Memorandum

To: Hadi Partovi, Code.org

CC: Trevor Packer, College Board

From: Jack Buckley, College Board

Date: 12/17/2014

**Re:**   **Preliminary Results of AP Computer Science Analyses**

Introduction

The purpose of this memorandum is to document some exploratory analyses linking participation in AP Computer Science to subsequent performance in SAT Mathematics and AP Calculus. Caveats:

1. All analyses are preliminary and have not gone through full Research replication and quality control. I take full responsibility for any errors in these analyses.

2. None of these analyses should be interpreted as causal. Although there appears to be a relationship between AP CS participation and subsequent outcomes, it is highly likely that this is the result of one or more omitted and confounding characteristics of students that are not able to be controlled for given this research design.

3. See 2 again.

Data
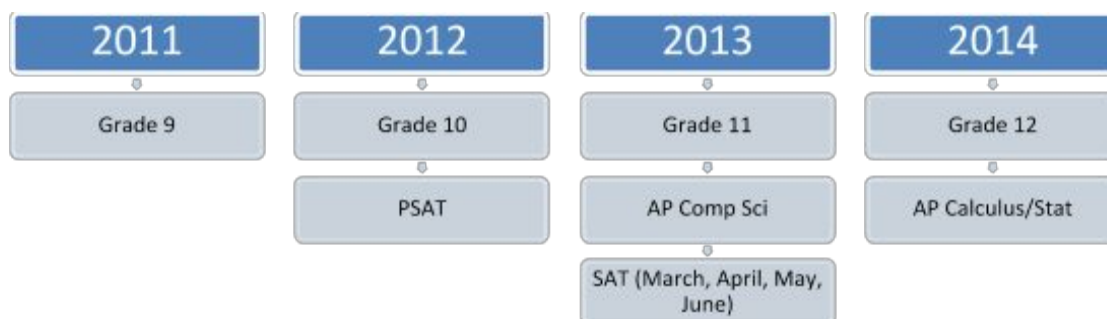
Data used in these analyses are for the graduating cohort of 2014. We matched PSAT and AP data to the SAT history file (which contains scores from all of the cohort's SAT administrations) after transposing all the scores for the SAT admins in the spring of students' junior year (i.e., as the AP course was ending).

Data elements:

- All AP scores for Computer Science (COMSCAGR) from the 2014 cohort, along with the year the students took the exam (COMSCAYR).
- AP Calculus AB, AP Calculus BC, and AP Statistics (both grade and year taken for all exams) as possible math outcomes.
- Math SAT scores from late in Junior year (March-June administrations) assuming effects would only be observed on SAT scores close to the completion of the 11th grade year.
- 10th grade P/N scores in math, verbal, and writing (pn_so_mrcn, pn_so_vrcn, pn_so_w) to use to control for prior achievement.
- Gender and ethnicity (9 category CB) for additional control variables.

Model 1: Simple linear regressions

The section reports on the results of two sets of regression models. First, we estimate:

$$SATM_i^{(a)} = \beta_0 + \beta_1 APCS_i + \varepsilon_i,$$

Where $SATM_i^{(a)}$ is student i's mathematics score on the SAT administered in a given month of 2013, a; $APCS_i$ is an indicator variable coded 1 if the student took the AP Computer Science exam before 2014 and 0 otherwise; and ε_i is normal, mean zero error term (so the β_1 coefficient in this model is equivalent to a t-test on a simple difference in means between AP-takers and non-takers).

We then estimate:

$$SATM_i^{(a)} = \beta_0 + \beta_1 APCS_i + \sum_{k=2}^{4}(\beta_k PSAT_i^{(k)}) + \beta_5 Female_i + \sum_{j=6}^{13}(\beta_j Race_i^{(j)}) + \varepsilon_i,$$

where the additional terms are the student's prior scores on the three sections of the PSAT/NMSQT (verbal, mathematics, and writing), an indicator variable whether or not the student is female, and a set of 8 indictor variables for the student's race/ethnicity with White (non-Hispanic) the excluded category. Partial results of the models (estimated via ordinary least squares) are presented in Table 1.

As Table 1 illustrates, the simple differences are all statistically significant at conventional levels and range from 131.7 to 198.6 points on the SAT Math scale.[1] This is well over one standard deviation and would be a very large impact *if* it were causally attributable to AP Computer Science participation.

TABLE 1: RELATIONSHIP BETWEEN AP COMPUTER SCIENCE AND SUBSEQUENT SAT MATH SCORES

	March	March	April	April	May	May	June	June
	AP CS Only	Full Model	AP CS Only	Full Model	AP CS Only	Full Model	AP CS Only	Full Model
AP CS	131.7	17.1	198.6	35.4	144.7	19.1	151.0	19.9
(Std. Err)	(1.84)	(0.93)	(13.39)	(6.53)	(2.12)	(1.07)	(2.06)	(1.06)
Adj. R^2	.027	.758	.010	.767	.018	.758	.020	.748
n		181068		21731		257672		265622

However, when we examine the results of the full model, including controls for prior PSAT scores as well as gender and race/ethnicity; we find a range of observed relationships between 17.1 and 35.4 SAT scale score points, all statistically significant at conventional levels. For example, students who took AP Computer Science before 2014, on average, score 17.1 points higher in Mathematics on the March SAT, controlling for prior achievement and some demographic

---

[1] Ordinarily, we use the most recent administration for each student as the best unbiased estimate of their SAT performance. Here, due to the preliminary nature of these analyses, I instead present results from each administration separately. The impact on the patterns observed here is minimal. Only partial results presented here; complete results available upon request.

characteristics.[2] These coefficients represent a range of effect sizes from .163 (March) to the unusually high .329 (April)[3], expressed in standard deviation units.

Even though these adjusted estimates are substantially smaller than the simple differences, it must be noted again that it is highly likely that there are unobserved and confounding characteristics such as subject matter interest, motivation, other student characteristics, or other school experiences that are correlated with both the decision to take the AP Computer Science exam *and* SAT Math performance. Accordingly these estimates should not be considered causal and, conservatively, be regarded as upper bounds of any true impact of AP CS on SAT Math.

Model 2: Ordinal logistic regression of other AP mathematics outcomes

It is also possible to consider the relationship between AP Computer Science course taking (or, here, exam taking as a proxy) and student performance on subsequent AP exams: AB Calculus, BC Calculus, and Statistics. For these outcome measures, however, it is not appropriate to estimate the simple linear regression models used above due to the ordinal nature of the AP scores (reported on the 1-5 integer scale).

As a more appropriate alternative, I estimate instead a series of ordinal logistic regression models of the form:

$$y_i^* = \beta_1 APCS_i + \sum_{k=2}^{4}(\beta_k PSAT_i^{(k)}) + \beta_5 Female_i + \sum_{j=6}^{13}(\beta_j Race_i^{(j)}) + \varepsilon_i$$

for a mean-zero error term with logistic distribution and an unobserved, latent variable y^* that maps to the observed AP grades, y:

$$y_i = 1 \quad if \ y_i^* \le \tau_1$$

$$y_i = 2 \quad if \ \tau_1 < y_i^* \le \tau_2$$

$$\vdots$$

$$y_i = 5 \quad if \ y_i^* > \tau_4,$$

where the τ are cutpoints to be simultaneously estimated along with the β's. The models were estimated using AP results for exams taken in 2014 examining students who took AP CS prior to

---

[2] As a limited sensitivity analysis, I also estimated propensity score matching models using the same covariates in the full regression model as predictors of the probability of taking AP CS, then matching AP CS student to their nearest neighbor on this metric for the March SAT taker subsample. The estimated average effect of treatment on the treated (ATET) for this model is 20.80 (1.16, $p < .01$) scale score points, compared to 17.1 in the regression model. This suggests that the estimates presented in table 1 are not sensitive to the linearity assumption of the simple model, but does not provide any additional robustness to omitted variable bias.

[3] The April administration represents data from only 21,731 relatively high ability students.

2014. Partial results of these models are presented in Table 2, below (complete results on request):

	AB Calculus	BC Calculus	Statistics
AP CS	.521	.343	.466
(Std. Err)	(.048)	(.047)	(.049)
Psd. R^2	.095	.088	.189
n	82145	33581	56411

In each of the models, the coefficient on AP Computer Science is positive and statistically significant at conventional levels, indicating a positive relationship between taking this exam and subsequent performance in other related AP subjects.

Although it is possible to interpret these coefficients on a odds ratio metric, it is generally more informative to consider differences in predicted probabilities of outcomes for the various categories. One method of this is to estimate predicted probabilities via stochastic simulation of the outcomes based on the estimates from the ordinal regression models and holding the various coefficients at values of interest. Here, for example, we can hold the PSAT/NMSQT scores at the sample mean values, and generate predicted probabilities of AP grades of 1-5 on the three other exams for white, male students (the modal type of student in these cases) for students both with and without AP Computer Science experience and then compare the differences. Figure 1, below, presents the results of these predictions.
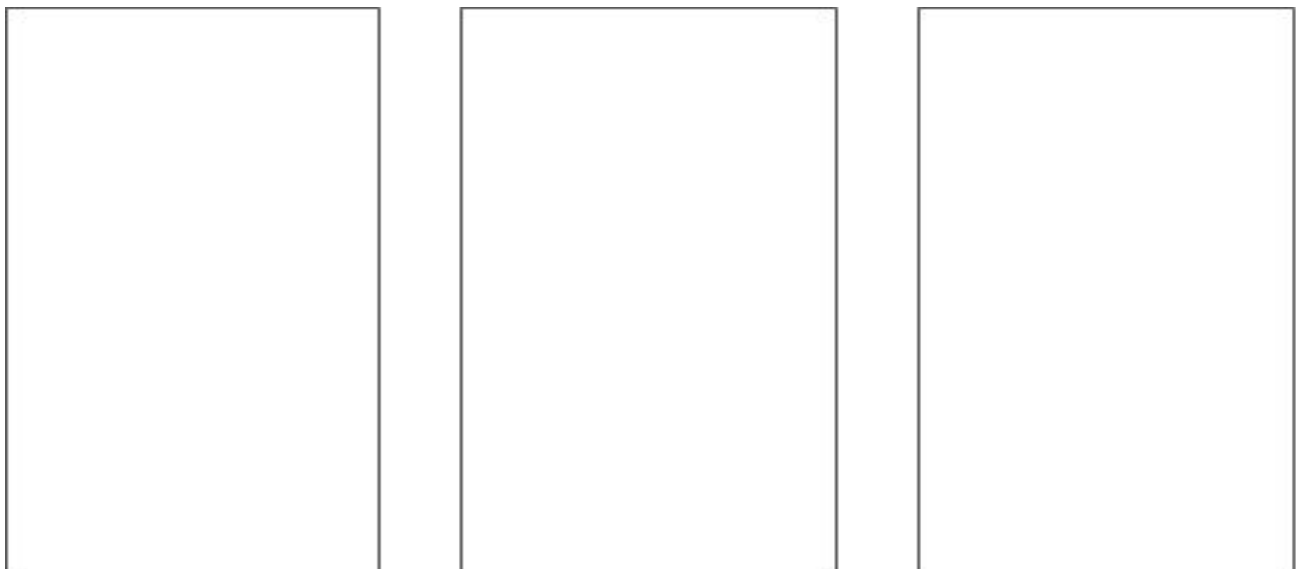
As the figure shows, the estimated "impact" of AP CS coursetaking (as proxied by AP CS exam participation) is somewhat heterogeneous across the three outcome courses—partly due to the very different score distributions for the subjects—but is generally positive in that students tend to be less likely to have predicted scores of 1 and 2 and more likely to have predicted scores of 4 and 5 (5 only for BC Calculus). In all three cases there are fewer scores of 3 for the AP CS students compared to their similar counterparts without AP CS.

As in the case of the linear regression models above, these results cannot be interpreted as causal. Although they suggest a relationship between AP CS coursetaking and subsequent AP success, it is again highly likely that there are additional confounding factors omitted from the model.