are an increasing concave function of the years of schooling $y = e^{g(s)}$, where $g_s \geq 0$ and $g_{ss} \leq 0$. Individuals pay a cost associated with attending school c_i (effort costs, and tuition costs), and discount the future at rate δ. Individuals chose the level of schooling s to maximize the present value of earning minus the present value of costs. As a further simplification, students choose between going to college for 4 years ($s=4$) or stopping schooling after high-school graduation ($s=0$). An individual chooses to attend college *iff*

$$V(4) \geq V(0) \equiv \int_4^\infty e^{g(4)} e^{-\delta t} dt - \int_0^4 c_i e^{-\delta t} dt \geq \int_0^\infty e^{g(0)} e^{-\delta t} dt \qquad [1]$$

The model illustrates that as c_i decreases, the individual's utility associated with attending college increases, so that a reduction in the costs associated with college will make individuals more likely to attend college as opposed to ending their schooling at high-school.

It is useful to think of taking AP courses as a way to reduce the costs to attending college (increased likelihood of admission, more financial aid, tuition savings due to college credit, and faster graduation). Because the increased AP participation will reduce individuals' costs of attending college, for any given level of applications, one would expect an increase in matriculation rates. The second way in which APIP may affect c_i is if students don't know their true ability to succeed at college i.e., they don't know their true value of c_i. Learning their true cost c_i as a result of taking AP courses could affect their college enrollment decisions. All else equal, if they learn that they are able to cope and enjoy the material, then they may be more likely to apply to college and vice-versa. The direction and magnitude of this information effect depends largely on three things: (1) how many students are on the margin of applying to college, (2) whether students on the margin are optimistic or pessimistic about their abilities and (3) the quality of AP instruction.

Consider the student's decision to take AP courses. Taking AP courses reduces her college costs from c_0 to c_1. Taking AP courses entails a cost $\tau(e_{s1})$. Assuming that students do not change their college going behaviors under the two regimes, a student only takes the AP course if the present discounted value of the college cost savings is greater than the cost to the student of taking AP courses. Specifically *iff*

$$\tau(e_{s1}) \leq \int_0^s (c_0 - c_1)e^{-\delta t}dt \text{ where } s \in \{0,4\}. \qquad [2]$$

This inequality illustrates that (1) students who are not going to college will have no cost reduction and will therefore not take AP courses, and (2) those that will go to college may take AP courses to reduce their costs. In fact, the college going decision and the AP taking decisions are made jointly, but the partial effects models capture the dynamics of the choice problem. Under the APIP, students earn money for passing AP exams, only after they have taken the AP courses. Thus students who take the AP exams will get a reward π at time 1. Under the APIP students will take the AP exams if the present value of the costs minus the reward is less than the savings in college costs or *iff*

$$\tau(e_{s1}) \leq e^{-\delta}\pi + \int_0^s (c_0 - c_1)e^{-\delta t}dt \text{ where } s \in \{0,4\}. \qquad [3]$$

The savings associated with just passing out of a semester of college, and the recouped labor force earnings is about a third of ones annual salary plus half of a years tuition costs[20]. If one were to add the increase in earnings potential associated with being more likely to go to college and getting into a more selective college, the benefits to AP taking would be even greater.

[20] According to the US Department of Education, average annual tuition costs in Texas are $8,057. http://www.ed.gov/students/prep/college/thinkcollege/early/students/edlite-college-costs.html
According to the US census bureau in 2005 workers 18 and over with a bachelor's degree earn an average of $51,206 a year. http://www.census.gov/Press-Release/www/releases/archives/education/004214.html
Half the tuition plus one third of the annual earning comes to $21,093.

In contrast, the average value of the rewards is less than $200 per student and the test fee reduction is at most $30.[21] Based on comparing [2] and [3], in order for the rewards to have the large impact on AP participation documented in Figure 1 would require: (A) the unlikely condition that several students were bunched right on the margin at which the benefits outweigh the costs and they were bumped over this margin by the prospect of the rewards, (B) the implausible scenario that students who are not interested in going to college are taking AP courses solely for the rewards, (C) students are myopic and value the rewards in the near future over the potential savings further in the future, (D) students were uninformed of the benefits to AP taking, (E) students were otherwise constrained from taking AP courses, or (F) students were more likely to take AP courses under the APIP for reasons unrelated to their present discounted value of income stream, such as a change in peer attitudes toward AP courses, or greater teacher encouragement. Given the relative magnitude of the potential gains to AP taking and the monetary rewards a student could earn, the large response suggests some inefficiency was ameliorated by the APIP.

V. Identification Strategy

Since program schools were not randomly selected, a critical question is which schools to use as an appropriate control group. Given that APIP schools may have been very different from non-APIP schools *before* APIP adoption, comparing the outcomes of APIP adopters to the outcomes of other schools would confound the effects of the program with the pre-adoption differences between adopters and non-adopters. As such, my basic strategy is to compare the outcomes for cohorts at APIP schools before APIP adoption to the outcomes of other cohorts from the same school after APIP adoption. By comparing cohorts within the same school, I am able to eliminate self-selection within a cohort. To account for policy or demographic changes that may affect all schools over time and other attributes that may change within a school over time, I compare the changes in outcomes associated with APIP adoption to the changes in outcomes for the schools that do not adopt the APIP over the same time period, while controlling for other observable school characteristics. This difference-in-differences (DID) strategy described above is implemented by estimating equation [4] by OLS.

[21] For students who are on free or reduced lunch the reduction is from $30 dollars to $15 dollars. The test fee reduction of $15 would be important if students were very credit constrained. While this is possible, the $15 reduction is much too small to generate the participation responses observed.

$$Y_{it} = \alpha + \beta \cdot Treat_{it} + X'_{it}\psi + \tau_t + \theta_i + \varepsilon_{it} \quad [4]$$

Where Y_{it} is the outcome for school *i* in year *t*, $Treat_{it}$ is an indicator variable that takes the value of 1 if the school has adopted the APIP by that particular year and 0 otherwise, X_{it} is a matrix of time varying school enrollment and demographic variables, τ_t is a year fixed effect, θ_i is a school-specific intercept and ε_{it} is an idiosyncratic error term. Since APIP schools may have already been on a trajectory of improving outcomes, or progressively getting worse, one would like to compare the changes in outcomes, for the treated schools to the changes in outcomes for other schools that were not treated but on a similar trajectory. I address this concern by estimating equation [4] while including a different linear time trend for the APIP eligible schools.[22] This strategy is valid so long as APIP schools have the same incoming distribution of students in their pre-program and post-program cohorts. I will present discussions of and tests of this restriction.

DID with school fixed effects on pre-selected sample of similar schools

Using all other schools as controls would only be valid if (a) all other Texas high-schools shared the same time shocks as APIP schools and (b) outcomes in APIP schools and other schools respond similarly to demographic and enrollment changes over time. Given the differences between APIP and non-APIP schools documented in Tables 1 and 2, this assumption is implausible.[23] While accounting for such heterogeneity can potentially be achieved by including flexible interactions of the covariates, a more elegant and robust solution is to restrict the control sample to schools that are similar to the treatment schools. The selection bias literature primarily uses two methods to find untreated units that are observationally similar to treated units. These methods rely on the assumption that treatment is exogenous between units that are observationally similar. One method is to match treated and control units based on similarity in their observable characteristics. The second method is to estimate each unit's

[22] The resulting baseline regression would be of the form $Y_{it} = \alpha + \beta \cdot Treat_{it} + \eta I_i^{APIP} \cdot t + X'_{it}\psi + \tau_t + \theta_i + \varepsilon_{it}$ All variables are defined as before and I_i^{APIP} is a school-constant indicator variable equal to 1 if school *i* is ever an APIP school and 0 otherwise. The variable *t* is the calendar year and η measures a differential linear trend for the APIP schools. A positive η would indicate that the APIP schools tend to have greater improvements in outcomes over time than other schools irrespective of adoption status, while a negative η would indicate the opposite.

[23] For example an inflow of low-income students to a suburban school may have a very different effect on outcomes than an inflow of low-income students to an inner-city school where most treated schools are located. In the presence of unobserved heterogeneity this is likely to lead to incorrect inferences and inconsistent estimates. Assuming the marginal effect to be the same across all schools will lead to a form of specification error that would lead to inconsistent estimates of the effect of APIP adoption.

likelihood of receiving treatment as a function of observable characteristics (propensity score) and match treated and untreated units with similar estimated propensity scores.

Using the first method, I select the sample of schools that were observationally similar to APIP schools during their pre-adoption years. As a measure of observational similarity, the distance between two schools *i* and *j* is $d_{i,j} = [(X_{it} - X_{jt})'V(X_{it} - X_{jt})]^{1/2}$ where *V* is the diagonal matrix constructed by putting the inverses of the variances of the covariates on the diagonal.[24] The vector X_{it} includes all school urbanicity and demographic variables for school *i* at time *t*. For each treated school (school that actually adopted the program), for each pre-adoption year, I select the seven other schools that did not adopt the program with the smallest distance from the treated school.[25] The control sample is comprised of all untreated schools that meet the distance criteria (similar covariates) for at least one APIP school during a pre-adoption year. The summary statistics of the treated schools and the control schools in Tables 3 and 4 show that the matched sample is observationally similar to the treated group of schools.

Using the second method, I estimate the propensity score using a probit model of treatment (APIP adoption) as a function of all the school demographic control variables and the first and second lags of the outcome variables. This would capture the fact that treatment may be determined not only by school demographics but also based on historical performance and trends in the outcome variables. The propensity score estimates are shown in Table 5. The estimated propensities vary for each school over time, since the covariates vary by year. I define the maximum estimated propensity score, across the years, to be the propensity score for that school. The distribution of the maximum estimated propensity scores are shown in Appendix Figure A1. A commonsense approach to restricting the sample would be to only use those control schools with estimated propensities that are within the range of estimated propensities of treated schools. Given the wide range of estimated propensities for the APIP schools, attributable to multiple selection criteria, this yields a very large sample (excluding only the poorest of matches). As such, results using this sample are not reported in the main text. However, Appendix note A3 describes the sample selection, its shortcomings and the estimated results using this sample.

[24] This is the efficient weight matrix suggested in Abadie, Drukker, Herr and Imbens (2001).

[25] Because this method matches each treated school in each year to seven closest untreated matches, an untreated school may be matched with several treated high-schools at any point in time, and could be matched with the same treated schools for multiple years. As such there is some overlap in the control schools used over treated schools.

The matched control groups and estimated propensities can also be used in the regression model to determine the counterfactual time trends or time effects treated schools would have experienced in the absence of the APIP. The underlying assumption is that schools with similar observable characteristics or similar estimated propensities would experience similar time trends and experience similar time shocks. Allowing for such heterogeneity in time trends or time-effects increases the efficiency of the estimates and reduces the likelihood of comparing the change in outcomes across schools that were on different trajectories. For regressions on the matched sample, I estimate a linear time trend for each matched group of schools to determine the counterfactual time trend for the treated school of that group (there are only 35 treated schools with pre-adoption data). The resulting model is equation [5].

$$Y_{it} = \alpha + \beta \cdot Treat_{it} + X'_{it}\psi + \sum_{g=1}^{35} \delta_g G_{ig} \cdot t + \theta_i + \varepsilon_{it} \qquad [5]$$

All variables are defined as before G_{ig} is a dummy variable denoting the school group (equal to 1 if the school is a control unit for treated school g or if it is school g and zero otherwise), and t is the year. Parameter δ_g measures a linear time trend for group g. In all other samples, I estimate individual year-effects for schools with different estimated propensities. I create ten propensity score decile bins and estimate a year effect for each bin (decile). The resulting model is equation [6].

$$Y_{it} = \alpha + \beta \cdot Treat_{it} + X'_{it}\psi + \sum_{q} \tau_{qt} \cdot Q_{iq} + \theta_i + \varepsilon_{it} \qquad [6]$$

The variables are defined as before but now Q_{iq} is a dummy variable denoting propensity score deciles, so that τ_{qt} is the year effect for schools in decile q. With the inclusion of the propensity score decile-year fixed effects there is no need to estimate a linear trend for the APIP schools.[26]

DID with school fixed effects on eligible sample only

While schools that are observationally similar to the treated APIP schools are likely to be good controls, there is always the concern that schools that adopt the program are observationally similar but differ in unobservable ways from those that do not. The final approach to restricting the sample which directly addresses this concern is to only use the schools that received treatment and those that have been selected to receive treatment in the

[26] Specifications that include a linear trend for the APIP schools fail to reject the null hypothesis of no differential time trend at the even the 20 percent level.

future. Given that the timing of implementation is driven by the availability and the preferences of donors, conditional on being chosen to adopt the APIP, the actual timing of implementation is random. Since all the schools that adopt the APIP are interested in having the program, using only this sample ensures that the control schools are of equal unobservable "willingness" or ability to implement the program. This approach also ensures that the issue of response heterogeneity does not bias the results albeit at the expense of smaller sample sizes and less statistical power. Within the sample of APIP schools identification hinges on the random timing of program adoption and relies on the plausible assumption that the schools that are APIP schools but have not yet adopted the program are similar in *unobservable* ways to the APIP schools that have adopted the program. The summary statistics for the schools that adopted the program and the schools that will adopt the program in 2007 or 2008 presented in Tables 6 and 7 show that the future APIP adopters are similar to the schools that adopt the APIP in sample. I estimate models of the form [6] for the sample of APIP schools.

VI. Results

I present the results in four sections. In the main results section, I document the effect of the APIP on the main high-school outcomes: SAT/ACT performance, college matriculation and high-school graduation. I show the importance of selecting an appropriate control group and accounting for differential time effects across treated and control schools. I show that the central findings are robust to any reasonable specification that take these factors into account and to removing the early adopting schools from the sample. I also test for selective migration. The logic of this test is that any increase in the number of students who matriculate in college as a result of selective migration into APIP schools after adoption would lead to an increase in the number of students who (a) graduate from high-school, (b) take the SAT/ACT and (c) achieve TAAS/TASP equivalency. As such, in most plausible scenarios, one can rule out the possibility that increases in the number of students with high SAT/ACT scores, or the number of college entrants is due to an inflow of high-ability students into APIP schools if there is no corresponding increase in the number of students graduating from high-school or taking the SAT/ACT exams.

In the second section I investigate how the APIP affects outcomes over time. I also show the effect of the APIP on AP outcomes such as AP course enrollment and AP/IB examination taking. In the third section I show the effect of the APIP on the AP/IB examination taking and SAT/ACT performance of student broken up by gender and ethnicity. Using data from a sample of students who did not change schools during their entire high-school career, I show that the effects of the APIP on SAT performance cannot be attributed solely to any selective-migration- i.e. an inflow of high-achieving student into APIP schools. Finally, in the fourth section, I present anecdotal evidence from discussions with guidance counselor s for why and how the APIP affects student outcomes. I then present the results of various tests of the hypothesis that students and teachers are responding to the incentives in manners that are consistent with revenue maximizing.

Main Results

The main results are summarized in Table 8. The variable of interest is the “treat” variable which indicates APIP adoption. Since all the main outcomes are in logs, the interpretation of the coefficient on treat is the percentage change in the outcomes associated with APIP adoption. The main results include school and year fixed effects, using the different samples and control methods. To ensure that when one finds effects of the APIP on some outcomes while not on others, that it is not the result of sample differences, all regressions are estimated on a balanced panel of schools based on the availability of SAT/ACT data (with the exception of the college enrollment). The sample sizes for college enrollment are much smaller than for the other variables, thus the college outcomes have considerably less power. All the outcomes with the exception of grade 12 enrollment are conditional on the natural log of grade 12 enrollment.

The top panel of Table 8 shows the effect of the APIP using all schools in Texas while not controlling for any difference in underlying time trend. In this naïve specification the APIP is associated with a 4.6 percent increase in the number of high-school graduates and a 9.1 percent increase in the number of students who matriculate in college in Texas. The second-highest panel, which shows the estimated APIP effect while allowing APIP schools to have a different time trend from the other schools shows that APIP adoption is associated with a 7 percent increase in the number of 12^{th} graders and a 16% increase in the number of students who score above 1100/24 on the SAT/ACT exams. The negative coefficients on the time trends for grade 12 enrolment and above criterion indicate that the APIP schools had declining shares of students with high SAT/ACT scores and declining 12^{th} grade enrollments (relative to non APIP schools).

These changing results underscore the importance of taking the fact that the APIP adopters may have been on a different time trajectory than comparison schools into account.

The third-highest panel shows the first attempt to restrict the control sample to an appropriate set of schools. The third panel shows the estimated effect when one restricts analysis to the matched sample of schools, and one allows each matched group to have its own time trend. Appendix Table A2 includes very similar estimates when using propensity score decile by year fixed effects. In this specification APIP adoption is associated with a 4.8 percent increase in grade 12 enrollment an 11 percent increase in scoring above criterion and a 5.9 percent increase in college matriculation. There is no corresponding statistically significant increase in high-school graduation or SAT/ACT taking.[27] The fourth-highest panel restricts analysis to only the APIP eligible schools and controls for differential time effects by including year fixed effects interacted with deciles of the estimated propensity score. These results are similar to those of the matched sample such that APIP adoption is associated with a 5.7 percent increase in grade 12 enrollment a 22 percent increase in scoring above criterion and a 6.7 percent increase in college matriculation with no corresponding statistically significant increase in high-school graduation or SAT/ACT taking. Given the large gap in time between the first schools to adopt in 1996 and those in 2001, readers may be concerned that the early adopting APIP schools are not a good control group for the schools that adopt the APIP after the year 2001. As a robustness check, the lowest panel shows the results when one removes the first ten schools to adopt the APIP from the analysis, to remove the possibility that the estimated benefits are being driven by an early adopter bias. Using only the late adopters, APIP adoption is associated with a 29 percent increase in scoring above criterion and a 9 percent increase in college matriculation. There is no corresponding statistically significant increase in 12 grade enrollment, high-school graduation or SAT/ACT taking.

The results in the lowest three panels (which take the potential endogeniety concerns into account) are rather similar. They all show improvement in SAT/ACT performance and increases in college matriculation with no effect on high-school graduation or SAT/ACT taking. While the results show no decrease in the graduation rate conditional on grade 12 enrollment, readers may be concerned that the program could have an effect of the graduation rate conditional on grade 10 enrollment since most high school dropout occurs before 12th grade. The program could also

[27] These results are very similar to the propensity score restricted sample results summarized in appendix note A3.

have affected TAAS/TASP equivalency since students who attain TAAS/TASP equivalency would graduate from high-school, but would not be scoring above criterion on the SAT/ACT. Table 9 shows that there was no effect of the APIP on the number of high-school graduates (conditional on twice lagged 10th grade enrollment) and no effect on TAAS/TASP equivalency using the APIP sample. While the APIP may have increased 12th grade enrollment, students at APIP schools were no more likely to graduate from high-school, take the SAT/ACT exams, or achieve TAAP/TASP equivalency than before adoption. As such, the increase in the number of students who score above 1100/24 on the SAT/ACT and the number of students who matriculate in college is very unlikely to be due to an inflow of high-ability students into APIP schools.[28] In fact, using only the late APIP adopters where the estimated benefits to APIP adoption are largest, there is no statistically significant increase in 12th grade enrollment and the coefficients on high-school graduates and taking SAT/ACT are negative. Since the coefficients in this specification have the wrong sign for selective migration to be the cause, one cannot argue that the lack of significance is due to a lack of statistical power.

Dynamic Effects

To capture the dynamic effects of the program, I estimate specifications that include indicator variables for the years since adoption of the program.[29] This is implemented by estimating equation [7] by OLS in the sample of APIP eligible schools.

$$Y_{it} = \alpha + \sum_{k=1}^{3+} \pi_k \delta_{k,it} + X'_{it}\psi + \sum_{q} \tau_{qt} \cdot Q_{iq} + \theta_i + \varepsilon_{it} \qquad [7]$$

All variables are as before, but now π_k is the estimated coefficient on an indicator variable equal to 1 if it is the k^{th} year of the program and 0 otherwise, where $k \in \{1, 2, 3+\}$. The coefficients will map out how the program affects the outcome variable after being in place for one year, two years and then three or more years. With more data one could map out some pre-treatment years and more post-treatment years, however, having only forty treated schools preclude such estimates. It should be noted that comparisons of first year estimates and the third year estimates will not necessarily be on a balanced sample of schools since not all schools are observed in their

[28] It should be noted that with this test, I cannot rule out the possibility that there was an inflow of high-ability students and simultaneous inflow of low-ability students such that in the aggregate they cancel out leading to no effect on these outcomes. Given that high-school graduation, SAT/ACT taking and TAAS/TASP equivalency measure three different margins, this implausible scenario would have to hold in all three cases.

[29] This approach has been used in Grogger (1995), Reber (2003), and Stevenson and Wolfers (2006).

first second and third years of adoption in the sample period.[30] As such, the coefficients on the "years since adoption" variables will reflect the dynamic effects of the program, and may also reflect differences in program response or implementation across schools.

Figure 3 shows the increase in Advanced Placement course enrollment relative to non-AP course enrollment. A student is counted as enrolled in a course if they spent one day in the course. As such, this will include students who may have sat in on an AP course for less than a week and then withdrew. Relative to non-AP course enrollment, AP course enrollment increased by about 6 percent the first year, 12 percent the second year and then by about 21 percent in subsequent years (relative to non-AP enrollment). Table 10 shows the effects of APIP adoption by years since adoption and reports the coefficients on the first year, second year and the third year and beyond. The result in Figure 3 is consistent with the findings in column 6 of Table 10 which shows that AP course enrollment increases by about 21 percent by the third year. While much of the increase in AP course enrollment takes place after two years, there is an immediate 2 percentage point increase in the percentage of students taking at least one AP/IB examination that increases to 4.2 percentage points after the first year. With a base of about 18 percent this represents a relative increase of 11 percent in year one and 23 percent by year two onwards. The fact that course enrollment numbers do not kick in until year three suggests that much of the *initial* increase in AP exam taking came from students who, in the absence of the APIP, would have taken the course but not taken the exam, or would have enrolled in and withdrawn from AP courses. By year three onwards however, it is apparent that much of the increase in AP/IB examination taking is due to an increase in students taking AP courses and taking AP examinations as the pre-AP aspects of the program took effect.

Columns 2 and 3 echo the results in Table 8, where there is no statistically significant effect on high-school graduates or taking the SAT/ACT exams. Column 4 shows a 19 percent increase in the number of students scoring above 1100/24 on the SAT/ACT exams the first year of the program, a 22 percent increase the second year, and a 33 percent increase by the third year of adoption (relative to no adoption). The first year effect is somewhat large in light of the fact that some seniors in the first year may have taken the SAT/ACT after only being exposed to the

[30] For example, schools that adopt the program in 2005 will be used to identify the effect of the first year of the program but will not be used to identify the effect of having the program for two or more years. Since the college matriculation data begin in the year 2002, none of the ten original Dallas schools will be used directly to identify the program effect on college going.

APIP for half the school year. This suggests that the improvements in SAT performance may not solely be due to a spillover from exposure to AP courses, and may reflect an increase in SAT/ACT effort or increased desire to get into a good college. Column 4 shows that there is about a 7 percent increase in college matriculation associated with APIP adoption that does not increase over time. Column 1 shows that APIP adoption is associated with large increases in 12th grade enrollment by year three. Since this is conditional on lagged 11th grade enrolment, this may be an increased student retention effect. The effects of the program in the first and second year are likely due to improvements in AP instruction and the effect of having the incentives. Other aspects of the APIP such as increasing the pipeline by improving pre-AP instruction would affect the outcomes of graduates after three or four years such that the coefficient on "third year and beyond" includes the effects of all aspects of the program. The large first year effects suggest that improvements in AP instruction and monetary incentives alone were able to increase the number of students who take AP/IB exams, improve SAT/ACT performance, and increase the number of students who matriculate in college in Texas.

Effects by Gender and Race

In this section I present results based on sub-samples based on gender and race. It has been established that AP participation of minorities and low-income students tend to be lower than that of middle-class white students at the same high-schools. Insofar as these differences reflect suboptimal student or teacher effort one would expect to see larger increases in AP participation among these groups. The analysis by gender is motivated by a growing literature documenting that females are more responsive to interventions than males[31] and that among adolescent, girls have more self-discipline and delay gratification more than boys [Silverman (2003); Duckworth, Lee and Seligman (2006)].

Due to the nature in which the data are available the outcomes are reported in percentages and cannot be analyzed in logs as for the main results. Table 11 shows the effect of APIP adoption on the percentage of 11th or 12th graders who took at least one Advanced Placement or International Baccalaureate examination. The results show that the campus-wide increases in the number of students in 11th and 12th grade who take AP or IB exams are driven primarily by increased participation for black and Hispanic students. The results do not show any statistically

[31] For example Kling, Ludwig, and Katz (2005); Anderson (2007); Angrist, Lang, and Oreopoulos (2007); Angrist and Lavy (2007).

significant effect of the APIP on the proportion of white students who take at least one AP or IB exam. This does not mean that the number of AP/IB exams did not increase for the white students at APIP schools, but that the number of white students affected was unchanged. It is entirely possible that those white students who took one AP exam now take more AP courses and exams. The results in columns 5 and 6 show that there are increases in AP and IB examination taking for both genders with no greater effect on girls.

Given the differences in the effect of the APIP on the number of 11^{th} and 12^{th} graders who take at least one AP or IB exam across ethnicities, one would expect to see the corresponding differences in the effect of the APIP on SAT/ACT performance. Table 12 shows the effect of the APIP on the percentage of non-special education high-school graduates who score above 1100/24 on the SAT/ACT examinations for the different groups. By the third year of the program there are positive effects of the program for all groups. Given that Hispanics and blacks are typically underrepresented at the top of the graduating class, they have more room for improvement. While the percentage point changes are similar for white black and Hispanic students, the differences in relative impact, however, are sizable. Compared to the base levels this represents about a 12% increase for whites a 50 percent increase for Hispanics and an 80 percent increase for blacks. The fact that the number of white students taking at least one AP/IB exam did not increase suggests that the gains in SAT/ACT performance experienced by white students may be due to their taking more AP courses, increasing their effort in their courses, increases in the quality of AP instruction, or all three.

To ensure that the improvements in SAT/ACT performance were not driven by selective migration, I obtained school aggregate counts of the number of white and Hispanic graduates scoring above 1100/24 on the SAT/ACT for the subset of high-school graduates who were at the same high-school for all four years of their career. Due to heavy data masking these data are not available for black students. Table 13 shows the effect of APIP adoption of the number of white and Hispanic graduates scoring above criterion. Due to data masking there are some missing values in the data that correspond to counts that are between 0 and 4. The findings are robust to assuming values of 0, 2 or 4 for masked data. Making the reasonable assumption that a masked count is equal to 2, the results in Table 13 show that by year three the APIP increases the number of Hispanic and white students scoring above 1100/24 on the SAT/ACT by 18 percent and 26 percent respectively. I also obtained counts for graduates scoring above 900 on the SAT or 19 on the ACT exams. By year three the APIP increases the number of Hispanic and white graduates scoring above 900/19 on the SAT/ACT by 38 percent and 26 percent respectively. There however seems to be no increase in the number of black students scoring above 900 on the SAT or 19 on the ACT. This provides conclusive evidence that the APIP improves SAT/ACT outcomes (at least for White and Hispanic students), and that the improvements are not due to selective migration. This result lends further credibility to the assumption of no selective migration for all outcomes.

Mechanisms and tests of distortions

Given the improvements in SAT/ACT performance and college matriculation one would expect to see that more students were actually exposed to more rigorous course material due to APIP adoption. For this to be the case, it must be that students did not simply divert their effort away from other advanced courses (such as dual enrollment courses) in order to take AP courses. If students and teachers were simply revenue maximizers one would observe a decrease in dual enrollment courses as a result of APIP adoption. In fact, Figure 4 shows no evidence of this kind of distortion, and there is a statistically insignificant *increase* in dual enrollment course taking associated with APIP adoption after the second year. This suggests that students and teachers *did not* game the incentives by simply substituting away from other advanced courses towards AP courses so that there was an overall increase in rigorous course participation.

Evidence from discussions with guidance counselors at three different APIP high-schools in Dallas strongly suggests there was a school-wide campaign to increase participation in AP courses after APIP adoption. At two of the three high-schools an additional guidance counselor was hired to improve the school's ability to identify those students who should be encouraged to take AP courses. At all three high-schools, the guidance counselors were given explicit instructions to identify those students who should be taking AP courses, and to advertise AP courses. A large part of this campaign is information. Guidance counselors and AP teachers sell the AP program to students who are interested in going to college based on the scholarships that one can earn based on AP scores, the tuition one could save by graduating at an accelerated pace, and also on the fact that AP courses will boost one's high-school GPA and make it more likely to be in the top ten percent of the high-school class and get into a good college. There is also evidence that certain barriers to taking AP courses have been removed; at one high-school, there used to be a minimum class rank that a student had to have in order to take AP courses, but after the APIP was adopted any interested student was allowed to take AP courses.

All guidance counselors mention a shift in student and teacher attitudes towards AP courses. AP courses are now considered difficult courses that anyone can take, as opposed to being available only for the very brightest of students. The AP English teacher who had 11 students in 1995 and 110 students in 2003 highlights the difference in participation. Counselors claim that the reasons for the large increases in AP participation had to do with student information, and increased access through teacher encouragement and increased teacher and guidance counselor recommendations. The financial incentives to students and teachers may be responsible for the increased student and teacher effort in AP courses, but these aspects of the program are downplayed by counselors.

The large increases in AP participation are difficult to reconcile with the theoretical framework put forward in section IV without alluding to several of the elements highlighted by guidance counselor s. The theory and the evidence thus far have suggested that students and teachers are not simply behaving like revenue maximizers. While the data are limited in scope, differences in the way the APIP was implemented across schools allow for certain hypotheses regarding incentives to be tested. The first test is based on the hypothesis that if students and teachers respond to a reward of 100 dollars, then they should be even more responsive to a reward of 200 or 500 dollars. This is likely to be true because (1) larger incentives may lead students and teachers to exert more effort, and (2) a larger incentive means that more students are likely to be on the margin where the incentives make them change their behaviors. Fortunately there is some variation in incentives paid that allow one to see if the effect of the APIP varies by the level of incentives paid. Given that the incentive levels are not exogenously determined[32] and that the sample of schools within each incentive level is relatively small, these differences could reflect differences in implementation and response to the program. As such these findings should not be taken as conclusive but regarded as part of a larger body of evidence.

The evidence supporting the notion that that the effects of the program are stronger in schools with higher incentives presented in Table 14 is mixed at best. In column 5, the college enrollment effect appears to only exist for those schools that pay \$500 per exam, with no effect for schools that pay \$100 per exam. Due to small sample sizes, a coefficient for the 101-499

[32] The levels of incentives are often higher in schools in which the expected propensity to take AP course was particularly low. Source: Executive Vice President of AP Strategies.

range could not be estimated. In column 6, it would appear that the APIP effect has a stronger effect on AP enrollment in schools that pay more than $100 than those that pay only $100 per exam. However, the effect is not monotonic, and the estimated effect on AP course enrolment is larger in schools that pay between 100 and 500 dollars than schools that pay 500 dollars per passing score. In column 7, the APIP effect appears to have a *larger* effect on the percentage of 11th and 12th graders taking AP or IB exams in schools that pay only $100 per exam than schools that pay more. It should be noted that the results in columns 5 through 7 are based on much smaller samples than the other outcomes, so these findings should be taken as suggestive rather than conclusive. The most reliable estimates are for scoring above criterion in column 4; where the sample is large enough to estimate coefficients for all three groups of schools. In column 4 the APIP effect on the number of students scoring above 1100/24 on the SAT/ACT appears to be the same across all incentive levels. In fact, the estimated effect is *smallest* in schools that pay the highest incentives. In sum, the results of Table 14 fail to support the hypothesis that the APIP effect is increasing in the size of the rewards, and suggest that the APIP effect may in fact be invariant to the level of incentives.

The second test is based on the hypothesis that if students and teachers were responding *solely* to the rewards, there should be a greater participation response in those subjects for which rewards are provided than other subjects for which there is no reward. In fact, one might expect a decline in course enrollments of AP courses for which rewards are not provided. There is some variation in the subjects that are rewarded that allow one to take this prediction to the data. There are a few schools that offer rewards for all AP subjects while most only offer rewards for Math Science and English. Figure 5 shows the number of students enrolled in an AP social science courses and Math Science and English courses for Tyler district (which adopted the APIP in 2002 and paid rewards for Math Science and English only) and for Pflugerville district (which adopted the APIP in 2003 and paid rewards for all AP subjects). As one can see, there is an increase in social studies enrollment in Tyler after APIP adoption, suggesting that providing rewards for math science and English did not reduce, but may have *increased* enrollment in social studies AP courses. In Pflugerville, there appears to be an increase in AP enrollment for all subjects. In a regression framework, I test whether the change in the ratio of social studies or humanities AP course enrollees to total AP course enrollees is larger in APIP adopting schools that pay rewards for social studies and humanities. If the APIP increased enrollment in all courses equally, then there should be no differential effect of the APIP in schools that pay rewards for all AP subject as opposed to paying only for math science and English. Even if there were no increase in total AP course enrollment, there may have been a substitution across subjects that would be captured by changes in the ratio of social studies and humanities AP course enrollees to total AP enrollees. The results in columns 1 and 2 of Table 15 show that while APIP adoption appears to have a statistically insignificant negative effect on the social studies or humanities to total AP enrolment ratio those schools that paid rewards for social studies had an even larger decrease in this ratio compared to those that only paid rewards for math, science and English. Insofar as paying rewards for math science and English affects social science and humanities enrollment, the point estimate suggests that it *increases* it, as one could glean from Figure 5. The result is inconsistent with the hypothesis that student and teachers would substitute away from AP course for which there are no rewards towards those for which there are rewards.

VII. Conclusions

The APIP is associated with increases in the number of high-school graduates who score above 1100/24 on the SAT/ACT examinations by about 30%, and the number of students who matriculate in college by about 8% after two years of adopting the program. These findings are robust across a variety of specifications that account for both observed and unobserved heterogeneity across treated and control schools - supporting a causal interpretation of the results. I show that the improvements in SAT/ACT performance are among students who did not change schools during their high-school career and I present evidence that the improvements in college matriculation are not driven by selective migration into APIP schools. The results show improvements in SAT and ACT performance across all ethnic groups and for both male and female students.

The improvements in SAT/ACT performance are likely the result of increased exposure to rigorous material. However, this could also be the result of increased effort in SAT/ACT if effort on APs and effort on other tasks are complementary or if students study harder for both AP and SAT to get into a good college. The fact that there was no increase on SAT/ACT taking suggests that the APIP may not affect students' college application decisions, and that the increased college matriculation was the result of the lower effective costs.[33] However, it is possible that there is an effect on college application behavior that is not picked up by SAT/ACT taking. The theoretical possibility that students and teachers would divert resources away from other tasks towards AP courses is not supported by the data. The APIP has no effect on the number of 12th graders who graduate from high-school, the number of 10th graders who graduate from high-school, or the proportion of high-school graduates who attain TAAS/TASP equivalency. This does not preclude the possibility that the APIP had ill-effects on other *unobserved* outcomes not captured in the data. The fact that there are some benefits with no measured ill effects suggests that, prior to adoption, the selection into AP courses may have been sub-optimal, so that marginal students who may have benefited from taking AP courses were denied from doing so.

The curricular changes and the early emphasis on pre-AP material would not affect the graduating seniors until a few years after the program had been adopted. As such, the changes that take place at year-one of the APIP are likely due to the incentives, the AP courses and improvements in AP instruction. I cannot rule out the possibility that there was an influx of quality teachers to the APIP schools exactly during the fist year of the APIP program. This would not downplay the success of the program, but would suggest that improvements in teacher inputs were a part of the story. The fact that there is an increase in program effect over time on AP course enrolment and SAT/ACT performance, suggests that the push to promote AP participation in early grades through emphasis on pre-AP courses and vertical teams may have been effective.

The improvements in AP instruction would have had little effect if there were not a concurrent increase in the number of students taking AP courses. The anecdotal evidence suggests that the APIP gave teachers the impetus to increase AP course enrollment, guidance counselors the incentive to advertise and inform students of the benefits to the AP program, and students the incentives to take them. Guidance counselors claim that the alignment of school, student, and teacher incentives had a strong effect on the culture and attitudes of both students and educators, which in turn led to improved student outcomes. The empirical tests suggest that the APIP was working through some mechanism other than students and teachers reacting

[33] Since low-income students are sensitive to tuition costs, the potential tuition savings associated with an increase in financial aid and the ability to earn college credit for passing AP scores would lead to an increase in matriculation even if there were no increased interest in college.

directly to the monetary incentives in a "carrot and stick" manner. The body of evidence is more consistent with explanations put forth by guidance counselor s such as changes in peer norms, teacher norms, increased emphasis on AP courses, and information on the benefits to taking AP courses. The findings are suggestive of some of reasons we observe suboptimal educational choices in low-income, low-performing schools. The fact that the AP/IB exam participation response was much larger (on the extensive margin) for black and Hispanic students suggests that they had low initial participation rates because (a) peer norms did not promote taking AP courses, (b) students from these populations were less likely to have good information on how to negotiate the college application process, (c) student expectations of likelihood of success may have been low due to sub-optimal teacher encouragement.

While I show that the program is likely to have lasting effects on students, since they are more likely to attend college, it would be useful to determine the long-term effects of the APIP by observing the students affected by the APIP when they go to college and into the labor force. The program costs between one and three hundred dollars per student who takes an AP exam. If this program increases a student's likelihood of attending college, increases the quality of college attended, and reduces the time it takes to graduate from college, the costs of the program on a per-student basis would be far less than the average increase in lifetime earnings. In addition to the private costs associated with having students attend college who are not college ready, Texas currently spends 80 million dollars per year to bring ill-prepared college students up the level at which they can cope with college level course material. Since the program could potentially reduce the demand for remedial courses while in college, there could be cost savings for this reason also. As such, the relatively small per pupil expenditure on the APIP may have high social returns due to both the sizable private returns for students and perhaps some cost savings for the local government.

References

Abadie, A and G. Imbens, (2002) “Simple and Bias-Corrected Matching Estimators for Average Treatment Effects” National Bureau of Economic Reserach Technical Working Paper 283

Abadie, A., D. Drukker, J. Herr, and G. Imbens. (2001), “Implementing Matching Estimators for Average Treatment Effects in Stata”, *The Stata Journal*, volume1, Number 1, pp. 1-18.

Adelman, C.,: “Answers in the Tool Box: Academic Intensity, Attendance Patterns and Bachelor's Degree Attainment” (1999), U.S. Department of Education.

Anderson, M., (September 2005) “Uncovering Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” MIT Department of Economics, mimeo.

Angrist, J., E. Bettinger, E. Bloom, E. King and M. Kremer. (2002), “Vouchers for Private Schooling in Colombia: Evidence from Randomized Natural Experiments,” *American Economic Review*, 92(5), pp. 1535-1558.

Angrist, J., D. Lang, and P. Oreopoulos. (October 2007) “Incentives and Services for College Achievement: Evidence from a Randomized Trial” IZA Discussion Paper No. 3134.

Angrist, J., and V. Lavy, (2002) “The effect of High School Matriculation Awards: Evidence from Randomized Trials”, National Bureau of Economic Research Working Paper 9389.

Angrist, J., and V. Lavy, (2007) “The Effects of High Stakes High School Achievement Awards: Evidence from a Group Randomized Trial”, MIT and Hebrew University of Jerusalem.

Angrist, J., and V. Lavy, (2004) “The Effect of High Stakes High School Achievement Awards: Evidence from a School-Centered Randomized Trial”, IZA Discussion Papers Number 1146.

AP Strategies (Oct 11 2006) “Creating partnerships for academic Excellence” http://www.apstrategies.org/download/apresults/APS%202006.pdf.

Atkinson, A., and S. Burgess, C. Bronwyn, and G. Paul, (2004) "Evaluating the Impact of Performance-related Pay for Teachers in England," The Centre for Market and Public Organization 04/113, Department of Economics, University of Bristol, UK.

Breneman, D., and W. Haarlow. (1997) “Remedial Education: Costs and Consequences.” Remediation in Higher Education. Washington, D.C.: Thomas B. Fordham Foundation.

Binder, M., P. Ganderton, and K. Hutchens, (2002) “Incentive effects of New Mexico's meritbased state scholarship program: Who responds and how?” In D. E. Heller and P. Marin (Eds.), “Who should we help? The negative social consequences of merit scholarships” (pp. 41-56). Cambridge, MA: The Civil Rights Project at Harvard University.

Card, D., (1995), “Earnings, Schooling, and Ability Revisited,” in Solomon Polachek, ed., *Research in Labor Economics*, 14.

College Board (August 2003) “SAT Verbal and Math Scores Up Significantly as a Record-breaking Number of Students Take the Test. Average Math Score at Highest Level in More than 35 Years” Press release 26858.

Cornwell, C., and K. Lee, and D. Mustard, (2003) "The Effects of Merit-Based Financial Aid on Course Enrollment, Withdrawal and Completion in College," IZA Discussion Papers 820, Institute for the Study of Labor (IZA)

Dodd, B., S. Fitzpatrick, R. De Ayala, and J. Jennings, (2002) “An Investigation of the Validity of AP Grades of 3 and a Comparison of AP and Non-AP Student Groups” College Entrance Examination Board, New York.

Duckworth, A., and M. Seligman, (2006) “Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Test Scores.” *Journal of Educational Psychology* 98,

198-208.
Dougherty, Chrys, Mellor, Lynn and Jian, Shuling (February 2006) "The Relationship between Advanced Placement and College Graduation." National Centre for Education Assessment 2005 AP Study Series, Report 1.
Eimers, M. (2003) "Dual Credit and Advanced Placement: Do They Help Prepare Students for Success in College?" MidAIR and AIR Conference.
Frederick, S., G. Lowenstein,, and T. O'Donoghue (2002), "Time Discounting and Time Preferences: A Critical Review", *Journal of Economic Literatur*e, XL, 351-401.
Geiser, S., and V. Santelices, (2004) "The Role of Advanced Placement and Honors Courses in College Admissions" Center for studied on higher education, Research and Occasional Paper Series: CSHE.4.04.
Glewwe, P., N. Ilias and M. Kremer: "Teacher Incentives" (2003) National Bureau of Economic Reserach Working Paper Number 9671.
Heckman, J., and S. Cameron (2001) "The Dynamics of Educational Attainment for Black, Hispanic, and White Males" *Journal of Political Economy*, volume 109, pp 455–499.
Hauser R., (1993) "Trends in College Entry among Blacks, Whites, and Hispanics." In *Studies of Supply and Demand in Higher Education*, edited by Charles T. Clotfelter and Michael Rothschild. Chicago: Univ. Chicago Press
Holmstrom, B. and P. Milgrom, (1991) "Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design" *Journal of Law, Economics and Organization* 7 (Special Issue), pp 24–52.
Hudgins, K. (May 2003) "Advanced Placement program proves it pays to study hard: A Kick Start for College" Fiscal Notes.
Jacob, B. and Levitt, S. (2003). "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics*. 118(3): 843-877.
Kane, T., (October 1994) "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education." *Journal of Political Economy*, volume 102, pp 878–911.
Katz L., J. R. Kling, and J. Ludwig; (February 2005) "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Mobility Experiment." ; *Quarterly Journal of Economics* 120, 87-130.
Klopfenstein, K. (2004) "The advanced placement expansion of the 1990s: How did traditionally underserved students fare?" *Education Policy Analysis Archives*, volume 12 number 68.
Kohn, A. (1999) "Punished by rewards the trouble with gold stars, incentive plans, A's, praise, and other bribes." Bridgewater, NJ, Replica Books.
Kremer, M., E. Miguel, and R. Thornton, (2004) "Incentives to Learn" National Bureau of Economic Research Working Paper No. W10971.
Lavy, V (December 2002) "Evaluating the Effect of Teachers' Group Performance Incentives on Students Achievements." *Journal of Political Economy*, volume 110(6), 1286–1318.
Lavy, V (2004) "Performance Pay and Teachers' Effort, Productivity and Grading Ethics" National Bureau of Economic Research Working Paper No. 10622.
Long , B., and E. Bettinger, (April 2006) "Addressing the Needs of Under-Prepared College Students: Does College Remediation Work?" *Under submission*.
Lyon, J., (Oct 9, 2007) "Program to offer students money for high test scores." Arkansas News Bureau
Mathews, J. (Friday, March 9, 2007) "Group to Offer AP Exam Extra Credit: $250 Passing

Scores on College-Level Test at Some Schools Would Bring Bonuses for Students and Teachers" Washington Post

Mathews, J., (Tuesday, August 10, 2004) "Paying Teachers and Students for Good Scores" Washington Post.

Medina, J., (October 15, 2007) "Making Cash a Prize for High Scores on Advanced Placement Tests" The New York Times.

Morgan, R., and Ramist, L. (1998) "Advanced placement students in college: An investigation of course grades at 21 colleges." Princeton, NJ: Educational Testing Service.

Eberts, R., K. Hollenbeck, and J. Stone, (August 2000) "Teacher Performance Incentives and Student Outcomes" W.E. Upjohn Institute Working Paper No. 00-65.

Reber, S. (2005) "Court-Ordered Desegregation: Successes and Failures in Integration Since Brown" *Journal of Human Resources*, vol. 40 no. 3, pp. 559-590.

Stevenson, B., and J. Wolfers, (2006) "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress" (2006) *The Quarterly Journal of Economics*, MIT Press, vol. 121(1), pp 267-288.

Silverman, I., (November 2003) "Gender Differences in the Delay of Gratification: A Meta-Analysis," *Sex Roles* 49, pp 451-463.

Tables and Figures

Figure 1

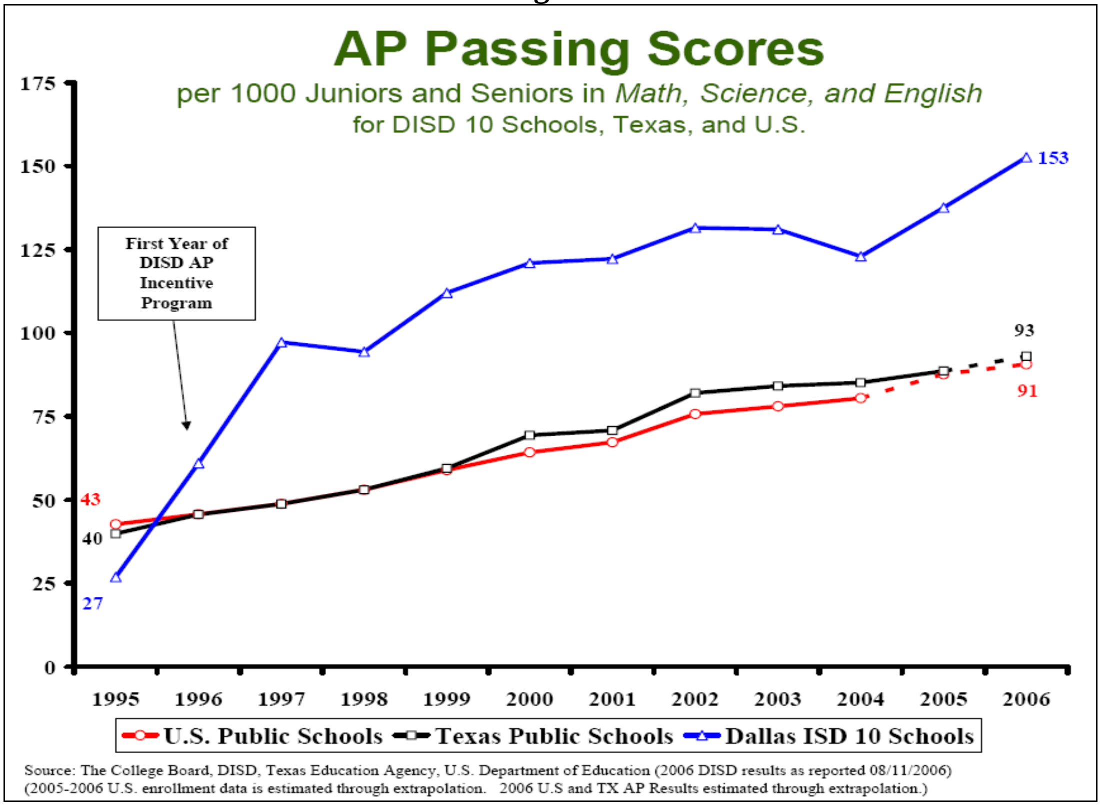
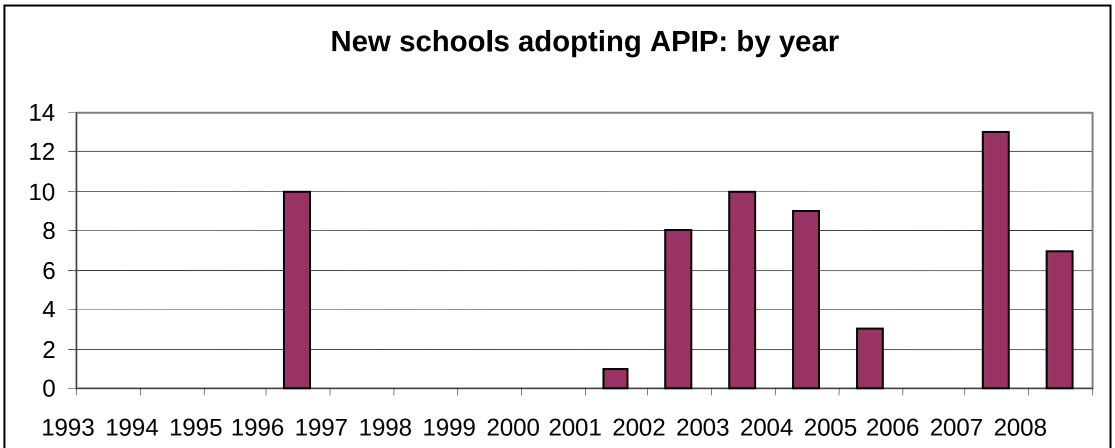


Figure 2



List of APIP districts: Abilene ISD, Amarillo ISD, Austin ISD (2008), Burkburnett ISD, Carrollton-farmers Branch ISD, City view ISD, Dallas ISD, Denison ISD, Galveston ISD, Houston ISD (2007), Pasadena ISD (2007), Pflugerville ISD, Tyler ISD, Wichita Falls ISD, Ysleta ISD (2007).

Table 1

Summary Statistics of Demographic Variables (Selected for APIP vs. Not Selected)

Sample	Schools Selected for Program				All Other High-Schools			
Years	1993-5	1996-9	2000-2	2003-5	1993-5	1996-9	2000-2	2003-5
Treated: Adopted APIP	0.00	0.16	0.23	0.57	0.00	0.00	0.00	0.00
	(0.0)	(0.37)	(0.42)	(0.5)	(0.0)	(0.0)	(0.0)	(0.0)
mobility	522.04	513.98	495.82	509.18	178.33	179.40	177.22	176.15
	(214.33)	(193.97)	(178.04)	(182.3)	(204.97)	(198.13)	(195.77)	(195.51)
%white	32.59	29.93	26.89	25.19	60.89	58.68	55.23	53.19
	(26.89)	(26.43)	(25.53)	(25.1)	(29.1)	(29.6)	(30.33)	(30.31)
%Black	30.79	29.85	28.05	25.42	10.39	10.29	11.05	11.32
	(30.18)	(29.18)	(27.66)	(24.98)	(15.89)	(15.52)	(17.08)	(17.08)
%Hispanic	33.30	37.00	41.71	46.19	27.42	29.61	32.23	33.88
	(25.81)	(26.7)	(27.44)	(27.62)	(28.57)	(29.03)	(29.4)	(29.28)
%Native American	0.30	0.34	0.36	0.38	0.25	0.28	0.31	0.38
	(0.28)	(0.31)	(0.32)	(0.39)	(0.76)	(0.91)	(0.76)	(1.04)
%Asian	3.02	2.88	2.99	2.83	1.05	1.11	1.19	1.24
	(3.43)	(3.48)	(3.94)	(3.78)	(2.68)	(2.79)	(2.83)	(2.96)
%Economically Disadvantaged	35.02	41.55	44.57	50.12	31.92	36.32	38.84	31.16
	(17.59)	(17.53)	(17.46)	(19.68)	(20.0)	(22.12)	(23.31)	(24.39)
%Limited English Proficiency	5.23	11.87	13.02	8.04	2.20	4.21	4.11	3.81
	(10.45)	(13.66)	(13.49)	(4.98)	(6.79)	(8.02)	(6.72)	(4.61)
City	0.90	0.90	0.91	0.87	0.17	0.18	0.22	0.24
	(0.3)	(0.3)	(0.29)	(0.33)	(0.38)	(0.39)	(0.42)	(0.43)
Urban Fringe	0.10	0.10	0.09	0.09	0.18	0.20	0.20	0.20
	(0.3)	(0.3)	(0.29)	(0.29)	(0.39)	(0.4)	(0.4)	(0.4)
Town	0.00	0.00	0.00	0.00	0.12	0.13	0.13	0.13
	(0.0)	(0.0)	(0.0)	(0.0)	(0.33)	(0.34)	(0.34)	(0.33)
Rural	0.00	0.00	0.00	0.04	0.49	0.46	0.44	0.43
	(0.0)	(0.0)	(0.0)	(0.19)	(0.5)	(0.5)	(0.5)	(0.5)
Grade 12 enrollment	302.86	315.49	332.82	338.79	136.10	140.58	145.96	152.60
	(133.26)	(132.27)	(139.12)	(149.85)	(160.26)	(168.45)	(177.44)	(188.62)
Lagged grade 11 enrollment	360.98	367.56	383.38	388.30	156.19	156.79	163.92	167.33
	(157.32)	(152.42)	(162.08)	(161.58)	(183.59)	(189.58)	(200.16)	(208.25)
Second lag grade 10 enrollment	-	440.40	458.28	460.57	-	178.46	181.31	178.83
	-	(175.54)	(178.61)	(176.21)	-	(213.21)	(218.66)	(220.33)
Graduates	275.58	294.84	331.86	346.39	126.27	135.04	144.35	153.65
	(127.58)	(125.64)	(136.09)	(151.48)	(148.28)	(159.85)	(174.35)	(187.69)

Standard errors in parentheses. The unit of observation is a school in a particular year.
Mobility is the number of students enrolled at the school for less than 83 percent of the school year

Table 2

Summary Statistics of Outcome Variables (Selected for APIP vs. Not Selected)

Sample	Schools Selected for Program				All other High-Schools			
Years	1993-5	1996-9	2000-2	2003-5	1993-5	1996-9	2000-2	2003-5
Take SAT/ACT	154.54	158.17	171.85	180.37	76.06	77.51	81.69	86.08
	(90.58)	(90.33)	(101.43)	(102.17)	(103.2)	(108.47)	(117.53)	(124.27)
Above Criterion on SAT/ACT	25.89	36.82	37.02	36.48	15.89	24.98	27.80	29.55
	(29.05)	(38.46)	(36.71)	(38.49)	(35.29)	(46.85)	(52.65)	(57.33)
College	-	-	136.69	140.68	-	-	90.21	95.48
	-	-	(86.26)	(91.81)	-	-	(101.9)	(108.45)
%Graduates above Criterion All	12.82	19.00	18.21	17.11	12.87	20.85	20.70	20.37
	(10.58)	(14.73)	(14.98)	(14.94)	(10.4)	(13.1)	(12.97)	(13.68)
%Graduates above Criterion Black	-	6.16	6.03	5.70	-	7.50	7.56	7.38
	-	(6.6)	(6.47)	(6.84)	-	(10.31)	(9.82)	(8.93)
%Graduates above Criterion White	-	34.96	35.65	36.31	-	26.95	27.69	27.49
	-	(18.02)	(16.93)	(17.39)	-	(14.33)	(13.82)	(14.73)
%Graduates above Criterion Hispanic	-	12.69	12.33	11.08	-	11.83	12.29	12.29
	-	(12.59)	(11.17)	(11.35)	-	(12.6)	(12.26)	(12.67)
%11 and 12 Graders taking AP or IB All	-	-	16.71	20.86	-	-	9.02	9.82
	-	-	(9.08)	(10.1)	-	-	(10.84)	(11.51)
%11 and 12 Graders taking AP or IB Black	-	-	10.15	13.21	-	-	5.66	6.72
	-	-	(7.01)	(8.48)	-	-	(9.15)	(10.45)
%11 and 12 Graders taking AP or IB White	-	-	26.08	29.48	-	-	11.88	12.87
	-	-	(18.01)	(16.99)	-	-	(13.85)	(14.97)
%11 and 12 Graders taking AP or IB Hispanic	-	-	11.88	15.42	-	-	7.05	7.92
	-	-	(8.75)	(8.99)	-	-	(9.6)	(10.07)
%9-12th graders taking college course	-	18.48	18.04	18.49	-	16.29	16.38	16.48
	-	(8.56)	(6.96)	(8.02)	-	(10.53)	(11.48)	(10.91)
% of Graduates who pass TAAS/TASP equivalency	35.64	35.00	55.65	64.78	41.28	42.54	64.87	73.34
	(16.01)	(13.87)	(13.9)	(13.88)	(15.85)	(14.31)	(14.03)	(14.15)
Observations	103	206	159	110	2349	5053	4060	2823

Standard errors in parentheses. The unit of observation is a school in a particular year.

Above Criterion on SAT/ACT is the number of high school graduates who score higher than 1100 on the SAT or 24 on the ACT. College is the count of high-school graduates who matriculate in college in the state of Texas the fall after graduation. A student who has attained TAAS/TASP equivalency has done well enough on the exit-level TAAS to have a 75% likelihood of passing the Texas academic Skills Program tests, which tests college readiness. IB examinations are the final exams for IB courses which are college level courses taken in high school.

Table 3

Summary Statistics for Demographic Variables (APIP Schools vs. Matched Schools)

Sample	Schools that Adopt the Program				Matched Sample			
Years	1993-5	1996-9	2000-2	2003-5	1993-5	1993-5	1993-5	1993-5
Treated	0.00	0.23	0.33	0.82	0.00	0.00	0.00	0.00
	(0.0)	(0.42)	(0.47)	(0.39)	(0.0)	(0.0)	(0.0)	(0.0)
mobility	478.03	480.86	458.73	477.26	418.21	396.33	392.84	387.60
	(196.27)	(181.93)	(173.97)	(190.46)	(231.26)	(193.69)	(188.52)	(193.03)
%white	33.81	32.11	29.75	29.05	46.07	45.06	42.21	39.76
	(30.81)	(30.15)	(28.53)	(28.0)	(32.05)	(31.67)	(31.03)	(30.26)
%Black	37.67	36.19	33.50	30.62	22.96	22.43	22.22	22.25
	(32.63)	(31.58)	(29.9)	(27.33)	(27.18)	(25.95)	(24.66)	(23.94)
%Hispanic	25.63	28.70	33.45	37.06	27.78	29.40	32.38	34.75
	(23.72)	(24.84)	(25.7)	(26.0)	(28.13)	(27.98)	(27.97)	(27.99)
%Native American	0.35	0.40	0.43	0.46	0.20	0.22	0.23	0.28
	(0.31)	(0.34)	(0.35)	(0.43)	(0.22)	(0.21)	(0.25)	(0.29)
%Asian	2.55	2.60	2.87	2.81	2.98	2.89	2.95	2.96
	(3.17)	(3.54)	(4.19)	(4.14)	(4.37)	(4.09)	(4.44)	(4.53)
%Economically Disadvantaged	35.56	41.23	42.60	45.82	26.90	32.21	37.51	40.27
	(19.3)	(18.49)	(16.6)	(18.95)	(20.17)	(22.35)	(24.49)	(25.67)
%Limited English Proficiency	4.74	11.88	13.02	7.10	3.39	6.61	6.84	6.01
	(10.91)	(15.25)	(14.94)	(4.75)	(8.27)	(9.49)	(8.91)	(5.49)
City	0.94	0.94	0.95	0.92	0.69	0.68	0.69	0.70
	(0.23)	(0.23)	(0.23)	(0.27)	(0.46)	(0.47)	(0.46)	(0.46)
Urban Fringe	0.06	0.06	0.05	0.05	0.23	0.23	0.22	0.22
	(0.23)	(0.23)	(0.23)	(0.22)	(0.42)	(0.42)	(0.42)	(0.42)
Town	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
	(0.0)	(0.0)	(0.0)	(0.0)	(0.08)	(0.08)	(0.08)	(0.08)
Rural	0.00	0.00	0.00	0.03	0.08	0.08	0.07	0.07
	(0.0)	(0.0)	(0.0)	(0.16)	(0.27)	(0.27)	(0.26)	(0.26)
Grade 12 enrollment	-	288.09	303.10	323.29	-	320.13	339.61	349.51
	-	(123.71)	(135.33)	(151.33)	-	(159.71)	(174.77)	(180.97)
Lagged grade 11 enrollment	327.20	341.64	340.63	375.53	342.53	365.38	381.54	384.03
	(145.18)	(140.17)	(147.24)	(155.73)	(177.47)	(182.12)	(193.16)	(195.93)
Second lag grade 10 enrollment	-	409.80	410.34	438.99	-	422.64	439.58	426.27
	-	(168.06)	(173.07)	(179.12)	-	(214.07)	(220.69)	(222.55)
Graduates	250.85	275.08	307.24	330.69	269.20	298.32	327.12	340.07
	(129.75)	(124.67)	(134.53)	(158.8)	(141.06)	(149.33)	(170.4)	(178.65)
Graduates/2nd lag grade 10	-	0.67	0.74	0.76	-	0.73	0.76	0.85
	-	(0.11)	(0.1)	(0.11)	-	(0.14)	(0.13)	(0.46)

Standard errors in parentheses. The unit of observation is a school in a particular year.

Table 4

Summary Statistics for Outcome Variables (APIP Schools vs. Matched Schools)

Sample	Schools Selected for Program				Matched Sample			
Years	1993-5	1996-9	2000-2	2003-5	1993-5	1993-5	1993-5	1993-5
Take SAT/ACT	137.96	147.05	157.56	166.13	162.93	174.87	189.35	194.40
	(90.24)	(93.2)	(91.4)	(93.89)	(108.29)	(109.42)	(126.02)	(127.79)
Above Criterion on SAT/ACT	24.28	36.94	37.63	37.97	34.57	49.60	53.73	55.69
	(30.43)	(42.65)	(39.75)	(39.15)	(42.54)	(50.46)	(58.47)	(62.41)
College	-	-	127.84	133.70	-	-	159.12	163.99
	-	-	(85.34)	(93.43)	-	-	(105.66)	(109.53)
%Graduates above Criterion All	12.16	18.58	18.30	18.37	15.01	22.06	21.64	21.42
	(10.68)	(14.89)	(15.37)	(15.38)	(11.63)	(14.21)	(14.56)	(15.13)
%Graduates above Criterion Black	4.58	5.63	5.95	5.60	4.95	7.85	7.64	7.40
	(6.57)	(6.07)	(6.58)	(6.87)	(5.9)	(9.2)	(8.17)	(8.0)
%Graduates above Criterion White	24.09	37.16	36.87	39.08	24.13	34.56	35.12	35.66
	(11.22)	(15.6)	(13.6)	(12.17)	(14.99)	(14.98)	(14.46)	(15.67)
%Graduates above Criterion Hispanic	6.63	12.85	12.55	12.02	7.65	14.16	13.55	13.82
	(8.74)	(11.77)	(11.57)	(11.85)	(9.21)	(12.0)	(11.25)	(12.04)
%11 and 12 Graders taking AP or IB All	-	-	18.04	22.79	-	-	13.01	15.41
	-	-	(7.96)	(9.4)	-	-	(8.17)	(9.1)
%11 and 12 Graders taking AP or IB Black	-	-	11.01	14.68	-	-	6.79	8.58
	-	-	(6.79)	(8.2)	-	-	(6.67)	(7.47)
%11 and 12 Graders taking AP or IB White	-	-	27.55	31.65	-	-	18.83	22.14
	-	-	(14.75)	(16.15)	-	-	(14.14)	(14.35)
%11 and 12 Graders taking AP or IB Hispanic	-	-	12.52	16.00	-	-	8.91	11.33
	-	-	(8.79)	(9.39)	-	-	(6.69)	(7.9)
%9-12th graders taking college course	-	19.45	18.12	18.76	-	18.57	19.47	19.34
	-	(9.09)	(6.78)	(8.22)	-	(9.36)	(10.98)	(9.02)
Observations	85	170	134	91	277	564	441	296

Standard errors in parentheses. The unit of observation is a school in a particular year.

Above Criterion on SAT/ACT is the number of high school graduates who score higher than 1100 on the SAT or 24 on the ACT. College is the count of high-school graduates who matriculate in college in the state of Texas the fall after graduation. IB examinations are the final exams for IB courses which are college level courses taken in high school.

Table 5

Estimation of the Propensity Score

Variable	Treated
Mid sized city	1.933
	[1.025]
Urban fringe of large city	-0.862
	[0.937]
Urban fringe of mid sized city	-5.535
	[1.855]**
Rural	2.155
	[1.312]
log mobility	0.169
	[0.125]
%while	-0.666
	[0.756]
%Native American	0.419
	[0.755]
%Asian	-0.675
	[0.756]
%Limited English Proficiency	0.072
	[0.007]**
%Black	-0.626
	[0.756]
Mid sized city*%Black	-0.049
	[0.010]**
Urban fringe of large city*%Black	0.01
	[0.007]
Urban fringe of mid sized city*%Black	-0.046
	[0.014]**
Rural*%Black	0.033
	[0.014]*
%Hispanic	-0.644
	[0.756]
Mid sized city*%Hispanic	-0.054
	[0.012]**
Urban fringe of large city*%Hispanic	0.016
	[0.009]
Urban fringe of mid sized city*%Hispanic	-0.079
	[0.020]**
Rural*%Hispanic	0.015
	[0.018]
%Economically Disadvantaged	-0.032
	[0.005]**
Mid sized city*%Economically Disadvantaged	0.039
	[0.012]**
Urban fringe of large city*%Economically Disadvantaged	-0.044
	[0.012]**
Urban fringe of mid sized city*%Economically Disadvantaged	0.046
	[0.017]**
Rural*%Economically Disadvantaged	-0.069
	[0.022]**

Variables Cont'd	
Log of grade 12 enrollment	-0.044
	[0.369]
Mid sized city*Log of grade 12 enrollment	-0.097
	[0.157]
Urban fringe of large city*Log of grade 12 enrollment	0.009
	[0.137]
Urban fringe of mid sized city*Log of grade 12 enrollment	0.96
	[0.334]**
Rural*Log of grade 12 enrollment	-0.569
	[0.213]**
first lag of log of grade 12 enrollment	-1.413
	[0.468]**
second lag of log of grade 12 enrollment	-1.096
	[0.482]*
first lag of log of grade 11 enrollment	0.411
	[0.394]
lag of log of SAT/ACT takers	0.12
	[0.304]
lag of above criterion	0.068
	[0.120]
lag of number of graduates	1.296
	[0.488]**
second lag of log of SAT/ACT takers	-0.206
	[0.306]
second lag of above criterion	0.045
	[0.114]
second lag of number of graduates	1.151
	[0.515]*
year==1997	-1.728
	[0.262]**
year==1998	-1.653
	[0.239]**
year==1999	-1.86
	[0.257]**
year==2000	-1.801
	[0.224]**
year==2001	-1.635
	[0.221]**
year==2002	-1.151
	[0.193]**
year==2003	-0.567
	[0.159]**
year==2004	-0.353
	[0.144]*
Constant	61.863
	[75.482]
Observations	5888

Robust standard errors in brackets
* significant at 5%; ** significant at 1%

Table 6

Summary Statistics for Demographic Variables (APIP Adopters vs. Future APIP Adopters)

Sample	Schools that adopt the APIP				Schools that will adopt APIP			
Years	1993-5	1996-9	2000-2	2003-5	1993-5	1993-5	1993-5	1993-5
Treated	0.00	0.23	0.33	0.82	0.00	0.00	0.00	0.00
	(0.0)	(0.42)	(0.47)	(0.39)	(0.0)	(0.0)	(0.0)	(0.0)
mobility	478.03	480.86	458.73	477.26	618.31	587.45	586.96	583.67
	(196.27)	(181.93)	(173.97)	(190.46)	(226.74)	(200.96)	(154.99)	(137.36)
%white	33.81	32.11	29.75	29.05	29.87	25.10	19.87	16.21
	(30.81)	(30.15)	(28.53)	(28.0)	(14.96)	(14.26)	(13.84)	(12.82)
%Black	37.67	36.19	33.50	30.62	15.54	15.77	14.65	13.26
	(32.63)	(31.58)	(29.9)	(27.33)	(15.62)	(15.71)	(14.27)	(11.55)
%Hispanic	25.63	28.70	33.45	37.06	50.32	55.41	62.00	67.48
	(23.72)	(24.84)	(25.7)	(26.0)	(22.06)	(20.95)	(20.2)	(18.05)
%Native American	0.35	0.40	0.43	0.46	0.20	0.20	0.19	0.18
	(0.31)	(0.34)	(0.35)	(0.43)	(0.16)	(0.16)	(0.16)	(0.17)
%Asian	2.55	2.60	2.87	2.81	4.06	3.51	3.30	2.87
	(3.17)	(3.54)	(4.19)	(4.14)	(3.81)	(3.28)	(3.24)	(2.86)
%Economically Disadvantaged	35.56	41.23	42.60	45.82	33.82	42.28	49.44	60.14
	(19.3)	(18.49)	(16.6)	(18.95)	(13.24)	(15.29)	(18.72)	(17.83)
%Limited English Proficiency	4.74	11.88	13.02	7.10	6.33	11.85	13.02	10.24
	(10.91)	(15.25)	(14.94)	(4.75)	(9.43)	(9.34)	(9.12)	(4.87)
City	0.94	0.94	0.95	0.92	0.81	0.81	0.80	0.76
	(0.23)	(0.23)	(0.23)	(0.27)	(0.4)	(0.39)	(0.4)	(0.44)
Urban Fringe	0.06	0.06	0.05	0.05	0.19	0.19	0.20	0.18
	(0.23)	(0.23)	(0.23)	(0.22)	(0.4)	(0.39)	(0.4)	(0.39)
Town	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Rural	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.06
	(0.0)	(0.0)	(0.0)	(0.16)	(0.0)	(0.0)	(0.0)	(0.24)
Grade 12 enrollment	-	288.09	303.10	323.29	-	379.16	408.69	385.88
	-	(123.71)	(135.33)	(151.33)	-	(130.02)	(121.79)	(134.5)
Lagged grade 11 enrollment	327.20	341.64	340.63	375.53	434.88	440.08	490.42	441.57
	(145.18)	(140.17)	(147.24)	(155.73)	(161.98)	(151.86)	(151.35)	(157.97)
Second lag grade 10 enrollment	-	409.80	410.34	438.99	-	523.79	574.13	532.59
	-	(168.06)	(173.07)	(179.12)	-	(167.53)	(144.19)	(126.86)
Graduates	250.85	275.08	307.24	330.69	330.47	338.67	392.33	383.03
	(129.75)	(124.67)	(134.53)	(158.8)	(105.02)	(117.24)	(121.34)	(127.62)
Graduates/2nd lag grade 10	-	0.67	0.74	0.76	-	0.66	0.69	0.71
	-	(0.11)	(0.1)	(0.11)	-	(0.13)	(0.11)	(0.14)

Standard errors in parentheses. The unit of observation is a school in a particular year.

Table 7

Summary Statistics for Outcome Variables (APIP Adopters vs. Future APIP Adopters)

Sample	Schools that adopt the APIP				Schools that will adopt APIP			
Years	1993-5	1996-9	2000-2	2003-5	1993-5	1993-5	1993-5	1993-5
Take SAT/ACT	137.96	147.05	157.56	166.13	191.31	182.85	206.95	213.61
	(90.24)	(93.2)	(91.4)	(93.89)	(81.15)	(78.84)	(116.42)	(114.0)
Above Criterion on SAT/ACT	24.28	36.94	37.63	37.97	29.48	36.53	35.51	32.99
	(30.43)	(42.65)	(39.75)	(39.15)	(25.81)	(27.23)	(28.2)	(37.29)
College	-	-	127.84	133.70	-	-	157.69	156.97
	-	-	(85.34)	(93.43)	-	-	(87.49)	(87.13)
%Graduates above Criterion All	12.16	18.58	18.30	18.37	14.27	19.93	17.98	14.17
	(10.68)	(14.89)	(15.37)	(15.38)	(10.36)	(14.45)	(14.15)	(13.61)
%Graduates above Criterion Black	4.58	5.63	5.95	5.60	6.26	7.55	6.26	5.97
	(6.57)	(6.07)	(6.58)	(6.87)	(6.34)	(7.69)	(6.19)	(6.88)
%Graduates above Criterion White	24.09	37.16	36.87	39.08	24.10	31.69	33.51	31.54
	(11.22)	(15.6)	(13.6)	(12.17)	(17.13)	(20.81)	(21.62)	(23.36)
%Graduates above Criterion Hispanic	6.63	12.85	12.55	12.02	9.32	12.40	11.85	9.13
	(8.74)	(11.77)	(11.57)	(11.85)	(14.29)	(14.02)	(10.37)	(10.12)
%11 and 12 Graders taking AP or IB All	-	-	18.04	22.79	-	-	13.45	16.35
	-	-	(7.96)	(9.4)	-	-	(10.83)	(10.36)
%11 and 12 Graders taking AP or IB Black	-	-	11.01	14.68	-	-	7.91	9.69
	-	-	(6.79)	(8.2)	-	-	(7.21)	(8.23)
%11 and 12 Graders taking AP or IB White	-	-	27.55	31.65	-	-	23.15	25.27
	-	-	(14.75)	(16.15)	-	-	(23.22)	(18.04)
%11 and 12 Graders taking AP or IB Hispanic	-	-	12.52	16.00	-	-	10.32	14.09
	-	-	(8.79)	(9.39)	-	-	(8.59)	(7.95)
%9-12th graders taking college course	-	19.45	18.12	18.76	-	16.32	17.82	17.87
	-	(9.09)	(6.78)	(8.22)	-	(6.84)	(7.47)	(7.61)
Observations	85	170	134	91	32	64	46	33

Standard errors in parentheses. The unit of observation is a school in a particular year.
Above Criterion on SAT/ACT is the number of high school graduates who score higher than 1100 on the SAT or 24 on the ACT. College is the count of high-school graduates who matriculate in college in the state of Texas the fall after graduation. IB examinations are the final exams for IB courses which are college level courses taken in high school.

Table 8

APIP Adoption effects by sample used (Before vs. After comparisons)

			Outcome Variables				
			1	2	3	4	5
Estimation Sample	Control for differential APIP School Trend? (Method)	Variable	log grade 12	log Graduates	log SAT/ACT takers	log Above Criterion	log Attend College
All schools	NO	Treat	0.02 [0.023]	**0.046** **[0.020]***	0.05 [0.038]	0.013 [0.061]	**0.091** **[0.027]****
All schools	YES (linear trend)	Treat	**0.07** **[0.020]****	**0.044** **[0.018]***	0.04 [0.039]	**0.166** **[0.053]****	**0.073** **[0.027]****
		APIP Trend	**-0.01** **[0.003]****	0 [0.004]	0.002 [0.004]	**-0.03** **[0.012]***	0.01 [0.012]
Matched	YES (Group Trends)	Treat	**0.048** **[0.021]***	0.018 [0.014]	0.032 [0.028]	**0.113** **[0.040]****	**0.059** **[0.031]+**
APIP schools	YES (Year Effects)	Treat	**0.057** **[0.018]****	0.029 [0.021]	0.004 [0.031]	**0.223** **[0.034]****	**0.067** **[0.027]***
APIP schools (Late adopters)	YES (Year Effects)	Treat	0.042 [0.026]	-0.005 [0.024]	-0.061 [0.058]	**0.296** **[0.072]****	**0.09** **[0.046]+**

* significant at 5%; ** significant at 1%, + significant at the 10% level

Robust standard errors in brackets. Standard errors clustered at school district level. Each estimate represents a separate regression for each sample. All regressions control for school demographic variables, school fixed effects and time fixed effects. Baseline regressions are from Appendix table A1, the matched regression results are from A2 and the APIP and late adopters APIP results are from Table A3

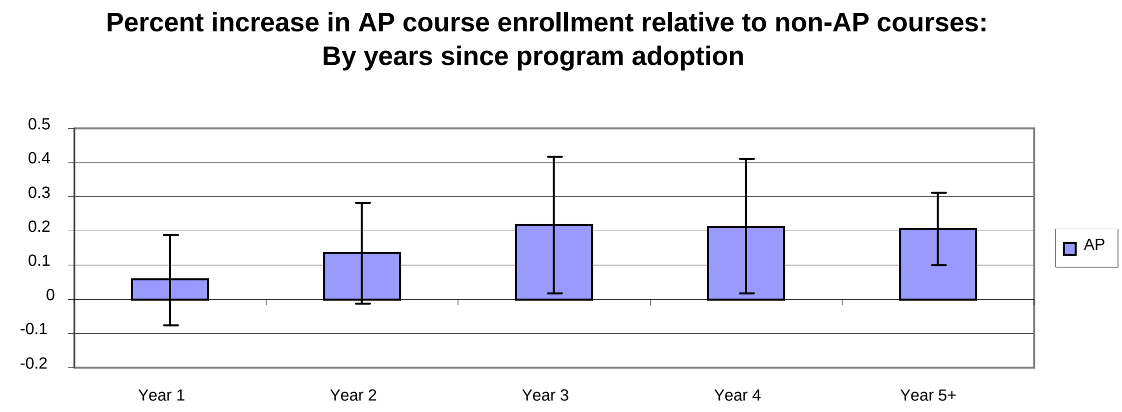
Table 9

Effect of APIP on graduates per 10th grader and TAAS/TAKS Equivalency

On APIP Sample	Log Graduates	% Graduates achieving TAAS/TASP Equivalency	% Graduates achieving TAAS/TASP Equivalency
Treat	0.02	0.642	1.384
	[0.017]	[1.764]	[1.234]
Year*Decile Fixed Effects?	YES	YES	YES
School Fixed Effects?	YES	YES	YES
Conditional on grade 10 enrollment	YES	NO	YES
Conditional on grade 12 enrollment	NO	YES	NO
Observations	480	575	357
Number of campus	57	57	54
R-squared	0.71	0.87	0.85

Robust standard errors in brackets. Clustered at the school district level

Figure 3



This figure shows the coefficients of the DID estimates of the effect of adoption on log AP-enrollment where the control group is log non-AP enrollment using the APIP sample for the years 1999 through 2004. Regressions include school and year fixed effects. See Table 10 for DID estimates of APIP adoption on the log of AP reenrollment (not relative to non-AP enrolment).

Table 10

Effects of First, Second, and Third and More Years after APIP adoption

Sample: schools that will have adopted APIP by 2008	1	2	3	4	5	6	7
	log grade 12 enrolment	log graduates	log SAT/ACT taking	ln above criterion in SAT/ACT	log attend college	Log AP course enrollment	%11 and 12th graders taking AP/IB exam
First year	0.029	0.027	0.003	**0.189**	**0.06**	0.069	**2.019**
	(0.02)	(0.018)	(0.028)	**(0.042)****	**(0.028)+**	(0.057)	**(0.923)***
Second year	**0.054**	0.027	0.016	**0.223**	**0.083**	0.027	**4.183**
	(0.020)*	(0.024)	(0.052)	**(0.036)****	**(0.027)****	(0.087)	**(1.100)****
Third year and beyond	**0.141**	0.038	-0.011	**0.333**	**0.067**	**0.217**	**4.237**
	(0.033)**	(0.034)	(0.033)	**(0.062)****	**(0.039)+**	**(0.084)***	**(1.323)****
Controls?	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
School fixed effects?	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year*Decile fixed Effects?	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Observations	578	578	578	578	226	210	226
Number of campus	57	57	57	57	57	56	57
R-squared	0.5	0.81	0.61	0.39	0.55	0.21	0.46

Robust standard errors in parentheses + significant at 10%; * significant at 5%; ** significant at 1%

Standard errors clustered at the school district level.

Table 11

APIP effect on % of 11 and 12th graders who took at least one AP or IB exam: by gender and ethnicity

Sample: schools that will have adopted APIP by 2008	1	2	3	4	5	6
	All	Black	White	Hispanic	Female	Male
First year	**2.019**	**2.505**	-0.203	**2.033**	**2.035**	**1.943**
	(0.923)*	**(0.619)****	(1.896)	**(1.035)+**	**(0.911)***	**(0.951)+**
Second year	**4.183**	**4.293**	-1.163	**5.309**	**4.337**	**3.705**
	(1.100)**	**(1.891)***	(2.43)	**(1.820)***	**(1.225)****	**(1.288)***
Third year and beyond	**4.237**	2.673	0.388	**6.099**	2.125	**5.657**
	(1.323)**	(1.717)	(2.058)	**(2.893)+**	(1.287)	**(1.665)****
School Fixed effects?	YES	YES	YES	YES	YES	YES
Controls?	YES	YES	YES	YES	YES	YES
Year*Decile fixed effects?	YES	YES	YES	YES	YES	YES
Observations	226	221	195	225	217	217
Number of campus	57	56	51	57	56	56
R-squared	0.46	0.37	0.31	0.32	0.45	0.38

Robust standard errors in parentheses + significant at 10%; * significant at 5%; ** significant at 1%

Standard errors clustered at the school district level.

Table 12

APIP effect on percentage of graduates who score above criterion (APIP sample)

Sample: Schools that will have adopted APIP by 2008	1	2	3	4	5	6
	All	Black	White	Hispanic	Female	Male
First year	**1.966**	0.623	**2.421**	1.433	1.232	**2.191**
	(0.592)**	(0.751)	**(1.384)+**	(1.341)	-0.893	**(0.968)***
Second year	**2.785**	1.091	**3.879**	**7.835**	**2.125**	**4.001**
	(0.486)**	(0.926)	**(1.180)****	**(2.449)****	**(1.018)+**	**(1.451)***
Third year and beyond	**4.014**	**5.077**	**4.19**	**6.54**	**4.704**	**5.887**
	(1.079)**	**(1.873)***	**(1.535)***	**(3.172)+**	**(1.905)***	**(2.457)***
School Fixed effects?	YES	YES	YES	YES	YES	YES
Controls?	YES	YES	YES	YES	YES	YES
Year*Decile fixed effects?	YES	YES	YES	YES	YES	YES
Observations	693	402	318	381	424	423
Number of campus	57	55	48	53	57	57
R-squared	0.42	0.25	0.32	0.16	0.27	0.31

Robust standard errors in parentheses + significant at 10%; * significant at 5%; ** significant at 1%

Standard errors clustered at the school district level.

Table 13

APIP effect on SAT/ACT performance of non movers: by ethnicity

Sample: All students who attended their 12th grade school for at least four years	1	2	3	4	5	6
Sub-sample	Black	Hispanic	White	Black	Hispanic	White
	Above Criterion	Above Criterion	Above Criterion	Over 900/19 on SAT/ACT	Over 900/19 on SAT/ACT	Over 900/19 on SAT/ACT
	Missing values assumed to be 2					
First year	-	0.018	**0.165**	0.04	**0.234**	0.005
	-	(0.047)	**(0.055)****	(0.052)	**(0.074)****	(0.047)
Second year	**-**	**0.12**	**0.242**	-0.12	**0.378**	0.045
	-	**(0.051)***	**(0.084)***	(0.117)	**(0.116)****	(0.063)
Year 3+	**-**	**0.18**	**0.262**	0.123	**0.385**	**0.257**
	-	**(0.038)****	**(0.15)+**	(0.092)	**(0.204)+**	**(0.096)***
	Missing values assumed to be 4					
First year	-	0.023	**0.124**	0.031	**0.183**	0.009
	-	(0.019)	**(0.046)***	(0.04)	**(0.058)****	(0.032)
Second year	**-**	**0.084**	**0.166**	-0.104	**0.294**	0.043
	-	**(0.027)****	**(0.075)***	(0.075)	**(0.089)****	(0.035)
Year 3+	**-**	**0.108**	**0.221**	0.043	**0.313**	**0.222**
	-	**(0.023)****	**(0.122)**	(0.065)	**(0.167)+**	**(0.063)****
	Missing values assumed to be 0					
First year	-	0.006	**0.254**	0.057	**0.342**	-0.003
	-	(0.113)	**(0.084)****	(0.091)	**(0.116)***	(0.096)
Second year	-	0.197	**0.405**	-0.156	**0.56**	0.051
	-	(0.13)	**(0.118)****	(0.227)	**(0.182)****	(0.139)
Year 3+	**-**	**0.335**	0.351	0.295	**0.54**	**0.334**
	-	**(0.120)***	(0.217)	(0.175)	**(0.292)+**	**(0.185)+**

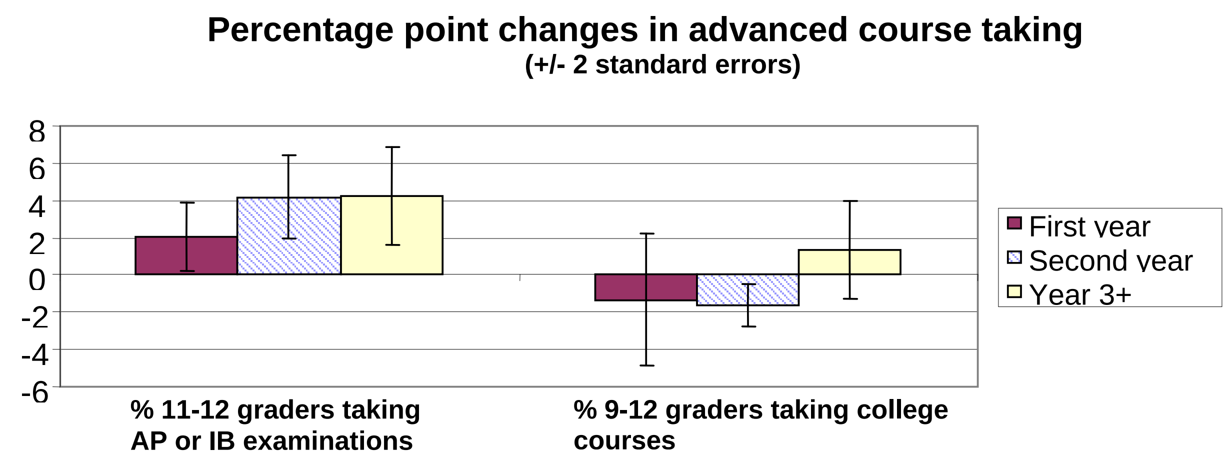
Note: All regression specifications include school level controls, school fixed effects, year fixed effects and propensity score decile time year fixed effects. Treatment effects are not estimated for above criterion for blacks because there are too many missing values. Since the data are heavily masked cells with counts fewer than 5 are not reported. To provide bounds for those masked observations missing values are assumed to be 4 and 0.

Due to data masking count fewer than 5 are coded as missing.

Robust standard errors in parentheses. Standard errors clustered at school district level.

* significant at 5%; ** significant at 1% ; + significant at 10 %

Figure 4



This figure shows the coefficient estimates of DID regressions using the APIP sample. The regressions include all school demographic controls, school fixed effects, and year*(propensity score decile) fixed effects. The standard errors are clustered at the school district level.

Table 14

Differences in Differences Regressions on APIP sample: by Incentive level

Sample: APIP schools only	1	2	3	4	5	6	7
	log grade 12	log Graduates	log SAT/ACT takers	log Above Criterion	log Attend College	log AP enrollment	% 11th and 12th grades taking AP or IB exams
$100 Per passing AP exam	**0.103**	**0.043**	0.012	**0.227**	-0.001	0.009	**3.702**
	[0.028]**	**[0.020]***	[0.027]	**[0.046]****	[0.049]	[0.095]	**[1.317]***
$101-$499 Per passing AP exam	-0.062	0.037	-0.098	**0.247**	-	**0.195**	-0.136
	[0.057]	[0.066]	[0.051]	**[0.069]****	-	**[0.063]****	[1.678]
500 Per passing AP exam	-0.016	-0.01	0.026	**0.203**	**0.108**	**0.136**	**1.951**
	[0.033]	[0.030]	[0.048]	**[0.045]****	**[0.027]****	**[0.032]****	**[0.571]****
Controls?	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
School fixed effects?	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Year*Decile fixed Effects?	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Observations	578	578	578	578	226	313	217
Number of campus	57	57	57	57	57	57	56
R-squared	0.49	0.81	0.62	0.39	0.56	0.17	0.4

Robust standard errors in brackets Standard errors clustered at the school district level

* significant at 5%; ** significant at 1%

Figure 5

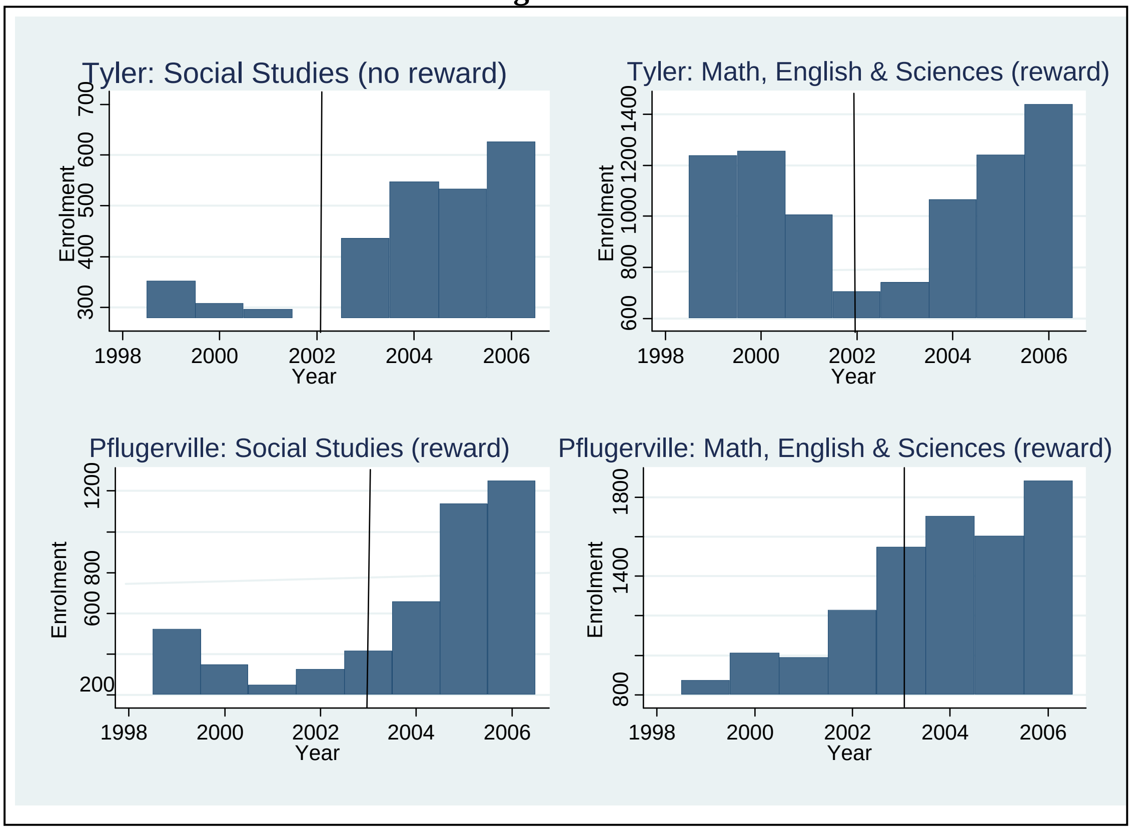


Table 15

APIPeffect on ratio of social studies AP enrolles to total AP enrolees: by subjects rewarded

	1	2	4	5
Treated	-0.036	-0.016	-0.025	-0.006
	[0.022]+	[0.022]	[0.025]	[0.026]
Treated*(rewards for social sciences)	-	-	-0.029	-0.026
	-	-	[0.035]	[0.035]
School Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Sample	All	APIP	All	APIP
Observations	557	211	557	211
R-squared	0.5	0.6	0.5	0.6

Robust standard errors in brackets
+ significant at 10%; * significant at 5%; ** significant at 1%

VIII. Appendix

Table A1

Baseline Difference-in-Differences Regressions with school fixed effects

Sample: All Texas High schools (with SAT taking data)	1	2	3	4	5	6	7	8	9	10
	log grade 12	log Graduates	log SAT/ACT takers	log Above Criterion	log Attend College	log grade 12	log Graduates	log SAT/ACT takers	log Above Criterion	log Attend College
treat	0.02	**0.046**	0.05	0.013	**0.091**	**0.07**	**0.044**	0.04	**0.166**	**0.073**
	[0.023]	**[0.020]***	[0.038]	[0.061]	**[0.027]****	**[0.020]****	**[0.018]***	[0.039]	**[0.053]****	**[0.027]****
year*APIP school	-	-	-	-	-	**-0.01**	0.001	0.002	**-0.03**	0.01
	-	-	-	-	-	**[0.003]****	[0.004]	[0.004]	**[0.012]***	[0.012]
mobility	-0.153	0.007	0.008	-0.015	-0.059	-0.152	0.007	0.008	-0.012	-0.06
	[0.013]**	[0.010]	[0.011]	[0.016]	[0.026]*	[0.014]**	[0.010]	[0.011]	[0.016]	[0.026]*
% white	0.025	0.045	0.034	-0.035	0.077	0.025	0.045	0.034	-0.036	0.076
	[0.029]	[0.019]*	[0.050]	[0.059]	[0.050]	[0.029]	[0.019]*	[0.050]	[0.060]	[0.050]
% black	0.025	0.043	0.028	-0.045	0.076	0.024	0.043	0.028	-0.046	0.075
	[0.030]	[0.019]*	[0.050]	[0.059]	[0.050]	[0.029]	[0.019]*	[0.050]	[0.060]	[0.050]
% Hispanic	0.024	0.045	0.028	-0.052	0.074	0.024	0.045	0.028	-0.052	0.074
	[0.029]	[0.019]*	[0.050]	[0.059]	[0.050]	[0.029]	[0.019]*	[0.050]	[0.060]	[0.050]
% Econ. Disadvantaged	-0.002	0.0002	0.001	-0.001	-0.001	-0.002	0.0003	0.0004	-0.001	-0.001
	[0.000]**	[0.000]	[0.001]	[0.001]	[0.001]	[0.000]**	[0.000]	[0.001]	[0.001]	[0.001]
% Limited English Proficiency	0.0001	0	-0.002	-0.003	0.0001	0.00004	0	-0.002	-0.003	0.00002
	[0.001]	[0.000]	[0.001]*	[0.002]*	[0.002]	[0.001]	[0.000]	[0.001]*	[0.002]*	[0.002]
lag of grade 11 enrollment	0.277	0.031	0.02	0.104	0.206	0.277	0.031	0.02	0.107	0.206
	[0.018]**	[0.012]**	[0.014]	[0.017]**	[0.033]**	[0.018]**	[0.012]**	[0.014]	[0.017]**	[0.033]**
log grade12 enrollment	-	0.774	0.75	0.564	0.473	-	0.774	0.75	0.561	0.473
	-	[0.032]**	[0.026]**	[0.034]**	[0.056]**	-	[0.032]**	[0.026]**	[0.034]**	[0.056]**
Constant	2.111	-3.644	-2.623	3.272	-6.513	2.113	-3.644	-2.624	3.287	-6.445
	[2.914]	[1.893]	[4.967]	[5.929]	[5.022]	[2.891]	[1.895]	[4.973]	[5.993]	[5.022]
Year Fixed Effects?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
School Fixed Effects?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	12808	12808	12808	12808	4538	12808	12808	12808	12808	4538
Number of schools	1434	1434	1434	1434	1286	1434	1434	1434	1434	1286
R-squared	0.31	0.7	0.41	0.26	0.21	0.31	0.7	0.41	0.26	0.21

Robust standard errors in brackets. Standard errors clustered at school district level.

* significant at 5%; ** significant at 1%

TableA2

Difference-in-Differences Regressions with school fixed effects on matched sample

Sample: APIP adopters & Matched schools	1	2	3	4	5	6	7	8	9	10
	log grade 12	log Graduates	log SAT/ACT takers	log Above Criterion	log Attend College	log grade 12	log Graduates	log SAT/ACT takers	log Above Criterion	log Attend College
treat	**0.054**	0.024	0.039	**0.135**	**0.064**	**0.048**	0.018	0.032	**0.113**	**0.059**
	[0.027]*	[0.018]	[0.033]	**[0.039]****	**[0.034]**	**[0.021]***	[0.014]	[0.028]	**[0.040]****	**[0.031]**
mobility	-0.01	0.045	-0.023	-0.201	-0.08	-0.015	0.049	-0.014	-0.169	-0.069
	[0.030]	[0.030]	[0.055]	[0.046]**	[0.060]	[0.027]	[0.026]	[0.047]	[0.043]**	[0.058]
% white	0.014	0.02	0.05	-0.304	-0.068	0.011	0.017	0.047	-0.309	-0.083
	[0.042]	[0.035]	[0.052]	[0.144]*	[0.091]	[0.043]	[0.032]	[0.052]	[0.140]*	[0.087]
% black	0.01	0.015	0.04	-0.312	-0.09	0.008	0.012	0.036	-0.317	-0.105
	[0.044]	[0.035]	[0.051]	[0.146]*	[0.093]	[0.044]	[0.032]	[0.052]	[0.142]*	[0.089]
% Hispanic	0.016	0.02	0.044	-0.331	-0.071	0.014	0.017	0.04	-0.337	-0.085
	[0.042]	[0.036]	[0.050]	[0.147]*	[0.092]	[0.043]	[0.033]	[0.051]	[0.142]*	[0.087]
% Econ. Disadvantaged	-0.001	0.0002	0.001	-0.001	0.0004	-0.001	0.001	0.001	-0.001	0.0002
	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]
% Limited English Proficiency	0.001	0.0004	-0.003	-0.001	0.001	0.0003	0.0004	-0.003	-0.001	0.001
	[0.001]	[0.001]	[0.001]**	[0.003]	[0.003]	[0.001]	[0.001]	[0.001]**	[0.003]	[0.003]
lag of grade 11 enrollment	0.365	-0.014	-0.05	0.232	0.642	0.364	-0.018	-0.053	0.218	0.645
	[0.045]**	[0.058]	[0.080]	[0.066]**	[0.148]**	[0.043]**	[0.060]	[0.081]	[0.062]**	[0.149]**
log grade12 enrollment	-	0.933	0.933	0.766	0.147	-	0.933	0.931	0.767	0.155
	-	[0.111]**	[0.151]**	[0.085]**	[0.185]	-	[0.111]**	[0.149]**	[0.083]**	[0.185]
Constant	2.666	-1.8	-4.222	30.036	8.308	2.93	-1.494	-3.889	30.487	9.536
	[4.125]	[3.839]	[5.158]	[14.792]*	[9.196]	[4.183]	[3.522]	[5.192]	[14.382]*	[8.743]
Year Fixed Effects?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
School Fixed Effects?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time Quartile Fixed Effects?	YES	YES	YES	YES	YES	NO	NO	NO	NO	NO
Control group trend?	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES
Observations	1999	1999	1999	1999	556	1999	1999	1999	1999	556
Number of schools	190	190	190	190	188	190	190	190	190	188
R-squared	0.4	0.74	0.54	0.38	0.36	0.45	0.77	0.59	0.41	0.44

Robust standard errors in brackets. Standard errors clustered at school district level.

* significant at 5%; ** significant at 1%

Table A3

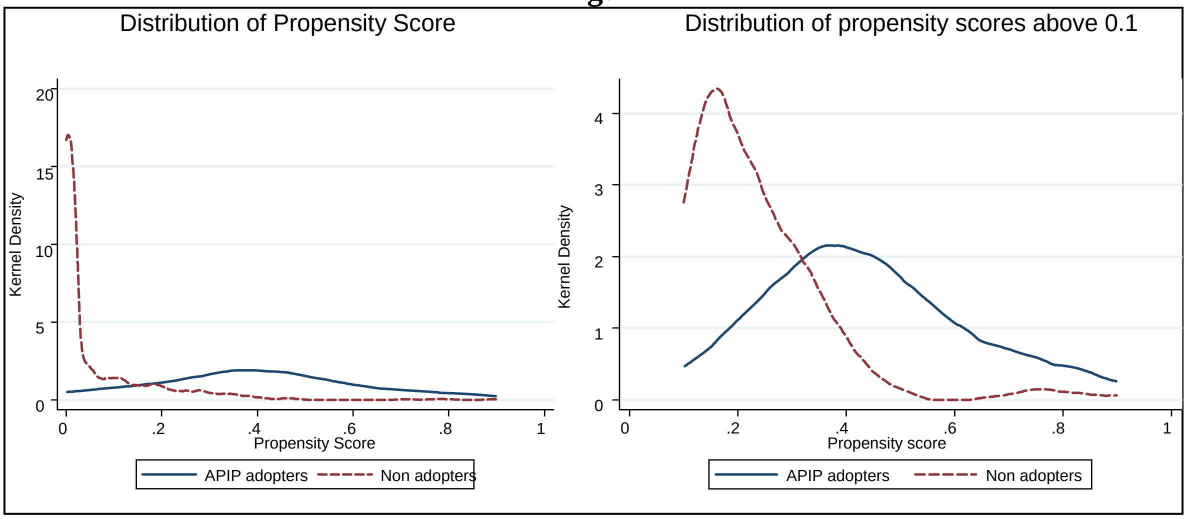
Difference-in-Differences Regressions sample of schools who will have adopted APIP by 2008

Treated observations used:	Using all adopters					Using only schools that adopted after 2000				
	1	2	3	4	5	6	7	9	9	10
	log grade 12 enrollment	log graduates	log take SAT/ACT	log above criterion	log college	log grade 12 enrollment	log graduates	log take SAT/ACT	log above criterion	log college
treat	**0.057**	0.029	0.004	**0.223**	**0.067**	0.042	-0.005	-0.061	**0.296**	**0.09**
	[0.018]**	[0.021]	[0.031]	**[0.034]****	**[0.027]***	[0.026]	[0.024]	[0.058]	**[0.072]****	**[0.046]+**
mobility	-0.028	0.035	0.075	0.037	0.087	-0.005	-0.009	0.075	-0.03	0.064
	[0.019]	[0.029]	[0.033]*	[0.086]	[0.131]	[0.035]	[0.039]	[0.048]	[0.206]	[0.189]
% white	0.083	0.033	0.043	0.142	-0.143	0.177	0.082	0.115	0.446	-0.153
	[0.104]	[0.055]	[0.068]	[0.137]	[0.123]	[0.044]**	[0.034]*	[0.119]	[0.146]**	[0.153]
% black	0.07	0.026	0.031	0.13	-0.145	0.166	0.079	0.105	0.427	-0.154
	[0.103]	[0.055]	[0.067]	[0.129]	[0.123]	[0.043]**	[0.034]*	[0.119]	[0.144]*	[0.153]
% Hispanic	0.089	0.033	0.03	0.104	-0.149	0.178	0.079	0.1	0.403	-0.158
	[0.101]	[0.056]	[0.070]	[0.127]	[0.124]	[0.042]**	[0.035]*	[0.121]	[0.150]*	[0.154]
% Econ. Disadvantaged	-0.001	0.004	0.008	-0.001	-0.003	-0.003	0.001	0.004	-0.003	-0.003
	[0.003]	[0.001]**	[0.002]**	[0.002]	[0.002]	[0.001]*	[0.000]	[0.002]	[0.002]	[0.002]
% Limited English Proficiency	-0.001	0.0004	-0.004	-0.004	0.003	-0.002	0.0003	-0.005	-0.006	0.004
	[0.001]	[0.001]	[0.001]**	[0.002]	[0.002]	[0.001]	[0.001]	[0.003]	[0.004]	[0.004]
lag of grade 11 enrollment	0.495	-0.162	-0.187	0.475	0.423	0.486	0.025	0.114	0.69	0.438
	[0.045]**	[0.102]	[0.152]	[0.127]**	[0.106]**	[0.047]**	[0.060]	[0.106]	[0.163]**	[0.106]
log grade12 enrollment	-	1.166	1.313	0.717	0.544		0.767	0.646	0.432	0.534
	-	[0.188]**	[0.281]**	[0.041]**	[0.165]**		[0.119]**	[0.159]**	[0.158]*	[0.201]
Constant	-2.533	-4.237	-6.171	-13.955	13.275	-11.715	-6.75	-9.587	-42.309	13.955
	[10.180]	[6.433]	[6.070]	[13.080]	[12.235]	[4.276]*	[3.726]	[11.220]	[14.315]*	[15.190]
Year*Decile Fixed Effects?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
School Fixed Effects?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	578	578	578	578	226	483	483	483	483	190
Number of schools	57	57	57	57	57	48	48	48	48	48
R-squared	0.48	0.81	0.61	0.39	0.55	0.55	0.79	0.51	0.43	0.56

Robust standard errors in brackets, Standard errors clustered at the school district level.

. + significant at 10% * significant at 5%; ** significant at 1%

Figure A1



Note A1

The Texas 10 percent rule was put in place in 1997 and it ensured that the top 10 percent of students from each high school would be guaranteed admission to a Texas public University. One would expect college matriculation to increase in school that have on average low achievement, such as the selected APIP schools, even if these schools did not adopt the APIP.

The Texas statewide Advanced Placement Incentive Program was introduced in 1999-2000. Under the statewide program, the state appropriated $21 million over the years 1998-2000 for the Texas AP incentive program, up from $3 million the previous biennium. The statewide program provides a $30 reduction in exam fees for all public school students who are approved to take the AP exams, teacher training grants up to $450, up to $3000 equipment and material grants for AP classes, and financial incentives to the school of up to $100 for each student who scores three or better on any AP exam. One would expect this policy to increase AP participation, and effort even if the APIP were not adopted by the APIP schools. [Source: Texas Education Agency Press Release: "Number of Advanced Placement exams taken by Texas students increases dramatically". August 23, 2000. http://www.tes.state.tx.us/press/pr000823.htm]

Note A2

In Texas, as of the 1998-9 academic year, students are only required to take 3 credits (often over 3 years starting in 9th grade) of mathematics and science and four credits in English to satisfy their high school graduation requirements. Therefore students, who have taken these courses by 10th or 11th grade, would either have a free period be taking some less rigorous elective, or taking a dual enrollment course at a college. If schools do not offer AP mathematics, students who have fulfilled the graduation requirements would either be involved in a dual enrollment course or take no math class at all. Student who had completed the science requirements would be involved in a dual enrollment science class, take a less rigorous science elective such as geology, anatomy, physiology, or have a free period. Those students who take AP English courses would be doing so *in lieu* of the standard high school English courses, or a dual enrollment college course. [Sources: Walter Dewar, Executive Vice President AP Strategies, and counselor s at several Dallas high schools]

Note A3

All schools for which the variables predict non-treatment perfectly are dropped from the estimation, and therefore eliminated from the subsequent analysis. By removing these de-facto bad matches the sample of estimation is reduced by 40%. For example, because no schools located in a small town or large town were ever treated or selected to be treated all schools located in a small town or large town are automatically removed from the sample. The estimates of this probit regression can be found in the Table 5. The estimated scores indicate that, conditional on being in the estimation sample, having large shares of limited English proficiency students in a large city or urban fringe of a large city is associated with program adoption. They also suggest that large schools *per se* is not associated with treatment but rather being a large school in an urban fringe as evidenced by the interaction between 12th grade enrollment and the urbanicity variables. The ethnicity variable suggest that while the treatment schools in large cities have large black and Hispanic enrollment shares, the treatment schools in urban fringes or rural areas have lower black and Hispanic enrollment shares. The estimates suggest that while on can predict APIP adoption, there may be a few different criteria used to select school rather than a single school profile.

One can see that there are several schools with very low estimated propensity scores in the prospective control group and a few with low propensities in the treated group. There are three outlying APIP schools with the three lowest propensities of 0.002 , 0.02 and 0.05. All other treated schools have estimated propensities at or above 0.1. Because the average estimated propensity for treated schools is 0.48, I remove the control schools with estimated propensities too low to be good candidate comparison schools. I use 0.1, (and .04) as the cut-offs for the control schools, and use only those control schools with an estimated propensity greater than or equal to 0.1 (and 0.04). By restricting the control schools to those that are similar to the treated schools in their likelihood of adopting the APIP, the assumption that the responses to other covariates and that the trend for the non-APIP schools is a reasonable predictor for the trend the APIP schools would have experience prior to adopting the program is much more plausible.

In addition to restricting the sample, one can also allow for differential time effects for schools that have different likelihoods of receiving treatment. By doing this one is allowing for the control schools that are similar to the treatment schools to have the same time effect. This is a more flexible and complete way of controlling for the fact that the schools that are most likely to have been treatment schools have a different time trajectory than schools that are less likely to have been chosen to be APIP schools. Restricting the sample and controlling for differential time effects, the regression below is estimated on the sample of APIP schools and the non-APIP schools with an estimated likelihood of treatment greater than 10% and 4%.

$$Y_{it} = \alpha + \beta \cdot Treat_{it} + \eta I_i^{APIP} \cdot t + X_{it}\psi + \sum_q \tau_{qt} \cdot Q_{iq} + \theta_i + \varepsilon_{it}$$

The variables are defined as before but now Q_{iq} is a dummy variable denoting propensity score deciles, such that $Q_{iq}=1$ if school i has a propensity score in decile q and zero otherwise. Therefore, τ_{qt} is a decile year effect, or the year effect for schools in decile q.

Note that there are only two outlying treatment schools that have an estimated propensity below this 0.04, and including these schools does not have any discernable effect on the results. Keeping all the schools with propensity scores above the minimum of the treated schools leaves in several poor matches for the vast majority of APIP schools. In any case any reasonable cut-off yields the same results. The results are summarized in appendix Table A5.

Table A5

Difference-in-Differences Regressions on Sub sample based on Propensity score

	Sample: APIP school & schools with propensity>=0.1					Sample: APIP school & schools with propensity>=0.04				
	1	2	3	4	5	6	7	8	9	10
	log grade 12	log Graduates	log SAT/ACT takers	log Above Criterion	log Attend College	log grade 12	log Graduates	log SAT/ACT takers	log Above Criterion	log Attend College
treat	**0.064**	0.028	0.016	**0.182**	**0.073**	**0.057**	0.025	0.015	**0.169**	**0.061**
	[0.024]*	[0.017]	[0.039]	**[0.042]****	**[0.028]***	**[0.022]****	[0.017]	[0.040]	**[0.043]****	**[0.026]***
year*APIP school	0.0004	0.0004	0.003	-0.014	0.008	0.001	0.001	0.003	-0.014	0.009
	[0.005]	[0.003]	[0.008]	[0.013]	[0.013]	[0.005]	[0.003]	[0.006]	[0.012]	[0.011]
mobility	-0.061	0.015	-0.019	-0.043	-0.026	-0.067	0.014	-0.016	-0.041	-0.014
	[0.039]	[0.012]	[0.024]	[0.031]	[0.065]	[0.020]**	[0.007]	[0.018]	[0.025]	[0.021]
% white	-0.016	0.063	0.115	-0.154	-0.156	-0.017	0.033	0.022	-0.072	-0.011
	[0.045]	[0.031]*	[0.040]**	[0.154]	[0.062]*	[0.022]	[0.017]*	[0.031]	[0.085]	[0.039]
% black	-0.023	0.058	0.103	-0.167	-0.16	-0.02	0.03	0.012	-0.083	-0.015
	[0.046]	[0.030]	[0.041]*	[0.155]	[0.062]*	[0.022]	[0.016]	[0.031]	[0.085]	[0.040]
% Hispanic	-0.018	0.063	0.109	-0.182	-0.16	-0.018	0.034	0.016	-0.094	-0.015
	[0.045]	[0.031]*	[0.039]**	[0.155]	[0.062]*	[0.022]	[0.017]*	[0.031]	[0.085]	[0.039]
% Econ. Disadvantaged	0.0001	0.001	0.001	-0.001	-0.001	0.00009	0.0003	0.001	0.0005	-0.001
	[0.001]	[0.001]	[0.001]*	[0.002]	[0.001]	[0.001]	[0.000]	[0.000]	[0.001]	[0.001]
% Limited English Proficiency	0.0001	0.0002	-0.004	-0.003	0.003	-0.001	0.0002	-0.002	-0.002	0.002
	[0.001]	[0.001]	[0.001]**	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]
lag of grade 11 enrollment	0.365	0.022	0.025	0.103	0.487	0.514	0.059	0.071	0.131	0.294
	[0.047]**	[0.039]	[0.063]	[0.054]	[0.100]**	[0.033]**	[0.038]	[0.056]	[0.047]**	[0.073]**
log grade12 enrollment	-	0.869	0.922	0.771	0.402	-	0.843	0.823	0.729	0.609
	-	[0.098]**	[0.116]**	[0.091]**	[0.132]**	-	[0.060]**	[0.077]**	[0.055]**	[0.084]**
Constant	6.15	-5.639	-10.928	15.247	15.54	4.83	-2.829	-1.625	6.585	1.095
	[4.592]	[3.186]	[3.965]**	[15.721]	[6.366]*	[2.201]*	[1.705]	[3.135]	[8.612]	[3.966]
Year*Decile Fixed Effects?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
School Fixed Effects?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1878	1878	1878	1878	733	5678	5678	5678	5678	2616
Number of schools	190	190	190	190	187	711	711	711	711	694
R-squared	0.38	0.77	0.58	0.32	0.41	0.49	0.79	0.51	0.22	0.44

Robust standard errors in brackets. Standard errors clustered at school district level.