A Little Now for a Lot Later

A Look at a Texas Advanced Placement Incentive Program

C. Kirabo Jackson

ABSTRACT

I analyze a program implemented in Texas schools serving underprivileged populations that pays both students and teachers for passing grades on Advanced Placement (AP) examinations. Using a difference-in-differences strategy, I find that program adoption is associated with increased AP course and exam taking, increases in the number of students with high SAT/ACT scores, and increases in college matriculation. The rewards don't appear to distort behaviors in undesirable ways, and I present evidence that teachers and students were not simply maximizing rewards. Guidance counselors credit the improvements to greater AP access, changes in social norms towards APs, and better student information.

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 has been expanded to over 40 schools in Texas and is targeted primarily to low-income, minority-majority school districts with a view toward improving college readiness. Across the United States, college matriculation and completion rates for low-income and mi-

C. Kirabo Jackson is an assistant professor of labor economics at the ILR School at Cornell University. *The author thanks Roland Fryer, Caroline Hoxby, and Lawrence Katz for their guidance and gratefully acknowledges helpful comments from two anonymous referees, Elias Bruegmann, Claudia Goldin, Li Han, Clement Jackson, Dale Jorgenson, Gauri Kartini Shastry, Katharine Emans Sims, Erin Strumpf, and Daniel Tortorice. The author thanks Walter Dewar and Gregg Fleisher of AP Strategies, and Nina Taylor, Perry Weirich, and Shawn Thomas of the Texas Education Agency. This article benefitted from conversations with Elizabeth Chesler, Paula Jackson, June Jackson, Derek Janssen, and Michael Jeffries. The usual disclaimer applies. The data used in this article can be obtained beginning six months after publication through three years hence from C. Kirabo Jackson, ILR School, 345 Ives Hall East, Cornell University, Ithaca, NY 14853 ckj5@cornell.edu. This article incorporates material previously circulated in Jackson (2007) and Jackson (2008).*
[Submitted July 2008; accepted February 2009]
ISSN 022-166X E-ISSN 1548-8004 © 2010 by the Board of Regents of the University of Wisconsin System

Name /jhr453/JHR45_03_A03/Mp_592 04/21/2010 09:00AM Plate # 0 pg 592 # 2

nority students are much lower than those for nonpoor whites.[1] These differences may be due, in part, to low-income and minority students having lower participation rates in advanced courses such as APs (Klopfenstein 2004) and being less college ready as a result. These lower participation levels may reflect inefficiencies that can be ameliorated by the APIP, such as imperfect student information, student myopia, lack of student readiness, suboptimal teacher effort, or suboptimal track placement.[2]

At the first 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. By 2002, those schools had 132 passes per 1,000 juniors and seniors taking math, science, and English, compared to 86 in Texas and 80 in the United States (AP Strategies). Due to the perceived success of this program, New Mexico and New York City have adopted similar programs while schools in Arkansas, Alabama, Connecticut, Kentucky, Massachusetts, Virginia, and Washington have received grants to replicate the APIP.[3] Cash incentive programs for students and teachers are growing in popularity and increasing in importance, but they are relatively new and are therefore not adequately studied nor understood. I aim to provide much needed evidence on the efficacy of such programs by evaluating one of the earliest cash incentive programs in the United States. Specifically, I answer the following questions: (1) Does the APIP increase AP course enrollment and exam taking? (2) Do the APIP incentives distort behaviors in undesirable ways? (3) Does the APIP affect longer-term outcomes such as college enrollment? (4) What are the mechanisms through which the APIP may operate?

The effect of the APIP reflects a combination of the effect of increased AP course taking and other aspects of the program such as teacher training and student and teacher incentives. The rationale behind the push to increase AP participation[4] is the observation that students who take rigorous courses and exams in high school, such as APs, have higher SAT scores (College Board 2003) and are more likely to enroll and be successful in college than their peers, as measured by GPA and graduation rates (Dodd et al. 2002; Dougherty et al. 2006; Eimers 2003; Geiser and Santelices 2004; Morgan and Ramist 1998). Because observationally similar students who take AP courses may differ in *unobservable* dimensions, such as motivation, and all these studies rely on observational data, it is unclear how much one can attribute to AP participation. The APIP, however, pays students and teachers for passing AP exam scores—producing exogenous variation in AP taking that is unrelated to students'

1. According to the August 2006 Current Population Survey, the percentage of white high school graduates or GED holders between the ages of 25 and 29 ever enrolled in some college program was 71 for whites, 60 for blacks, and 52 for Hispanics. The implied two- or four-year college completion percentages are 68 for whites, 51 for blacks, and 53 for Hispanics.

2. In Card (1995), educational attainment is the result of a lifetime utility maximization problem based on available information. In this framework, new information could change students' educational choices. See Frederick, Loewenstein, and O'Donoghue (2002) for a discussion of time discounting and time preference.

3. See Lyon (2007), Mathews (2007), and Medina (2007), www.nationalmathandscience.org.

4. In his 2004 State of the Union Address, President Bush announced a plan under the No Child Left Behind Act to support state and local efforts to increase access to AP courses and exams (http://www.whitehouse.gov/). Several states have programs with the same objective. For example, the Western Consortium for Accelerated Learning Opportunities consisting of Arizona, Colorado, Hawaii, Idaho, Montana, New Mexico, Oregon, South Dakota, and Utah.

Name /jhr453/JHR45_03_A03/Mp_593 04/21/2010 09:00AM Plate # 0 pg 593 # 3

intrinsic motivation. As such, improvements in secondary and postsecondary outcomes associated with APIP adoption may reflect, in part, the causal effect of taking AP courses.

The second important aspect of the APIP is the incentives for students and teachers. Research indicates that students and teachers respond to incentives in rewarded tasks.[5] However, some studies suggest that providing incentives to students without the cooperation of teachers may be ineffective (Angrist and Lavy 2002; Kremer, Miguel, and Thornton 2004). 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 others. 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 when they are rewarded for having high grade-point averages (Binder, Ganderton, and Hutchens 2002; Cornwell et al. 2005; Glewwe, Ilias, and Kremer 2003;). While the APIP may improve AP outcomes, the incentives may lead teachers and students to divert resources away from non-AP activities in ways that could hurt non-AP students and undermine overall student achievement. This underscores the importance of looking at a broad range of AP *and* non-AP outcomes while also looking at longer-term outcomes such as college enrollment. The fact that improved AP outcomes could simply reflect increased test-taking effort as opposed to increased knowledge is further reason to look at a broad range of outcomes.

Using aggregate school-level data, I identify the program effect by comparing the change in outcomes of cohorts within the same school, before and after APIP adoption, to the change in outcomes for cohorts in comparison schools over the same time period. By comparing cohorts within the same school, I eliminate self-selection within a cohort—the self-selection that ordinarily makes one student enroll in AP courses and another not to do so. Since the program 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 comparison group those schools that had already decided to adopt the APIP but had not yet had the opportunity to implement it. My comparison group also helps me account for potentially confounding statewide policies.[6] Because program adoption was not *random*, I present falsification tests indicating that the timing of adoption was likely *exogenous*. My identification strategy is valid so long as schools have the same incoming distribution of students in their preprogram and postprogram cohorts. I test this important restriction empirically and can rule out most plausible scenarios of changing student selection into program schools.

5. This list includes Angrist, Bettinger, Bloom, King, and Kremer (2002); Angrist, Lang, and Oreopoulos (2009); Angrist and Lavy (2002, 2007); Atkinson et al. (2004); Eberts, Hollenbeck, and Stone (2000); Kremer, Miguel, and Thornton (2004); and Lavy (2002, 2004). Some psychologists have documented 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 (1999) for an overview of this research.

6. For a description of such policies see the appendix, note A1.

Name /jhr453/JHR45_03_A03/Mp_594 04/21/2010 09:00AM Plate # 0 pg 594 # 4

I find that cohorts in schools affected by the program had more students taking AP or International Baccalaureate (IB) exams by the first year of APIP adoption and increases in AP course enrollment by the third year of adoption. Looking to a broader range of outcomes, by the second year of adoption affected cohorts had more students with high SAT/ACT scores and more students who matriculated at a college in Texas. While there are no differences by gender, some specifications suggest that the relative improvements may be largest for minority students. I find little evidence that APIP schools diverted resources away from non-AP students or activities. I present several pieces of evidence that suggest that the response was not simply due to students and teachers maximizing their cash rewards. This evidence is corroborated by guidance counselors' claims that the increased AP participation was due to increased encouragement from teachers, better student information, and changes in teacher and peer norms. The findings contribute to the education incentives literature by presenting some of the first evidence that a well-designed cash incentive program for students and teachers can improve short-term and longer-term outcomes.[7] The results also suggest that increasing participation and effort in rigorous courses such as APs can be effective in improving overall student outcomes.

The remainder of this paper is as follows: Section II describes the APIP; Section III describes the data used; Section IV lays out a theoretical framework within which to think about the APIP effect on AP participation and the other outcomes; Section V motivates and describes the empirical strategy; Section VI analyzes the results, and Section VII provides conclusions.

II. Description of the Advanced Placement 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.[8] 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 of 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 (College Board). 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 5 being the highest and 3 generally regarded as a passing grade. AP courses are taught during regular class time and generally substitute for another course in the same

7. Angrist and Lavy (2007) found that cash incentives improved educational outcomes for girls (including postsecondary enrollment). This is the only other study of a cash incentive program for high school students.
8. While this is true in general, some highly selective colleges only allow students to use AP credits to pass out of prerequisites, but not toward regular graduation credit.

Name /jhr453/JHR45_03_A03/Mp_595 04/21/2010 09:00AM Plate # 0 pg 595 # 5

subject (AP chemistry instead of 11th grade science, for example), for another elective course, or for a free period. While AP courses count toward 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.[9]

The APIP is run by AP Strategies, a nonprofit organization based in Dallas, and 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. It 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 teaches students and trains other AP teachers. Vertical teams also include teachers who teach grades that precede those in which AP courses are offered. For example, a vertical team might create a math curriculum starting in 7th grade designed to prepare students for AP calculus in 12th grade. 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).

The APIP's monetary incentives are intended to encourage participation and induce effort in AP courses. Lead teachers receive between $3,000 and $10,000 as an annual salary bonus, and a further $2,000 to $5,000 bonus opportunity that is based on results. Pre-AP teachers receive an annual supplement of between $500 and $1,000 per year for extra work. AP teachers receive 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 discretionary bonuses of up to $1,000 based on results. While the amount paid per passing AP score and the salary supplements are well defined in each school, there is variation across schools. Overall, the APIP can deliver a considerable increase in compensation for teachers.[10]

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 a free or reduced lunch program would pay $15 (instead of $30), while those who are not would pay $30 (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 take 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. A student who passes several AP examinations during 11th and 12th grades can earn several hundred dollars. For example, one student earned $700 in his junior and senior years for passing scores in AP examinations (Mathews 2004). Since students must attend the AP courses *and* pass the AP exams to receive the rewards, students who do not take the AP courses would not take the exams in an attempt to earn the cash rewards.

9. Discussion with the executive vice president of AP Strategies and with counselors at several Dallas schools, 2006. For a description of the high school graduation requirements see appendix, note A2.

10. 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, when 49 of her 110 students received a 3 or higher, she earned $11,550 for participating in the program; this was a substantial increase in annual earnings (Mathews 2004).

Name /jhr453/JHR45_03_A03/Mp_596 04/21/2010 09:00AM Plate # 0 pg 596 # 6

This aspect of the incentives makes them relatively difficult to game and likely to increase overall student learning.

As a general rule, adoption of the APIP works as follows. First, schools that are interested in implementing the APIP approach AP Strategies and are put on a list.[11] AP Strategies then tries to match interested schools to a donor. When a private donor approaches AP Strategies, he or she selects which schools to fund from within the group of willing schools. In most cases the donor wants schools within a specific district.[12] Once a willing school has been accepted by the donor, preparations are made (such as training and creation of curricula) and the program is implemented the following calendar year.[13] It takes about two years to fully implement the APIP after a school expresses interest.

The donors choose the subjects that will be rewarded and ultimately determine the size of the financial rewards. While there are differences across schools, most schools reward English, mathematics, and sciences. There is variation in the timing across schools of when the program is introduced, which I exploit to identify the effect of the program. As illustrated in Figure 1, 41 schools adopted the APIP between 1995 and 2005 and 61 schools would have adopted the program by 2008, so the number of treated units is relatively small.[14] Since donors chose the schools, donor availability and donor preferences are the primary reasons for variation in the timing of program implementation. To quote the 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." Since several districts compete for the same donor, donor preferences determine the districts or the schools within a district that will adopt the program in any given year.[15] I argue that the exact timing of which schools adopt the program, within the group of willing schools, is orthogonal to *changes* in unobserved school characteristics. I test this assumption empirically in Section VI.

The total 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 from $100 to $300. Private donors pay for 60 to 75 percent of the total costs of the program, and the district covers the remainder. Districts typically pay for teacher training and corre-

11. There are a few exceptions. Schools in Austin were approached by the donor to adopt the APIP in 2007. Also, five schools in Dallas secured a donor before approaching AP Strategies.

12. For example: The first ten Dallas schools were chosen based on proximity to AP Strategies; ST Microelectronics is located in the Carrolton-Farmers community and funded that district's schools; the Priddy Foundation specifically requested the Burkburnett and City View schools; anonymous donors specifically requested Amarillo and Pflugerville schools; the Dell Foundation (headquartered in Austin) funds the Austin and Houston programs; and the remaining Dallas schools were funded by the O'Donell Foundation to complete the funding of Dallas Independent School District (ISD).

13. The seven schools to 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 rewards were provided.

14. A post hoc power analysis reveals that after removing school and year means, the residual variance is sufficiently small so that 30 schools (with both pre- and post-treatment data) are sufficient to detect effects similar to those found with more than 0.80 power. See appendix note A4 for a more detailed discussion of this.

15. For example, in 2005 four high schools were chosen by The Michael and Susan Dell Foundation from a list of seven willing Houston schools. The remaining three schools may adopt the program at a later date.

Name /jhr453/JHR45_03_A03/Mp_597 04/21/2010 09:00AM Plate # 0 pg 597 # 7

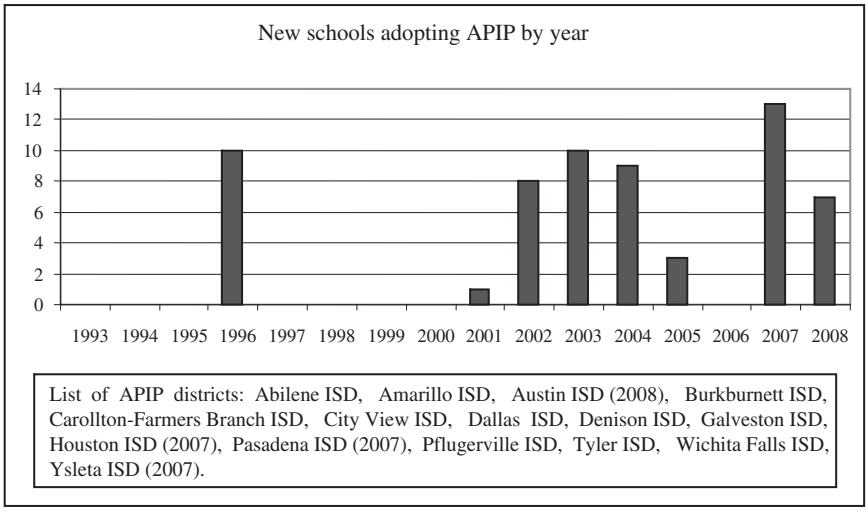

Figure 1
Number of Schools Adopting the APIP by Year

sponding travel, release time, and some of the supplies and equipment costs. The donors fund the cash rewards to students and teachers, 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, districts may be able to fund their 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.

III. Data

The data on school demographics, high school graduation rates, and college entrance examinations come from the Texas Education Agency's Academic Educational Indicator System (AEIS). I use these high-school-level data for the academic years 1994 (1993–2004) through 2005 (2004–2005). All years refer to academic years. Urbanicity data come from the Common Core. College enrollment data for 2002 through 2005 come from the Texas Higher Education Coordinating Board. The final data set combines these publicly available data with a list of program schools obtained from AP Strategies. Forty schools adopted the APIP in sample, while 59 were scheduled to have adopted the program by 2008. Since these data were created for state accountability purposes, several statistics are not in their ideal form (for example, the percentage of high school graduates in a given year who have taken the SAT/ACT exams as opposed to the number of students taking the SAT/ACT exams in a given year). Wherever possible, I have computed other

Name /jhr453/JHR45_03_A03/Mp_598 04/21/2010 09:00AM Plate # 0 pg 598 # 8

variables that are easier to interpret in a regression setting. In some cases, such computations were not possible and the reported statistics are used as is.

Summary statistics in Table 1 show sample means of the variables for 1993 through 1999 and 2000 through 2005 for the APIP schools that adopted the program in sample (APIP schools), schools that were selected to adopt the program but had not yet done so in sample (future APIP schools), and schools that were never selected to adopt the APIP by 2008 (non-APIP schools). Restricting the data to high schools with 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 in a particular year, and all the variables are defined for a particular academic year.[16] Since all schools adopt the APIP at the beginning of the academic year, I define a school to be treated as a school that will have been exposed to the APIP for at least one full academic year. For example, while the first schools to adopt the APIP did so in the 1995 *calendar* year, they are coded as first being treated in the 1996 (1995–96) *academic* year.

Schools that were selected for the APIP were quite different from schools that have not yet been selected and may never be selected for the APIP. The APIP schools and future APIP schools had average total enrollments in 2000 through 2005 of 1,731 and 2,068 respectively—much larger than the average enrollment of 715 students for non-APIP schools. Between 2000 and 2005, 77 percent of the APIP schools and 67 percent of the future APIP schools were in a large or mid-sized city compared to only 20 percent for non-APIP schools.[17] During the same time period, only 29 percent of the students at APIP schools and 14 percent of students at future APIP schools were white compared to 53 percent for non-APIP schools. However, the minority students in APIP schools were largely black, while Hispanics made up the majority of students at future APIP schools. Also between 2000 and 2005, about ten percent of students were limited English proficient at APIP and future APIP schools compared to less than 4 percent at non-APIP schools.

For most outcomes, the APIP schools and future APIP schools performed worse than other Texas high schools. The variable "graduates" is the total number of high school graduates completing at least the minimum high school program.[18] The ratio of 10th graders to graduates (two years later) was much lower in APIP and future APIP schools (about 0.75) than in non-APIP schools (about 0.84). The ratio of AP course enrollment (which counts each student in each AP course) to school enrollment between 2000 and 2005 was 0.23 for APIP schools and 0.2 for future APIP schools, as compared to 0.45 for non-APIP schools. Despite this, the percentage of 11th and 12th graders taking at least one AP or IB exam was 20 percent in APIP schools, 15 percent in future APIP schools, and only 9.62 percent in non-APIP schools. This suggests that a greater proportion of 11th and 12th graders at APIP schools took the examination after having taken the course, or that in non-APIP schools many AP course enrollees were in 9th or 10th grade. For all schools, white

16. Since some variables are not available for all years, and some schools did not exist during all years, sample sizes and composition may vary slightly over time.

17. Due to changes in neighborhoods, the urbanicity variables *do* change within schools over time.

18. This includes special education students completing an Independent Education Plan. Note: Students must pass the exit-level TAAS exam typically taken in 10th grade to graduate. Since the APIP affects 11th and 12th graders, I do not look at this outcome.

Table 1
Summary Statistics

Summary statistics for demographics for APIP schools and other comparison groups						
	APIP Schools		Future APIP Schools		Non-APIP Schools	
	1993–1999	2000–2005	1993–1999	2000–2005	1993–1999	2000–2005
Treat	0.15	0.74	0.00	0.00	0.00	0.00
	(0.36)	(0.44)	(0.0)	(0.0)	(0.0)	(0.0)
Enrollment	1,629.02	1,731.97	2,108.04	2,068.36	716.85	751.56
	(649.81)	(710.62)	(625.75)	(511.64)	(781.97)	(833.36)
Mobility	479.30	469.03	592.63	568.74	178.19	180.73
	(184.28)	(187.08)	(205.23)	(155.26)	(199.44)	(197.25)
Percent white	32.67	28.71	26.69	17.30	59.38	53.36
	(30.31)	(27.93)	(14.59)	(12.95)	(29.46)	(30.42)
Percent black	36.68	31.98	15.69	13.50	10.32	11.30
	(31.86)	(28.49)	(15.6)	(12.43)	(15.64)	(17.08)
Percent Hispanic	27.68	36.04	53.71	66.05	28.92	33.67
	(24.46)	(26.02)	(21.35)	(18.98)	(28.9)	(29.5)
Percent Asian	2.58	2.34	3.69	2.51	1.09	1.12
	(3.41)	(3.96)	(3.45)	(2.96)	(2.76)	(2.98)
Percent free lunch	34.19	38.68	34.62	48.08	30.42	35.51
	(22.97)	(23.9)	(20.8)	(27.44)	(23.97)	(26.25)
Percent limited English	9.50	10.16	10.01	11.84	3.57	3.83
	(14.33)	(12.61)	(9.68)	(10.21)	(7.71)	(6.8)
Grade 12	280.96	315.48	374.02	391.47	139.45	156.89

(*continued*)

Name /jhr453/JHR45_03_A03/Mp_600 04/21/2010 09:00AM Plate # 0 pg 600 # 10

Table 1 (*continued*)

Summary statistics for demographics for APIP schools and other comparison groups						
	APIP Schools		Future APIP Schools		Non-APIP Schools	
	1993–1999	2000–2005	1993–1999	2000–2005	1993–1999	2000–2005
	(125.96)	(147.63)	(126.18)	(125.88)	(166.03)	(186.44)
Lag Grade 11	329.44	359.90	439.05	450.64	156.39	172.80
	(142.7)	(156.15)	(152.53)	(151.79)	(187.78)	(206.57)
2nd lag Grade 10	397.85	426.05	519.11	539.03	177.57	187.65
	(170.2)	(182.02)	(170.5)	(140.23)	(211.62)	(222.18)
City	0.902	0.769	0.813	0.672	0.182	0.197
	(0.23)	(0.42)	(0.39)	(0.47)	(0.39)	(0.4)
Rural	0.00	0.013	0.00	0.025	0.489	0.373
	(0.0)	(0.11)	(0.0)	(0.16)	(0.5)	(0.48)
High school graduates	267.00	321.09	335.94	378.75	132.25	156.08
	(126.6)	(148.86)	(112.83)	(123.29)	(156.32)	(183.87)
AP course Enrollment	389.08	414.68	412.84	405.55	270.03	339.58
	(218.79)	(252.67)	(240.34)	(264.61)	(318.23)	(353.11)
Percent 11 and 12th graders taking AP/IB (all)	–	20.42	–	15.07	–	9.62
	–	(8.98)	–	(10.32)	–	(11.62)
Percent 11 and 12th graders taking AP/IB (black)	–	12.90	–	9.09	–	6.38
	–	(7.72)	–	(7.65)	–	(10.27)
Percent 11 and 12th graders taking AP/IB (Hispanic)	–	14.19	–	12.50	–	7.68
	–	(9.28)	–	(8.23)	–	(10.24)
Percent 11 and 12th graders taking AP/IB (white)	–	29.60	–	23.88	–	12.58
	–	(15.54)	–	(20.08)	–	(14.78)

Graduates taking SAT/ACT	143.02	160.03	184.67	208.73	76.05	82.49
	(92.11)	(92.27)	(79.29)	(114.72)	(106.82)	(120.35)
Graduates scoring above 1100/24 on SAT/ACT	31.72	36.77	33.18	33.46	20.78	27.52
	(39.39)	(39.4)	(26.84)	(32.1)	(43.35)	(54.61)
College enrollees (any college in Texas)	–	131.84	–	151.89	–	94.48
	–	(89.91)	–	(83.78)	–	(108.44)
Percent graduates above 1100/24 on SAT/ACT (all)	16.44	18.25	18.05	15.81	18.04	20.70
	(13.94)	(15.33)	(13.44)	(13.88)	(12.8)	(13.53)
Percent graduates above 1100/24 on SAT/ACT (black)	5.28	5.80	7.13	6.00	6.48	7.68
	(6.25)	(6.67)	(7.27)	(6.39)	(9.6)	(9.92)
Percent graduates above 1100/24 on SAT/ACT (white)	32.65	37.76	29.16	31.34	23.74	27.78
	(15.51)	(13.04)	(19.9)	(22.39)	(14.44)	(14.46)
Percent graduates above 1100/24 on SAT/ACT (Hispanic)	10.78	12.33	11.38	10.34	9.54	12.46
	(11.22)	(11.66)	(14.11)	(10.18)	(11.58)	(12.77)
Percent graduates who pass TAAS/TASP	35.814	59.997	37.836	58.207	42.281	65.114
	(16.54)	(15.31)	(10.67)	(12.36)	(16.21)	(17.7)
Percent 9th–12th graders taking college courses	15.518	18.408	13.177	17.987	14.917	16.527
	(8.82)	(7.37)	(7.61)	(7.34)	(9.86)	(11.45)
Number of schools	40		18		1,413	

Notes: School demographic data, high school graduation, and SAT/ACT data are available for 1993 through 2005; AB/IB examination data are available for years 2001 through 2005; college matriculation data are available for years 2001 through 2005; and AP course enrollments are available for 2000 through 2005. As such, the number of observations varies from variable to variable. While the regressions are estimated on a balanced sample, the sample averages show all the available data. Future APIP schools are those schools that will have adopted the APIP by 2008 but had not done so by 2005.
Mobility is the number of students who were enrolled in the school for less than 83 percent of the school year. Students have attained TAAS/TAKS equivalency if they have done well enough on the exit-level TAAS to have a 75 percent likelihood of passing the Texas Academic Skills Program tests for college readiness. College is the count of students who matriculated in college the fall after high school graduation in any school in Texas.

Name /jhr453/JHR45_03_A03/Mp_602 04/21/2010 09:00AM Plate # 0 pg 602 # 12

11th and 12th graders were much more likely to have taken at least one AP or IB exam than were minorities. The variable "college enrollees" is the total number of students who graduated that academic year who enrolled in *any* college (public or private) in Texas the fall of the following academic year. Between 2000 and 2005, the ratio of college attendees to graduates was 0.41 for the APIP schools and 0.40 for future APIP schools, compared to over 0.60 for non-APIP schools. During this same time period, the ratio of high school graduates to 12th graders scoring above 1100/24 on the SAT/ACT (based on exams taken at any time) was roughly 0.12 for APIP schools and 0.09 for future APIP schools compared to over 0.17 for non-APIP schools. In all schools, white high school graduates outperformed both black and Hispanic students in the SAT/ACT examinations.[19]

IV. Theoretical Background

This section analyzes the APIP through the lens of the principal agent multitask model and the Becker-Rosen schooling decision model. While not exhaustive, these models highlight important mechanisms through which the APIP may affect AP participation, non-AP outcomes, and the decision to apply to and enroll in college.

A. The Effect of the APIP on AP Outcomes

Suppose student AP output and student non-AP output (other educational output) are both functions of student and teacher effort in AP courses and student and teacher effort in other educational tasks, respectively. Under the APIP, teacher pay is more closely tied to the AP output of their students. The gains to a student of taking and doing well on AP exams are also greater under the APIP. A principal agent multitask model (Holmstrom and Milgrom 1991) predicts that where good AP performance is more likely with higher teacher and student effort, both teachers and students will exert more effort to improve student AP output. Therefore, one would expect (a) an increase in teacher effort to recruit students to take AP courses, (b) an increase in teacher effort to improve the quality of their instruction, (c) an increase in student AP exam taking, (d) an increase in student effort to perform well in AP exams, and (e) an increase in AP course enrollment.

B. The Effect of the APIP on Non-AP Outcomes

A principal agent multitask model also predicts that students and teachers may divert effort away from other academic pursuits *in order to* increase their AP output. For example, to increase AP effort teachers may spend less time on their non-AP courses, schools may expend fewer resources to non-AP students, students may spend less time preparing for their non-AP courses and students may withdraw from other

19. The raw number of nonspecial education graduates scoring above 1100/24 on the SAT/ACT for the campus is available. However, SAT/ACT performance is presented as the percentage of nonspecial education graduates scoring above 1100/24 on the SAT/ACT broken down by ethnicity. The number of nonspecial education graduates by ethnicity is not provided.

Name /jhr453/JHR45_03_A03/Mp_603 04/21/2010 09:00AM Plate # 0 pg 603 # 13

courses in order to take AP courses. As such, if there are no complementarities between AP output and non-AP output, and if students and teachers solely maximize their rewards, the APIP *could* have deleterious effects on the academic outcomes of non-AP students and on non-AP outcomes. The lack of complementarities between AP output and non-AP output is unlikely, however, since students gain knowledge and skills through taking AP courses that they can use in other courses and in other academic spheres, and teachers may learn general teaching skills in their AP teacher training that make them better teachers in their non-AP classrooms. Since the incentive effect and the spillover/complementarity effect work in opposite directions, the effect of the APIP on non-AP outcomes, and therefore on educational outcomes overall, is theoretically ambiguous. This requires an empirical treatment, which I present in Section VI.

C. The Effect of the APIP on College Decision

To analyze how the APIP may affect the college-going decision, consider the Becker-Rosen model. The log of earnings is an increasing concave function of the years of schooling $y = e^{g(s)}$. Individuals pay a cost to attend school c_i (effort costs, tuition costs, foregone earnings), and discount the future at rate δ. Individuals choose the level of schooling s to maximize the present value of earnings minus costs. As a simplification, students choose between going to college for four years ($s = 4$) or stopping schooling after high school ($s = 0$). An individual chooses to attend college iff

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

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 after high school. It is useful to think of taking AP courses as a way to reduce college costs (increased likelihood of admission, more financial aid, tuition savings due to college credit, faster graduation, and signal to colleges about ability or motivation). Because the increased AP participation (due to the APIP) will reduce an individual's costs of attending college, for any given level of applications one would expect an increase in matriculation rates. The second way the APIP may affect the college-going decision is if students are unaware of their true ability to succeed at college. Learning their true cost c_i as a result of taking AP courses could affect their college enrollment decisions.[20] All else equal, if they learn that they are able to cope with and enjoy the material, they may be more likely to apply to college and vice versa. The direction and magnitude of this information effect depends on (1) the number of students on the margin of applying to college, (2) whether marginal students are optimistic or pessimistic about their abilities, and (3) the quality of AP instruction.

20. This idea is similar to Costrell (1993), who models the information value of matriculating in college to learn one's suitability. He argues that this could explain the low college completion rates among certain populations.

Name /jhr453/JHR45_03_A03/Mp_604 04/21/2010 09:00AM Plate # 0 pg 604 # 14

D. The Student Decision to Take AP Courses

Policies that provide monetary rewards to students for taking and doing well on AP exams presume that students are not *already* making optimal decisions. It is therefore instructive to lay out a rational choice framework of AP participation to highlight how and why the APIP might affect student decisions. Consider a college-bound student's decision to take AP courses. Taking AP courses reduces college costs from c_0 to c_1 and entails a cost τ. In the absence of the APIP, a student only takes AP courses if the present discounted value of the college cost savings is greater than the private cost of taking AP courses. Specifically iff

$$(2) \quad \tau \leq \int_0^s (c_0 - c_1) e^{-\delta t} dt \text{ where } s \in \{0,4\}$$

Under the APIP, students who take AP courses earn a reward π for passing AP exams at time *1*. Under the APIP, students take AP exams if the present value of the costs minus the rewards is less than the savings in college costs or iff

$$(3) \quad \tau \leq e^{-\delta}\pi + \int_0^s (c_0 - c_1) e^{-\delta t} dt \text{ where } s \in \{0,4\}$$

The savings associated with passing out of a semester of college and the recouped labor force earnings is about a third of one's annual salary plus half of a year's tuition costs.[21] 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. In contrast, the average value of the rewards is less than $200 per student and the test fee reduction is at most $30.[22] The key insight of this model is that irrespective of the size of the costs (effort or otherwise) to taking AP courses, the *marginal* change in lifetime benefits due to the APIP is small, so that if students were behaving optimally, there would be a small participation response. Any sizable AP participation response would require that (1) students were myopic and overvalued the rewards in the near future, (2) students were uninformed about the benefits to AP taking, or (3) students were more likely to take AP courses under the APIP for reasons unrelated to their present discounted value of income stream, such as changes in constraints, increases in the supply of AP courses, changes in peer attitudes toward AP courses, or greater teacher encouragement.[23] The testimonial evidence presented in section VI suggests Explanation 3.

21. According to the U.S. Department of Education, average annual tuition costs in Texas were $8,057 in 2007. According to the U.S. Census Bureau, in 2005 workers 18 and over with a bachelor's degree earned an average of $51,206 a year. One-third of the annual earnings plus half the tuition cost comes to $21,093.
22. For students who are on free or reduced lunch, the reduction is $15. This reduction would only be very important if students were severely credit constrained.
23. A large AP participation response could also indicate that (4) several students were bunched right on the margin at which the benefits outweigh the costs or (5) students who were not interested in going to college were taking AP courses solely for the rewards. Any smoothness assumption rules out (4). Guidance counselors claim that all AP students have college aspirations, ruling out (5).

Name /jhr453/JHR45_03_A03/Mp_605 04/21/2010 09:00AM Plate # 0 pg 605 # 15

E. The Timing of Effects

Since the cash incentives may lead to increased student interest and increased encouragement from teachers, one might expect to see improvements in many outcomes immediately. Since the APIP rewards affect students in 11th and 12th grades, however, any effects that operate through the *duration* of exposure to AP courses (such as SAT/ACT scores) should be larger in the second year than in the first. Many effects may show incremental improvements over time since the APIP also includes training for pre-AP teachers; there may be "learning-by-doing" effects; and the size of the pool of prepared prospective AP students may increase over time.

V. Identification Strategy

My basic strategy is to identify the effect of the program using a difference-in-differences (DID) methodology: Comparing the *change* in outcomes across exposed and unexposed cohorts for schools that adopted the APIP to the change in outcomes across cohorts for schools that did not adopt the APIP over the same time period. This strategy has the benefit of (1) removing any pretreatment differences that may exist between schools that decided to adopt the APIP and those that did not and (2) removing self-selection *within* a cohort by making cross-cohort comparisons (that is, I do not compare AP students to non-AP students within a cohort). This strategy relies on the assumption that the difference in outcomes across cohorts for the comparison schools is the same, in expectation, as what the adopting schools would have experienced had they not adopted the APIP. For the changes in comparison schools to be a credible counterfactual for what the APIP schools would have experienced in the absence of the APIP, the comparison schools must be similar to the APIP adopting schools in both observable and *unobservable* ways. Because APIP schools and non-APIP schools have very different observable characteristics (see Table 1), using all other Texas high schools as the comparison group would be misguided.

Due to a scarcity of donors, AP Strategies could not implement the APIP in all interested schools at the same time. This allows me to restrict the estimation sample to only those schools that either adopted the APIP or will have adopted the APIP by 2008—using the change in outcomes for APIP schools that did not yet have the opportunity to implement the program (future APIP schools) as the counterfactual change in outcomes. Table 1 shows that the APIP schools and future APIP schools are similar in most respects. One notable difference is that the APIP schools have large black enrollment shares while future APIP schools have large Hispanic enrollment shares. Such differences are not unexpected since the sample of treated schools is small and the timing of APIP adoption was not completely random. This sample restriction has two important benefits: (1) since APIP-willing schools are observationally similar, they are likely to share common time shocks, and (2) since all schools that agreed to adopt the APIP were similarly motivated and interested, restricting the sample in this way avoids comparing schools with motivated principals who want to adopt the APIP to schools with unmotivated principals who have no interest in the program. This sample restriction controls for school self-selection

Name /jhr453/JHR45_03_A03/Mp_606 04/21/2010 09:00AM Plate # 0 pg 606 # 16

on *unobserved* characteristics and allows for a consistent estimate of the average treatment effect on the treated (ATT).[24]

Within this subsample of APIP-willing schools, identification relies on the assumption that the exact timing of APIP implementation was exogenous to other within-school *changes*. Since all willing schools had to wait for a donor to adopt the APIP, and timing of *actual* adoption relies on idiosyncratic donor preferences and availability, this assumption is plausible. Readers may worry that if donors selected schools based on the enthusiasm of school principals and administrators, or if some schools expressed interest before others, then the timing of adoption may not be orthogonal to school characteristics. Since donor choices were not random, I cannot entirely rule this out. However, it is important to note that all regressions use within-school variation so that differences in time-constant school enthusiasm or motivation will not confound the results. The only problem would be if adoption were coincident with *changes* in unobserved school enthusiasm or motivation. As noted earlier, it takes about two years to implement the APIP after expressing interest. In section VI, I show that the timing of when a school is likely to expresses interest in the APIP is *not* associated with improved outcomes, while the timing of *actual* adoption is, suggesting that the assumption of exogenous timing of adoption is valid.

This difference-in-differences (DID) strategy described above is implemented by estimating equation [4] by OLS on the APIP and future APIP sample.

$$(4) \qquad Y_{it} = \alpha + \beta \cdot Treat_{it} + X'_{it}\psi + \tau_t + \theta_i + \varepsilon_{it}$$

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 the idiosyncratic error term. The coefficient on $Treat_{it}$ captures the average before/after difference in outcomes. Since outcomes may not improve immediately, the simple before/after comparison may understate the effect of APIP adoption. As such, I also estimate more flexible models that allow the effect to vary by years since implementation.

VI. Results

I present the results in four sections. In Section VIA, I document the effect of adopting the APIP on AP course enrollment, AP/IB exam taking, high school graduation, SAT/ACT performance, and college matriculation. I extend the analysis to investigate how APIP adoption affects outcomes over time. In Section VIB, I discuss various potential threats to validity and present falsification tests showing that the identification strategy is likely valid. In section VI.C, I present

24. Estimation results are very similar using all other high schools as the comparison group. Using such comparison groups, however, makes the results susceptible to school selection on unobserved characteristics.

Name /jhr453/JHR45_03_A03/Mp_607 04/21/2010 09:00AM Plate # 0 pg 607 # 17

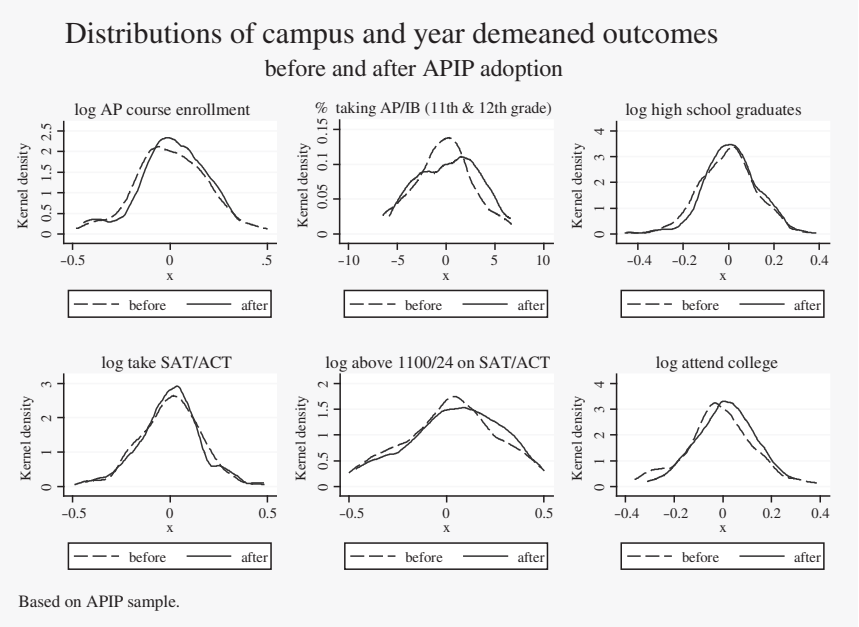

Figure 2
Distribution of Outcomes Before and After APIP Adoption

certain effects by gender and ethnicity and show that the effects of the APIP on SAT/ACT performance *cannot* be solely attributed to selective migration. Finally, in section VI.D I present evidence from discussions with guidance counselors for why and how the APIP affects student outcomes, and I present the results of various tests of the hypothesis that students and teachers responded to the incentives in manners consistent with revenue maximizing.

A. Main Results

Figure 2 plots the distribution of the outcomes after removing year means and school means for the APIP-willing schools *before* and *after* adoption. It is visually apparent that the distributions of the residuals after APIP adoption are to the right of the distributions before APIP adoption for AP course enrollment, the percentage of 11th or 12th graders taking at least one AP or IB exam, the number of students scoring above 1100/24 on the SAT/ACT, and the number of students enrolling in college in Texas. In contrast, the distributions for the number of high school graduates and students taking the SAT/ACT exams are largely unchanged before and after APIP adoption. This suggests that the APIP increased AP course enrollment, AP/IB exam taking, scoring above 1100/24 on the SAT/ACT, and college matriculation, but had no effect of SAT/ACT taking or graduating from high school. While not formal

Name /jhr453/JHR45_03_A03/Mp_608 04/21/2010 09:00AM Plate # 0 pg 608 # 18

statistical tests, Figure 2 provides an instructive graphical preview of the regression findings.

The main regression results are summarized in Table 2. The reported variable of interest is "treat," which indicates that the APIP had been adopted for at least one academic year. I present the effect of APIP adoption on each outcome in a separate row. Each column represents a different specification so that each column row entry presents the results from a separate regression model. Since AP enrollment data are only available for 2000 onward and AP/IB and college data are only available for 2001 onward, these variables have smaller sample sizes than the other variables.[25] The results would be qualitatively similar but less precise if one were to use the largest possible balanced sample with data from 2001 through 2005. All results are based on the APIP-willing schools only.[26]

1. Effect on AP Outcomes

The AP outcomes used in this paper are the log of AP course enrollment and the percentage of 11th and 12th graders taking AP or IB exams. The naïve specifications in Models 1, 2, 7, and 8 show that while APIP adoption is associated with increases in AP course enrollment and increases in the percentage of 11th and 12th graders who took AP or IB exams, after controlling only for year fixed effects and school fixed effects, APIP adoption is associated with a statistically insignificant 8.2 percent increase in AP course enrollment and a 2.55 point increase in the percentage of 11th and 12th graders who took AP or IB exams. Since the results in Models 1, 2, 7, and 8 do not control for any changes in enrollment or demographics over time, these results may not reflect a causal relationship. In the first preferred specification (Models 3 and 9), I test for the effect of APIP adoption net of any possible effect it may have on 11th or 12th grade enrollment. As such, I control for the size of the affected cohort—so that I control for 11th grade enrollment, 12th grade enrollment, lagged 11th grade enrollment, and lagged 10th grade enrollment. The results in Models 3 and 9 show that APIP adoption is not associated with a statistically significant increase in log AP course enrollment, but is associated with a 2.3 point increase in the percentage of 11th and 12th graders taking an AP or IB exam. To increase efficiency, I estimate a less restrictive model that allows observationally similar schools to have different time shocks. Specifically, I estimate a propensity score for each school based on the school demographic variables and historical values of the outcomes, assign each school its maximum estimated propensity score for all years, put each school into deciles of the maximum score, and include year fixed effects for each maximum propensity score decile.[27] The results are presented in Models 4 and 10. As expected, the findings are qualitatively similar; however, the standard errors are smaller for key variables.

Since the APIP could affect outcomes by affecting the size of the 11th and 12th grade class enrollments, while not changing the proportion of 11th or 12th graders

25. The data are set up so that if any one outcome has data in that year, all other outcomes also have data for that year (except for college going and AP/IB exam taking, which have smaller sample sizes).
26. Results using different samples are not appreciably different.
27. See appendix, note A3, for details of how the propensity score is estimated.

Table 2
Main APIP adoption Effects

Effect of APIP adoption on main outcomes (before/after comparison). Sample: APIP schools and future APIP-schools								
	log of AP course enrollment							N
	1	2	3	4	5	6	Placebo treatment	
Treat	0.162 [0.063]**	0.082 [0.069]	0.058 [0.075]	0.0951 [0.0732]	0.078 [0.072]	0.038 [0.067]	0.038 [0.067]	318
	Percent 11 and 12 graders taking AB/IB exams							
	7	8	9	10	11	12	Placebo treatment	
Treat	5.638 [0.919]**	2.557 [1.051]*	2.382 [0.874]*	2.2206 [0.676]*	2.233 [0.878]*	2.198 [0.893]*	−0.298 [1.211]	222

(*continued*)

Name /jhr453/JHR45_03_A03/Mp_610 04/21/2010 09:00AM Plate # 0 pg 610 # 20

Table 2 (*continued*)

Effect of APIP adoption on main outcomes (before/after comparison). Sample: APIP schools and future APIP-schools								
	log of graduates							
	13	14	15	16	17	18	Placebo treatment	
Treat	0.227 [0.030]**	0.106 [0.036]**	0.0117 [0.0184]	0.0118 [0.0184]	0.009 [0.017]	0.005 [0.023]	−0.022 [0.021]	578
	log of SAT/ACT takers							
	19	20	21	22	23	24	Placebo treatment	
Treat	0.189 [0.037]**	0.063 [0.046]	−0.0435 [0.0408]	−0.0464 [0.0541]	−0.046 [0.054]	−0.045 [0.045]	−0.049 [0.030]	578
	log score above 1100/24 on SAT/ACT							
	25	26	27	28	29	30	Placebo treatment	
Treat	0.167 [0.058]**	0.232 [0.070]**	0.1361 [0.0527]*	0.1976 [0.0253]**	0.198 [0.025]**	0.199 [0.046]**	0.09 [0.084]	578

	log attend college							
	31	32	33	34	35	36	Placebo treatment	
Treat	0.126 [0.035]**	0.117 [0.041]**	0.0496 [0.0283]+	0.0534 [0.0294]+	0.062 [0.030]+	0.048 [0.041]	− 0.085 [0.057]	226
School fixed effects	YES	YES	YES	YES	YES	YES	YES	
Year fixed effects	NO	YES	YES	YES	YES	YES	YES	
Controls included	NO	NO	1	2	2*	2**	1	
Decile-by-year effects	NO	NO	NO	YES	YES	YES	YES	

Robust standard errors in brackets. Standard error clustered at the district level.
+ significant at 10 percent; * significant at 5 percent; ** significant at 1 percent.
All regressions are based on a panel of 57 schools. Each coefficient estimate is the result of a different regression. Each column represents the same specification, while each row indicates a different dependent variable. Demographic controls are log school enrollment, percent white, percent black, percent Hispanic, percent Asian, percent Native American, percent free lunch, percent LEP, and the log of mobile students. Specification 1 includes school controls and affected cohort size (Grade 12, Grade 11, lagged Grade 11 and lagged Grade 10 enrollment for AP outcomes, and Grade 12 and lagged Grade 11 enrollment for non-AP outcomes based on the cohort of high school graduates). Specification 2 includes all variables in Specification 1 and year-by-propensity score decile fixed effects (the preferred model). Specification 2* includes all variables in Specification 2 with the exception of 11th and Grade 12 enrollment for AP outcomes and omit Grade 12 enrollment for the non-AP outcomes based on the cohort of high school graduates. Specification 2** includes all variables in Specification 2 but excludes all grade enrollment variables. Placebo treatment uses the second lead of actual adoption as the treatment. This variable denotes two years before actual APIP adoption.

Name /jhr453/JHR45_03_A03/Mp_612 04/21/2010 09:00AM Plate # 0 pg 612 # 22

who take AP courses or take AP/IB exams, the fifth column shows regression results that do not control for *contemporaneous* 11th and 12th grade enrollment. This model controls for the *initial* size of the affected cohorts (lagged 10th grade and lagged 11th grade enrollments) but does not control for the size of the cohort after APIP adoption. The results in Models 5 and 11 are almost identical to those in Models 3 and 4, and 9 and 10, indicating that the APIP did not increase AP course enrollment immediately after adoption (even if we allow effects through increases in cohort size). Models 6 and 12 shows estimates where none of the 10th, 11th or 12th grade enrollments are included. As one can see, this has little effect on the estimated coefficients.

2. Effect on Non-AP Outcomes

Rows 3 through 6 in Table 2 present the results for the non-AP outcomes. I will focus on the preferred models (Columns 3 through 6) that control for school fixed effects, year fixed effects, and school level controls for enrollment and demographic changes over time. Because these outcomes are based on the cohort of high school graduates, I control for cohort size by including 12th grade enrollment and the lag of 11th grade enrollment as covariates. All of these outcomes are in logs so that the interpretation is the percent increase in the outcome associated with APIP adoption. The basic DID results that control for school fixed effects, year fixed effects, the number of mobile students, student demographic variables, the log of grade 12 enrollment, and the lag of the log of Grade 11 enrollment are presented in the third column (Models 15, 21, 27 and 33). The results show that while the APIP does not have a statistically significant effect on the number of high school graduates or graduates who take the SAT/ACT exams, APIP adoption is associated with a 13 percent increase in the number of students scoring above 1100/24 on the SAT/ACT and a marginally statistically significant 4.96 percent increase in the number of students matriculating in college. The fourth column (Models 16, 22, 28 and 34) presents the results that allow schools with different estimated propensity scores to have different time effects. As with the AP outcomes, this model has little effect on the estimated coefficients, but it reduces the standard errors of the regression substantially thereby increasing statistical power. All subsequent models include these year effects for different propensity groups. The fifth column (Models 17, 23, 29 and 35) presents regression results that do not control for 12th grade enrollment, which could be affected by exposure to the APIP. Not controlling for 12th grade enrollment has no discernable effect on the results. The sixth column (Models 18, 24, 30 and 36) shows results that do not control for 12th grade or lagged 11th grade enrollment—again, the parameter estimates are virtually unchanged by this omission. These results make clear that any effects of the APIP on the outcomes of interest are not working through effects on cohort size. However, since controlling for cohort size does reduce the size of the standard errors without affecting the coefficients, all subsequent analyses include controls for cohort size.

3. Dynamic Effects

Because certain aspects of the APIP may take time to have an effect, I show these main outcomes during the first year, second year, and third plus year of the APIP.

Name /jhr453/JHR45_03_A03/Mp_613 04/21/2010 09:00AM Plate # 0 pg 613 # 23

Specifically, I estimate specifications like Equation 4, replacing the before/after indicator for APIP adoption with indicator variables denoting whether it is the first year of the APIP, the second year of the APIP, or the third year and beyond of the APIP. As before, all estimates are relative to before APIP adoption. The coefficients will map out how the program affects outcomes over time. With more data one could map out pretreatment years and more post-treatment years, however, data limitations preclude such estimates. Note that comparisons of first year estimates and third year estimates will not be on a balanced sample of schools since not all schools are observed in their first, second, and third years of adoption in sample.[28] 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.

Table 3 reports the coefficients on the indicator variables denoting the years since APIP adoption. Column 1 shows that while there is no statistically significant APIP effect on AP course enrollment in the first year, by the third year there is a statistically significant 33 percent increase. Column 2 shows that there is a 1.8 point increase in the percentage of 11th and 12th graders taking at least one AP or IB exam the first year of the APIP and about a 3.5 point increase thereafter. Taken together, the results show that much of the increased AP exam taking in the first two years was not driven by more students taking AP courses, but was likely a result of marginal students who already took AP courses but would not have taken the exams, being more likely to take the AP exams. By the third year and beyond, however, increases in the percentage of 11th and 12th graders taking AP/IB exams appear to be driven by both an increase in AP course taking, and an increase in AP/IB exam taking conditional on AP course taking.

Column 4 shows that APIP adoption may be associated with a marginally statistically significant decrease in SAT/ACT taking after the second year. While this could indicate that after two years students at APIP schools are nine percent less likely to take the SAT, given that the finding is not statistically significant at the five percent level, this may simply be noise. Since high school graduates in the first year of the APIP may only have had one year of AP course taking, while the second cohort and beyond would have had two years, one might expect the APIP effect to increase substantially from the first year, to the second year and beyond. Also, since students may have taken the SAT/ACT before the end of the school year, the SAT/ACT effects in the first year may reflect only half a year of APIP exposure for certain students. Therefore, one may expect relatively small first year SAT/ACT effects compared to the size of second year effects. This is consistent with what one observes. Column 5 shows statistically significant 15.7, 22.7, and 27.6 percent increases in the number of students scoring above 1100/24 on the SAT/ACT in the first, second, and third years and beyond, respectively. The relatively large first year effect on scoring above 1100/24 on the SAT/ACT is surprising given that it may not reflect a full year's exposure to the APIP for many students. The relatively large

28. For example, schools that adopted 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.

Name /jhr453/JHR45_03_A03/Mp_614 04/21/2010 09:00AM Plate # 0 pg 614 # 24

Table 3
APIP Adoption Effects Over Time

	Effect of the APIP over time					
	Sample: schools that adopt APIP and future-APIP adopters					
	1	2	3	4	5	6
	log of AP course enrollment	percent 11th and 12th graders taking AB/IB exams	log of graduates	log of SAT/ACT takers	log score above 1100/24 on SAT/ACT	ln attend college
First year	0.0805	1.8791	0.0099	−0.0315	0.1578	0.0521
	[0.0787]	[0.9338]$^{+}$	[0.0148]	[0.0451]	[0.0331]**	[0.0315]
Second year	0.0942	3.8755	0.0107	−0.0482	0.2275	0.0602
	[0.0928]	[1.5498]*	[0.0234]	[0.0868]	[0.0262]**	[0.0250]*
Third year and beyond	0.3358	3.4861	0.0196	−0.0915	0.2764	0.0726
	[0.1419]*	[1.4229]*	[0.0350]	[0.0513]$^{+}$	[0.0387]**	[0.0403]$^{+}$
Observations	318	222	578	578	578	226

	Sample: schools that adopt APIP after 2000 and future-APIP adopters					
	log of AP course enrollment	percent 11th and 12th graders taking AB/IB exams	log of graduates	log of SAT/ACT takers	log score above 1100/24 on SAT/ACT	ln attend college
	7	8	9	10	11	12
First year	0.0715	1.3633	0.0016	−0.0494	0.2383	0.0588
	[0.0932]	[1.2020]	[0.0269]	[0.0593]	[0.1020]*	[0.0427]
Second year	0.1011	3.2194	−0.0051	−0.1194	0.3204	0.0859
	[0.1230]	[1.9485]	[0.0331]	[0.0973]	[0.0931]**	[0.0477]+
Third year and beyond	0.3173	2.5845	0.0297	−0.2184	0.5791	0.113
	[0.1775]+	[2.3064]	[0.0371]	[0.1101]+	[0.1104]**	[0.0352]**
Observations	265	190	483	483	483	190

Robust standard errors in brackets. Standard error clustered at the district level.
+ significant at 10 percent; * significant at 5 percent; ** significant at 1 percent.
Columns 1 through 6 are based on a panel of 57 schools. Columns 7 through 12 are based on 48 schools. All regressions are based on the preferred specification using the sample of APIP and future-APIP schools. All models include school demographic and enrollment variables, school fixed effects, and time fixed effects for each propensity score decile.

Name /jhr453/JHR45_03_A03/Mp_616 04/21/2010 09:00AM Plate # 0 pg 616 # 26

first year effect would be consistent with the APIP having a positive effect on general motivation or effort rather than improving SAT/ACT performance through increases AP exposure alone. Column 6 shows that APIP adoption is associated with a statistically significant six percent increase in college matriculation in year two and a marginally statistically significant 7.25 percent increase in the third year and beyond.

B. Endogeneity Concerns and Falsification Tests

Section VIA shows that APIP adoption is associated with increased AP participation, improved SAT/ACT performance, and increased college going. There remain, however, three potential threats to validity that should be addressed. Specifically, (1) early adopters may differ from late adopters in ways that bias the results, (2) the timing of adoption may be endogenous, and (3) selective migration (an inflow of smart, motivated students) into APIP schools could drive the results.

To ensure that the findings are not driven by comparing the first set of early adopters (in 1996), who may have already had more rapid improvement in outcomes, to the later adopters (after 2000), Columns 7 through 12 of Table 3 show the dynamic regressions on the APIP-willing sample after removing the early-adopting schools. If the early adopters were more motivated so that the results were driven by school selection, the results should be smaller or nonexistent in the sample of late adopters. In fact, the effects in Columns 7 through 12 are *larger* than those using the full APIP sample, suggesting that comparing late to early adopters was not driving the results; if there is any bias in comparing late to early adopters, such bias would likely attenuate rather than inflate the APIP effects. The results show that APIP adoption is associated with statistically significant increases in AP course enrollment, the number of students scoring above 1100/24 on the SAT/ACT, and the number of students matriculating in college. The loss of statistical significance for AP/IB exam taking between Columns 2 and 7 is clearly the result of larger standard errors rather than a lack of an effect since the coefficients are very similar.

The second concern is that the exact timing of adoption may be endogenous. Within the group of APIP schools, adoption of the APIP could reflect *changes* in school motivation if schools that became more motivated over time adopted the APIP first because they expressed interest first. It typically takes two years between when a school first expresses interest in the APIP and when it is implemented. Therefore, if unobserved *changes* in motivation led schools to express interest in the APIP, one should observe improved results for APIP schools two years *before* actual APIP adoption. The last Column of Table 2, labeled "placebo treatment," shows the coefficient on the two-year lead of adoption.[29] None of the coefficients are statistically significant and the signs are often the opposite of the actual adoption estimates. This suggests that the APIP schools only saw improvements *after* they adopted APIP and not before (when they would have expressed a desire to implement the APIP)—as one would expect if the timing of adoption was exogenous.

29. I use two years before adoption because the schools will have been in the implementation stages (for example, training, announcement) in the t-1 year, but will have had no exposure (and would probably not have been chosen by a donor) in t-2. As such, t-2 is the best placebo year.

Name /jhr453/JHR45_03_A03/Mp_617 04/21/2010 09:00AM Plate # 0 pg 617 # 27

The last concern is that the results were driven by selective migration. I present three pieces of evidence showing that this was not the case. First, if schools that adopted the APIP had an inflow of high-ability students, there would have been associated with APIP adoption an increase in school enrollment and in Grade 12 enrollment. To test for this possibility, I estimate the effects of APIP adoption on Grade 12 and school enrollment. I control for the lag of Grade 12 enrollment and the lag of school enrollment. Columns 1 and 2 of Table 4 show that while the point estimates are positive, there is no statistically significant relationship between APIP adoption and Grade 12 enrollment or school enrollment.

Readers may still worry that the results could have been driven by changes in the selectivity of in-migration. Therefore, I test for *changes* in selective migration. Suppose the only effect the APIP had was to attract smart, motivated students to enroll (transfer) in program schools. If this were so, one would observe increases in the number of high school graduates, the number of students who attain Texas Academic Skills Program (TASP) equivalency, the number of students taking the SAT/ACT *in addition* to the improved SAT/ACT results, and the number of students matriculating in college. The results in Tables 3 and 4 show that this did not occur. The estimates in Columns 3, 4, and 5 of Table 4 show that there was no statistically significant effect of the APIP on the number of students achieving TASP equivalency or the number graduates, and that the point estimate for the number of student taking the SAT/ACT exams is *negative*.[30] The findings are inconsistent with the effects of the APIP being the result of selective migration.

Another way to show robustness to selective migration is to aggregate up to the school-district level and run district-level regressions. Since most cross-school migration would occur within districts rather than across districts, results based on district-level variation in treatment intensity would be robust to within district selective migration. To show robustness to selective migration, one would like to show that as APIP treatment intensity increases within a district, there are improvements relative to other districts. Table 5 presents these district-level regressions where the variable of interest is the share of treated schools in the district at that point in time, and all variation is based on within-district variation over time. The district-level regression results are similar to the main results in Table 2, suggesting that selective migration within a district was not driving the main results. Since selective migration is such an important concern, in Section VIC I show that the SAT/ACT results are robust to looking only at the sample of students who did not transfer and were in the same high school for all four years. These tests present strong evidence against the selective migration hypothesis.

C. Effects by Gender and Race

In this section, I present results based on subsamples based on gender and race. It has been established that AP participation of minorities and low-income students

30. Because readers may wonder if APIP adoption is associated with increases in the graduation rate (conditional on grade 10 enrollment), I also estimate such a model. Column 6 of Table 4 shows weak evidence that APIP adoption may increase graduation rates (not conditional on Grade 11 or Grade 12 enrollment) after two years.

Name /jhr453/JHR45_03_A03/Mp_618 04/21/2010 09:00AM Plate # 0 pg 618 # 28

Table 4
First Test for Selective Migration

Tests for selective migration, effect on graduation rates, and substitution away from advanced courses							
Sample: Schools that will have adopted APIP by 2008							
	1	2	3	4	5	6[a]	7
	log grade 12 enrollment	log school enrollment	percent students earning TASP	log graduates	log of SAT/ ACT takers	log graduates	percent 9th–12th graders taking advanced courses
First year	0.0037	0.0139	2.0803	0.0099	−0.0315	0.0289	−1.174
	[0.0113]	[0.0146]	[1.5921]	[0.0148]	[0.0451]	[0.0179]	[1.453]
Second year	0.0036	0.0171	−0.2053	0.0107	−0.0482	0.021	−2.24
	[0.0154]	[0.0210]	[1.8429]	[0.0234]	[0.0868]	[0.0239]	[1.262]
Third year and beyond	0.0355	0.0372	1.2166	0.0196	−0.0915	0.0509	−0.885
	[0.0321]	[0.0244]	[2.5610]	[0.0350]	[0.0513]$^{+}$	[0.0287]$^{+}$	[1.358]

Name /jhr453/JHR45_03_A03/Mp_619 04/21/2010 09:00AM Plate # 0 pg 619 # 29

Grade 12 enrollment	N	N	Y	Y	Y	N	Y
Lag Grade 12 enrollment	Y	Y	N	N	N	N	N
Second lag of Grade 12 enrollment	Y	Y	N	N	N	Y	N
School enrollment	N	N	Y	Y	Y	N	Y
Lag school enrollment	Y	Y	N	N	N	N	N
Second lag of school enrollment	N	N	N	N	N	Y	N
Lag of Grade 11 enrollment	Y	Y	Y	Y	Y	N	Y
Observations	463	463	578	578	578	463	578

Robust standard errors in brackets. Standard errors clustered at the school district level.
+ significant at 10 percent; * significant at 5 percent; ** significant at 1 percent.
All regressions are based on a panel of 57 schools. All regressions include school fixed effects, year by propensity score decile fixed effects, and school demographic variables.
a. This model does not include current or lagged Grade 12 or school enrollment, so that these estimates are conditional on grade 10 enrollments two years prior. This provides weak evidence that the APIP may increase the proportion of 10th graders who subsequently graduate.

Name /jhr453/JHR45_03_A03/Mp_620 04/21/2010 09:00AM Plate # 0 pg 620 # 30

Table 5
Second Test for Selective Migration

	District level regressions					
	No controls for enrollment					
	1	2	3	5	5	6
	log of AP course enrollment	percent 11th and 12th graders taking AB/IB exams	log of graduates	log of SAT/ ACT takers	log score above 1100/24 on SAT/ACT	ln attend college
Share treated	0.002 [0.084]	5.177 [1.003]**	0.056 [0.035]	0.036 [0.048]	0.212 [0.074]**	0.092 [0.044]*
	Controls for the second lag of 10th grade enrollment					
	7	8	9	10	11	12
	log of AP course enrollment	percent 11th and 12th graders taking AB/IB exams	log of graduates	log of SAT/ ACT takers	log score above 1100/24 on SAT/ACT	ln attend college
Share treated	−0.042 [0.084]	5.023 [1.002]**	0.026 [0.023]	−0.014 [0.038]	0.158 [0.068]*	0.076 [0.037]*

Name /jhr453/JHR45_03_A03/Mp_621 04/21/2010 09:00AM Plate # 0 pg 621 # 31

Controls for the second lag of 12th grade enrollment

	13	14	15	16	17	18
	log of AP course enrollment	percent 11th and 12th graders taking AB/IB exams	log of graduates	log of SAT/ ACT takers	log score above 1100/24 on SAT/ACT	ln attend college
Share treated	−0.053	4.928	−0.003	−0.045	0.105	0.033
	[0.083]	[1.012]**	[0.019]	[0.036]	[0.063]+	[0.037]
Observations	85	85	171	156	156	74

Robust standard errors in brackets. Standard error clustered at the district level.
+ significant at 10 percent; * significant at 5 percent; ** significant at 1 percent;
The results are based on a panel of 15 school districts. All regressions include district fixed effects and year fixed effects. Columns 1 through 6 have no other controls, Columns 7 through 12 include the second lag of Grade 10 enrollment, and Columns 13 through 18 show results that control for 12 grade enrollment at the district level.

Name /jhr453/JHR45_03_A03/Mp_622 04/21/2010 09:00AM Plate # 0 pg 622 # 32

tends to be lower than that of middle-class white students at the same high schools (Klopfenstein 2004). Insofar as these differences reflect suboptimal student or teacher effort, one might expect 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 adolescents, girls have more self-discipline and delay gratification more than boys (Duckworth, Lee, and Seligman 2006; Silverman 2003).

Due to the nature in which the data are available, the outcomes are reported in percentages. Table 6 shows the effect of APIP adoption on the percentage of 11th or 12th graders who took at least one AP/IB exam. The results show that the campus-wide increases in AP/IB exam-taking were driven by increased participation for black and Hispanic students. The results do not show any statistically significant effect of the APIP on the proportion of white students who took 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 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 were increases in AP/IB exam taking for both genders with no greater effect on girls.

Given the differences in the effect of the APIP on the number of 11th and 12th graders across ethnicities who took at least one AP or IB exam, one would expect to see the corresponding differences in the effect of the APIP on SAT/ACT performance. Table 7 shows the effect of the APIP on the percentage of non-special education high school graduates who scored above 1100/24 on the SAT/ACT examinations for the different groups. By the third year of the program there were positive effects for all groups. Given that Hispanics and blacks are typically under-represented 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 (around five points), the differences in relative impact, however, are sizable. Compared to the base levels this represents about a 25 percent increase for whites, a 50 percent increase for Hispanics, and a near 100 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 have been 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 (that is, those students who did not migrate). Due to heavy data masking, these data are not available for black students.[32] Table 8 shows the effect of APIP adoption

31. For example, Anderson (2005); Angrist, Lang, and Oreopoulos (2009); Angrist and Lavy (2007); and Katz, Kling, and Ludwig (2005).

32. To preserve student confidentiality, the Texas Education Agency does not release data that is based on a sample of fewer than 5 students. If statistics for more than one group are requested, data are removed where statistics can be inferred from combining publicly available data and any of the other data requested. The process of removing such data is referred to as "masking."

Table 6
Effects on AP/IB Examination Taking by Gender and Ethnicity

APIP effect on percentage of 11th 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	1.8791	2.505	−0.203	2.033	2.035	1.943
	[0.9338]+	[0.619]**	[1.896]	[1.035]+	[0.911]*	[0.951]+
Second year	3.8755	4.293	−1.163	5.309	4.337	3.705
	[1.5498]*	[1.891]*	[2.43]	[1.820]*	[1.225]**	[1.288]*
Third year and beyond	3.4861	2.673	0.388	6.099	2.125	5.657
	[1.4229]*	[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 by 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

Robust standard errors in brackets.
+ significant at 10 percent; * significant at 5 percent; ** significant at 1 percent.
Standard errors clustered at the school district level.
Control variables include all school demographic and enrollment variables.

Name /jhr453/JHR45_03_A03/Mp_624 04/21/2010 09:00AM Plate # 0 pg 624 # 34

Table 7
Effect on SAT/ACT Performance by Gender and Ethnicity

APIP effect on percentage of graduates who scored above 1100/24 on SAT/ACT						
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 by 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

Robust standard errors in brackets.
+ significant at 10 percent; * significant at 5 percent; ** significant at 1 percent; Standard errors clustered at the school district level.
Control variables include all school demographic and enrollment variables.

Table 8
Robustness to Selective Migration

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
Subsample	Black	Hispanic	White	Black	Hispanic	White
	log above 1100/24 on SAT/ACT	log above 1100/24 on SAT/ACT	log above 1100/24 on SAT/ACT	log above 900/19 on SAT/ACT	log above 900/19 on SAT/ACT	log above 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]
Third year and beyond	–	0.18	0.262	0.123	0.385	0.257
	–	[0.038]**	[0.15]$^{+}$	[0.092]	[0.204]$^{+}$	[0.096]*

(*continued*)

Name /jhr453/JHR45_03_A03/Mp_626 04/21/2010 09:00AM Plate # 0 pg 626 # 36

Table 8 (*continued*)

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
Subsample	Black	Hispanic	White	Black	Hispanic	White
	log above 1100/24 on SAT/ACT	log above 1100/24 on SAT/ACT	log above 1100/24 on SAT/ACT	log above 900/19 on SAT/ACT	log above 900/19 on SAT/ACT	log above 900/19 on SAT/ACT
			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]
Third year and beyond	–	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]
Third year and beyond	–	0.335	0.351	0.295	0.54	0.334
	–	[0.120]*	[0.217]	[0.175]	[0.292]+	[0.185]+

Robust standard errors in brackets. Standard errors clustered at school district level.
+ significant at 10 percent; * significant at 5 percent; ** significant at 1 percent.
Notes: All regressions are based on 384 observations from 57 schools. All regression specifications include school-level controls, school fixed effects, year fixed effects, and propensity score decile by year fixed effects. Due to data masking count fewer than five are coded as missing. Treatment effects are not estimated for above 1100/24 for blacks because there are too many missing values. To provide bounds for those masked observations, missing values are assumed to be four and zero.

Name /jhr453/JHR45_03_A03/Mp_627 04/21/2010 09:00AM Plate # 0 pg 627 # 37

on white and Hispanic graduates scoring above 1100/24 on the SAT/ACT. Due to data masking there are some missing values that correspond to counts that are between 0 and 4. Table 8 shows that 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 8 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. However, there does not seem to be an 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.

D. 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, students could not have simply diverted their effort away from other advanced courses (such as dual enrollment courses, or college courses taken while in high school) 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. However, Column 7 of Table 4 shows little evidence of this kind of distortion: By the third year of the APIP, one sees a very small statistically insignificant decrease in dual enrollment course taking. This suggests that students and teachers *did not* game the incentives by substituting away from other advanced courses toward AP courses, resulting in an overall increase in rigorous course participation.

Evidence from discussions with guidance counselors at three different APIP high schools in Dallas strongly suggests there were school-wide campaigns 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 schools, the guidance counselors were given explicit instructions to identify those students who should be taking AP courses and to encourage AP participation. A large part of this campaign involved providing information. Guidance counselors and AP teachers sold the AP program to students who were interested in going to college, citing the scholarships one could earn based on AP scores, the tuition one could save by graduating at an accelerated pace, and the potential increase in high school GPA, which could increase the student's likelihood of being in the class's top ten percent and gaining admittance into a good college. There is also evidence that certain barriers to taking AP courses were 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 these courses. All guidance counselors mentioned a shift in student and teacher attitudes toward AP courses. AP courses are now considered difficult courses that anyone can

Name /jhr453/JHR45_03_A03/Mp_628 04/21/2010 09:00AM Plate # 0 pg 628 # 38

take, as opposed to being available only for the very brightest of students (one AP teacher noted that she now has to turn students away). The example of 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, increased access through teacher encouragement, and increased teacher and guidance counselor recommendations. The financial incentives to students and teachers may have been responsible for the increased student and teacher effort in AP courses, but these aspects of the program were downplayed by the counselors.

The large increases in AP participation are difficult to reconcile with the theoretical framework put forward in section IV without reference to several of the elements highlighted by guidance counselors. The theory and evidence thus far have suggested that students and teachers were 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. If the results are solely due to revenue maximizing behaviors, one might expect the APIP effect to be monotonically increasing in the size of the cash rewards. I test this prediction in a regression model by interacting the "treat" variable with the levels of the rewards ($100 per exam, $101–$499 per exam, and $500 per exam). Note that schools with higher student rewards also paid higher teacher rewards. Since the incentive levels were not exogenously determined and the sample of schools within each incentive level group is small, 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.

Evidence presented in Table 9 supporting the notion that that the effects of the program were stronger in schools with higher incentives is mixed at best. Column 1 shows that the schools that paid between $101 and $499 per exam had the greatest AP enrollment response, and the coefficient for schools that paid $500 per exam is the smallest. In column 2, the schools that paid $100 per exam had the *largest* AP/IB exam taking response. Column 5 shows that while all schools had improved SAT/ACT results, schools that paid $500 per exam had the smallest improvements and those that paid in the middle range had the most. College going in column 6 is the only outcome where the highest reward group has the largest effect. However, the effect is not monotonic: the middle incentive group actually has a negative coefficient while schools that pay $100 per exam have a positive coefficient. For none of the outcomes are the effects monotonically increasing in the size of the rewards and in only one outcome is the effect largest where the rewards are largest. In sum, the results of Table 9 do not support the hypothesis that the APIP effect is increasing in the size of the rewards.

The last test is based on the hypothesis that if students and teachers were responding *solely* to the rewards, there should have been a greater participation response in subjects for which rewards were provided than in subjects for which there was no reward. In fact, one might expect a decline in enrollments in AP courses for which rewards were not provided. There were a few schools that offered rewards for all AP subjects while most only offered rewards for math, science, and English. I test whether the increase in the ratio of the number of math, science, and English AP course enrollees to the number of total AP course enrollees is smaller in APIP-

Table 9
The Effects of APIP Adoption by Incentive Level

The effects of APIP adoption by incentive level						
Sample: schools that will have adopted APIP by 2008						
	1	2	3	4	5	6
	log AP enrollment	percent 11th and 12th graders taking AP or IB exams	log Graduates	log SAT/ACT takers	log Above 1100/24 on SAT/ACT	log Attend College
$100 per passing AP exam	0.062	3.882	0.033	−0.041	0.204	0.043
	[0.135]	[1.609]*	[0.023]	[0.030]	[0.047]**	[0.042]
$101–$499 per passing AP exam	0.199	0.547	−0.002	−0.159	0.302	−0.11
	[0.058]**	[2.040]	[0.044]	[0.069]*	[0.057]**	[0.062]+
$500 per passing AP exam	0.021	1.347	−0.026	0.017	0.114	0.074
	[0.051]	[0.653]+	[0.025]	[0.073]	[0.051]*	[0.022]**
Controls?	YES	YES	YES	YES	YES	YES
School fixed effects?	YES	YES	YES	YES	YES	YES
Year by decile fixed effects?	YES	YES	YES	YES	YES	YES
Observations	313	222	578	578	578	226
Number of campus	57	57	57	57	57	57

Robust standard errors in brackets. Standard errors clustered at the school district level.
+ significant at 10 percent; * significant at 5 percent; ** significant at 1 percent.
Control variables include all school demographic and enrollment variables.

Name /jhr453/JHR45_03_A03/Mp_630 04/21/2010 09:00AM Plate # 0 pg 630 # 40

adopting schools that paid rewards for social studies and humanities, by interacting "treat" with an indicator for whether the schools rewarded social studies AP exams. Again, this evidence should be taken as suggestive. The preferred model yields a coefficient of -0.003 with a standard error of 0.04 on the "treat" variable and a coefficient of 0.11 with a standard error of 0.029 on the interaction with whether the school gave rewards for social studies. The interaction is statically significant and is *positive*, suggesting that students took a *greater* share of math, science, and English AP courses at schools that gave rewards for social science courses. This result is inconsistent with the hypothesis that students and teachers substituted away from AP courses for which there were no rewards toward those for which there were rewards.

VII. Conclusions

Using a carefully selected group of comparison schools within which APIP adoption is likely exogenous, I find that the APIP is associated with increases in the number of students taking AP courses and the number of students taking AP/IB exams. Moreover, APIP adoption is associated with increases in the number of students scoring above 1100/24 on the SAT/ACT exams and the number of students who matriculate in college. I show that these results are not driven by comparing early adopters to late adopters, endogenous timing of APIP adoption, or selective migration. While there are no significant differences by gender, I find substantially increased AP/IB exam taking for black and Hispanic students. I also find improvements in SAT/ACT performance across all ethnic groups and for both males and females.

The improvements in SAT/ACT performance were likely the result of increased exposure to rigorous material, but could also have been the result of increased effort in SAT/ACT if students studied harder for both AP and SAT/ACT in order to get into college. The lack of an increase in 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 college costs (tuition, effort, time, etc.). However, it is possible that there was an effect on college application behavior that was not picked up by SAT/ACT taking. The theoretical possibility that students and teachers would divert resources away from other tasks toward AP courses is not supported by the data. The APIP had no effect on the number of 12th graders who graduated from high school, the number of 10th graders who graduated from high school, or the proportion of high school graduates who attained TASP equivalency. However, the APIP may have had negative effects on other *unobserved* outcomes. 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 suboptimal, so that marginal students who may have benefited from taking AP courses were not doing so.

The curricular changes and the early emphasis on pre-AP material would not have affected graduating seniors until a few years after the program had been adopted. Therefore, the changes that took place at year one of the APIP were likely due to

Name /jhr453/JHR45_03_A03/Mp_631 04/21/2010 09:00AM Plate # 0 pg 631 # 41

the incentives, the AP courses, and improvements in AP instruction.[33] The fact that there is an increase in program effect over time suggests that the push to promote AP participation in early grades through emphasis on preAP courses and vertical teams may also 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 exams. 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 AP program's benefits, 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 directly to the monetary incentives in a "carrot and stick" manner. The body of evidence is more consistent with explanations put forth by guidance counselors, such as changes in peer norms and teacher norms, increased emphasis on AP courses, and increased information given to students on the benefits to taking AP courses. The findings are suggestive of some of the 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 may have had low initial participation rates because (1) peer norms did not promote taking AP courses, (2) they were less likely to have good information on the college application process, and (3) student aspirations may have been low due to suboptimal teacher encouragement. The sum of the evidence suggests that student or teacher incentives alone would not have been as effective since it was likely the combination of increased teacher effort, increased student effort, *and* better instruction that lead to improved outcomes over time. This is consistent with the finding of larger gains when college students were provided with incentives and services than with incentives alone (Angrist, Lang, and Oreopoulos 2009), and with individual student incentives being less effective when teachers were not aware (Angrist and Lavy 2002).

While these results are encouraging in light of the rapid growth of similar programs, the long-term effects of the APIP on college and labor-force outcomes is unknown. The program costs about $200 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 will 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, as of 2006, Texas was spending $80 million per year to bring ill-prepared college students up the level at which they could cope with college-level course material. Since the program could potentially reduce the demand for remedial courses while in college, this could provide cost

33. I cannot rule out the possibility that there was an influx of quality teachers to the APIP schools during the first 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.

Name /jhr453/JHR45_03_A03/Mp_632 04/21/2010 09:00AM Plate # 0 pg 632 # 42

savings as well. 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 local governments.

References

Anderson, Michael L. 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, September. Unpublished.

Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. 2002. "Vouchers for Private Schooling in Colombia: Evidence from Randomized Natural Experiments." *American Economic Review* 92(5):1535–58.

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2009. "Incentives and Services for College Achievement: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics* 1(1):136–63.

Angrist, Joshua, and Victor Lavy. 2002. "The effect of High School Matriculation Awards: Evidence from Randomized Trials." National Bureau of Economic Research Working Paper 9389.

———. 2007. "*The Effects of High Stakes High School Achievement Awards: Evidence from a Group Randomized Trial*." MIT and Hebrew University of Jerusalem. Unpublished.

AP Strategies. 2006. "*Creating Partnerships for Academic Excellence*." October 11.

Atkinson, Adele, Simon Burgess, Bronwyn Croxson, and Paul Gregg. 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, U.K.

Binder, Melissa, Philip T. Ganderton, and Kristin Hutchens. 2002. "Incentive Effects of New Mexico's Merit Based State Scholarship Program: Who Responds and How?" In *Who Should We Help? The Negative Social Consequences of Merit Scholarships*, eds. D. E. Heller and P. Marin, 41–56. Cambridge, Mass.: The Civil Rights Project at Harvard University.

Card, David. 1995. "Earnings, Schooling, and Ability Revisited." *Research in Labor Economics* 14:23–48.

College Board. 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, August.

Cornwell, Christopher M., Kyung Hee Lee, and David B. Mustard. 2005. "The Effects of Merit-Based Financial Aid on Course Enrollment, Withdrawal and Completion in College." *Journal of Human Resources* 40(4):895–917.

Costrell, Robert M. 1993. "An Economic Analysis of College Admission Standards." *Education Economics* 1(3):227–41.

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, Angela Lee, and Martin E. P. 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, Lynn Mellor, and Shuling Jian. 2006. "The Relationship between Advanced Placement and College Graduation." National Centre for Education Assessment 2005 AP Study Series, Report 1, February.

Name /jhr453/JHR45_03_A03/Mp_633 04/21/2010 09:00AM Plate # 0 pg 633 # 43

Eberts, Randall, Kevin Hollenbeck, and Joe Stone. 2000. "Teacher Performance Incentives and Student Outcomes." W. E. Upjohn Institute Working Paper No. 00-65, August.

Eimers, Mardy. 2003. "Dual Credit and Advanced Placement: Do They Help Prepare Students for Success in College?" Mid-AIR and AIR Conference. Tampa, Florida.

Frederick, Shane, George Lowenstein, and Ted O'Donoghue. 2002. "Time Discounting and Time Preferences: A Critical Review." *Journal of Economic Literature* 40:351–401.

Geiser, Saul, and Veronica Santelices. 2004. "The Role of Advanced Placement and Honors Courses in College Admissions." Center for Studies on Higher Education, Research and Occasional Paper Series, CSHE.4.04.

Glewwe, Paul, Nauman Ilias, and Michael Kremer. 2003. "Teacher Incentives" National Bureau of Economic Research Working Paper No. 9671.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design" *Journal of Law, Economics and Organization* 7 (Special Issue): 24–52.

Hudgins, Karen. 2003. "Advanced Placement Program Proves It Pays to Study Hard: A Kick Start for College." Fiscal notes, May.

Jackson, C. Kirabo. 2007. "A Little Now for a Lot Later: A Look at a Texas Advanced Placement Incentive Program." Cornell Higher Education Research Institute Working Paper No. 107.

Jackson, C. Kirabo. 2008. "Cash for Test Scores: The impact of the Texas Advanced Placement Incentive Program." *Education Next* 8(4): 71–77.

Jacob, Brian A., and Steven D Levitt, 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics* 118 (3):843–77.

Katz, Lawrence F., Jeffrey R. Kling, and Jens Ludwig. 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Mobility Experiment." *Quarterly Journal of Economics* 120 (February): 87–130.

Klopfenstein, Kristin. 2004. "The Advanced Placement Expansion of the 1990s: How Did Traditionally Underserved Students Fare?" *Education Policy Analysis Archives 12, no. 68*.

Kohn, Alfie. 1999. *Punished by Rewards the Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*. Bridgewater, N.J.: Replica Books.

Kremer, Micheal, Edward Miguel, and Rebecca Thornton. 2004. "Incentives to Learn." National Bureau of Economic Research Working Paper No. W10971.

Lavy, Victor. 2002. "Evaluating the Effect of Teachers' Group Performance Incentives on Students' Achievements." *Journal of Political Economy* 110(6):1286–1318.

———. 2004. "Performance Pay and Teachers' Effort, Productivity and Grading Ethics." National Bureau of Economic Research Working Paper No. 10622.

Lyon, John. 2007. "Program to Offer Students Money for High Test Scores." Arkansas News Bureau, October 9.

Mathews, Jay. 2004. "Paying Teachers and Students for Good Scores." *Washington Post*, August 10.

Mathews, Jay. 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*, March 9.

Medina, Jennifer. 2007. "Making Cash a Prize for High Scores on Advanced Placement Tests." *New York Times*, October 15.

Morgan, Rick, and Len Ramist. 1998. "Advanced Placement Students in College: An Investigation of Course Grades at 21 Colleges." Princeton, N.J.: Educational Testing Service.

Silverman, I. W. 2003. "Gender Differences in the Delay of Gratification: A Meta Analysis." *Sex Roles* 49 (November): 451–63.

Name /jhr453/JHR45_03_A03/Mp_634 04/21/2010 09:00AM Plate # 0 pg 634 # 44

Note A1

The Texas ten percent rule was put in place in 1997 and ensured that the top ten percent of students from each high school in the state would be guaranteed admission to a Texas public university. One would expect college matriculation rates to have increased in schools 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 academic year 1999–2000. Under the statewide program, the state appropriated $21 million over the years 1998–2000 for the Texas APIP, 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 of up to $450, up to $3,000 in equipment and material grants for AP classes, and financial incentives to the schools of up to $100 for each student who scores 3 or better on any AP exam. One would expect this policy to increase AP participation and effort even if the APIP was not adopted by the selected 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 1998–99 academic year, students were only required to take 3 credits (often over 3 years starting in 9th grade) in mathematics and science and 4 credits in English to satisfy their high school graduation requirements. Therefore, students who had taken these courses by 10th or 11th grade would either have had a free period, be taking some less rigorous elective, or be taking a dual enrollment course at a college. If schools did not offer AP mathematics, students who had fulfilled the graduation requirements would either have been involved in a dual enrollment course or have taken no math class at all. Students who had completed the science requirements would have been involved in a dual enrollment science class; taken a less rigorous science elective such as geology, anatomy, or physiology; or had a free period. Those students who took AP English courses would have been 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 counselors at several Dallas high schools)

Note A3

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 (on the full sample). This captures 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 esti-

Name /jhr453/JHR45_03_A03/Mp_635 04/21/2010 09:00AM Plate # 0 pg 635 # 45

mates 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. Because no schools located in a small or large town were ever treated or selected to be treated, all schools located in small or large towns were automatically removed from the sample. The estimates of this probit regression can be found in table A2.

Note A4

For the post hoc power calculation I assume 30 treated school observation and 30 untreated school observations. (In actual fact there are 40 treated observations for some outcomes and more than 40 untreated observations for all outcomes). Since I condition on school effects and year effects, I use the variance of the within school residual (after taking out year effects) as the relevant variance of the outcome. I assume equal variance, and independence across groups (since I am using the within school variation) for ease of computation. Table A3 presents the results of this post hoc power analysis. In the last column, I show the effect size that one would expect to detect with 80 percent power (at the 5 percent significance level) in a two-sample *t*-test. For all outcomes for which I find effects (except college enrollment), this minimum effect size lies within the range of the actual estimated effects. For college enrollment I calculated that the highest estimated effect would have been detected with more than 70 percent power at the 5 percent significance level, and at the ten percent level with 80 percent power.

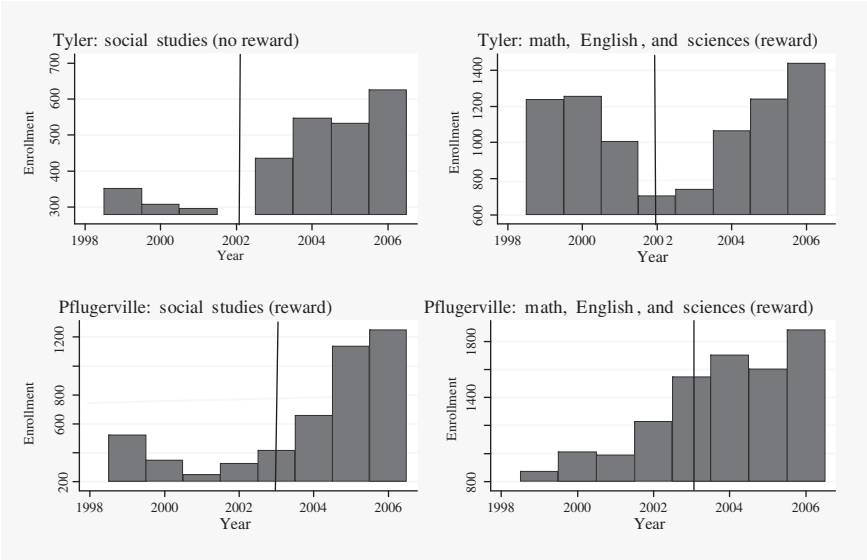

Figure A1
Number of Exams Taken per Year by Subjects Rewarded

Name /jhr453/JHR45_03_A03/Mp_636 04/21/2010 09:00AM Plate # 0 pg 636 # 46

Table A1
Difference-in-difference estimates of APIP adoption on main outcomes (sample: APIP schools and future-APIP schools)

	1	2	3	4	5	6	7	8	9	10	11	12
	log AP course enrollment	percent 11th and 12 graders taking AB/IB exams	log graduates	log SAT/ACT takers	log score above 1100/24 on SAT/ACT	log attend college	log AP course enrollment	percent 11th and 12 graders taking AB/IB exams	log graduates	log SAT/ACT takers	log score above 1100/24 on SAT/ACT	log attend college
Treat	0.058	2.382	0.0117	−0.0435	0.1361	0.0496	0.095	2.206	0.0118	−0.0464	0.1976	0.0534
	[0.075]	[0.874]	[0.0184]	[0.0408]	[0.0527]*	[0.0280]	[0.073]	[0.676]	[0.0184]	[0.0541]	[0.0253]	[0.0294]
Log enrollment	0.585	7.716	0.5801	0.6347	−0.029	0.1447	0.583	8.456	0.5947	0.6775	−0.0295	0.0778
	[0.572]	[4.879]	[0.1276]	[0.1775]	[0.1848]	[0.2327]	[0.588]	[6.132]	[0.1298]	[0.1776]	[0.1867]	[0.2308]
Percent white	−0.188	−0.332	−0.0339	−0.0703	0.0003	0.099	−0.131	−0.744	−0.0574	−0.0689	0.0986	0.0431
	[0.277]	[5.159]	[0.0556]	[0.0574]	[0.1947]	[0.1554]	[0.286]	[6.210]	[0.0643]	[0.0530]	[0.1927]	[0.2035]
Percent black	−0.189	−0.728	−0.0412	−0.0807	0.0025	0.0915	−0.141	−1.107	−0.0655	−0.0804	0.091	0.0328
	[0.272]	[5.196]	[0.0563]	[0.0575]	[0.1854]	[0.1534]	[0.283]	[6.278]	[0.0646]	[0.0537]	[0.1829]	[0.2016]
Percent Hispanic	−0.211	−0.637	−0.0384	−0.0859	−0.0259	0.1074	−0.161	−1.095	−0.0631	−0.0861	0.065	0.0479
	[0.282]	[5.251]	[0.0558]	[0.0602]	[0.1825]	[0.1583]	[0.289]	[6.379]	[0.0635]	[0.0550]	[0.1807]	[0.2073]
Percent Asian	−0.2	0.169	−0.0399	−0.0875	−0.0227	0.1468	−0.159	−0.195	−0.064	−0.0869	0.064	0.0837
	[0.272]	[5.836]	[0.0591]	[0.0623]	[0.1759]	[0.1539]	[0.283]	[7.035]	[0.0685]	[0.0586]	[0.1708]	[0.1976]
Percent Native American	−0.1	2.986	−0.0036	−0.0506	0.0925	0.1673	−0.068	2.369	−0.0143	−0.0449	0.2029	0.1285
	[0.233]	[4.594]	[0.0519]	[0.0925]	[0.2663]	[0.1807]	[0.256]	[4.881]	[0.0555]	[0.0827]	[0.2761]	[0.2390]
Percent free lunch	0.003	0.056	0.0033	0.0039	−0.0027	−0.0028	0.005	0.053	0.0033	0.0038	−0.0029	−0.0022
	[0.004]	[0.085]	[0.0009]	[0.0014]	[0.0016]	[0.0022]	[0.004]	[0.107]	[0.0010]	[0.0013]	[0.0020]	[0.0023]
Percent LEP	−0.002	−0.084	−0.0015	−0.0013	−0.0084	−0.001	0	−0.012	−0.0014	−0.0008	−0.0072	−0.001
	[0.007]	[0.278]	[0.0007]	[0.0017]	[0.0044]	[0.0036]	[0.006]	[0.329]	[0.0010]	[0.0018]	[0.0038]	[0.0034]

Name /jhr453/JHR45_03_A03/Mp_637 04/21/2010 09:00AM Plate # 0 pg 637 # 47

Log Mobile	0.271	3.277	−0.0461	−0.0094	0.1362	−0.025	0.281	3.53	−0.0517	−0.0201	0.1444	0.0218
	[0.163]	[2.716]	[0.0240]	[0.0404]	[0.0970]	[0.1475]	[0.170]	[3.081]	[0.0244]	[0.0396]	[0.1044]	[0.1490]
Lag log of Grade 11	−0.414	1.079	0.3916	0.3508	0.5307	0.4168	−0.455	0.843	0.4247	0.363	0.5489	0.4342
	[0.249]	[6.353]	[0.0377]	[0.1359]	[0.1143]	[0.1698]	[0.258]	[6.145]	[0.0373]	[0.1484]	[0.1243]	[0.2030]
Log Grade 12	0.243	3.926	0.3891	0.4789	0.4771	0.4419	0.247	3.712	0.3441	0.4368	0.4382	0.4389
	[0.180]	[3.106]	[0.0630]	[0.2498]	[0.1579]	[0.2584]	[0.158]	[3.254]	[0.0623]	[0.2799]	[0.1763]	[0.2817]
Lag log of Grade 10	−0.209	−9.844	–	–	–	–	−0.187	−11.096	–	–	–	–
	[0.232]	[5.447]	–	–	–	–	[0.240]	[5.634]	–	–	–	–
Log grade 11	0.285	1.13	–	–	–	–	0.311	2.339	–	–	–	–
	[0.315]	[7.654]	–	–	–	–	[0.291]	[8.844]	–	–	–	–
School effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Decile-year effects	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES
Schools	57	57	57	57	57	57	57	57	57	57	57	57
Observations	318	222	578	578	578	226	318	222	578	578	578	226

Robust standard errors in brackets. Standard errors are clustered at the school district level. Grade 10, Grade 11, and Grade 12 in the variable name indicates Grades 10, 11, and 12 enrollments respectively.

Name /jhr453/JHR45_03_A03/Mp_638 04/21/2010 09:00AM Plate # 0 pg 638 # 48

Table A2
Estimation of the Propensity Score

Variable	Treated
Midsized 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]
Percent white	−0.666
	[0.756]
Percent Native American	0.419
	[0.755]
Percent Asian	−0.675
	[0.756]
Percent limited English proficient	0.072
	[0.007]
Percent black	−0.626
	[0.756]
Midsized city*percent black	−0.049
	[0.010]
Urban fringe of large city*percent black	0.01
	[0.007]
Urban fringe of midsized city*percent black	−0.046
	[0.014]
Rural*percent black	0.033
	[0.014]
Percent Hispanic	−0.644
	[0.756]
Midsized city*percent Hispanic	−0.054
	[0.012
Urban fringe of large city*percent Hispanic	0.016
	[0.009]
Urban fringe of midsized city*percent Hispanic	−0.079
	[0.020]
Rural*percent Hispanic	0.015
	[0.018]

(*continued*)

Name /jhr453/JHR45_03_A03/Mp_639 04/21/2010 09:00AM Plate # 0 pg 639 # 49

Table A2 (*continued*)

Variable	Treated
Percent economically disadvantaged	−0.032
	[0.005]
Midsized city*percent economically disadvantaged	0.039
	[0.012]
Urban fringe of large city*percent economically disadvantaged	−0.044
	[0.012]
Urban fringe of mid-sized city*percent econ. disadvantaged	0.046
	[0.017]
Rural*percent economically disadvantaged	−0.069
	[0.022]
log of Grade 12 enrollment	−0.044
	[0.369]
Midsized 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 midsized 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 1100/24 on SAT/ACT	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 1100/24 on SAT/ACT	0.045
	[0.114]
Second lag of number of graduates	1.151
	[0.515]
Observations	5888

Robust standard errors in brackets.
Model includes school and year effects.