ONLINE APPENDIX FOR
Improving Academic Performance through Conditional Benefits: Open/Closed Campus Policies in High School and Student Outcomes

Shirlee Lichtman-Sadot*

July 2016

I am grateful to Caroline Hoxby, Ran Abramitzky, Matthew Harding, Moshe Justman, Roy Mill, John Pencavel, Karine van der Beek, participants at Ben-Gurion's Economics Department Brown Bag Seminar, the Hebrew University's Applied Microeconomics Seminar, and the 2014 AEFP annual conference for their helpful advice and comments. Noah Khassis provided valuable research assistance. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. 630714.

*Dept. of Economics, Ben-Gurion University, Beer-Sheva 84105, Israel. E-mail shirlees@bgu.ac.il.

Table A.1: Characteristics of High Schools who Responded to Survey vs. High Schools Targeted Initially

Variable	Survey Respondents	Target Schools w/o Response	P-Value for T-Test - Differences
Number of Schools	476	603	
Number of Students in School	1570.5	1624.7	0.356
	(878.89)	(1011.21)	
Free and Reduced Price Meal	0.434	0.552	0.000
Eligibility - October 2011	(0.258)	(0.264)	
Charter Schools	0.057	0.088	0.052
	(0.232)	(0.283)	
Fraction Hispanic - 2011	0.409	0.534	0.000
	(0.260)	(0.277)	
Fraction Black - 2011	0.061	0.078	0.010
	(0.091)	(0.116)	
Fraction White - 2011	0.359	0.250	0.000
	(0.255)	(0.242)	
Dropout Rate, 9th Grade - 2011	0.731	1.279	0.000
Weighted Mean	(1.023)	(1.733)	
Dropout Rate, 10th Grade - 2011	0.858	1.455	0.000
Weighted Mean	(1.169)	(2.033)	
Dropout Rate, 11th Grade - 2011	1.180	1.948	0.000
Weighted Mean	(1.604)	(2.568)	
Dropout Rate, 12th Grade - 2011	3.764	5.631	0.000
Weighted Mean	(3.905)	(5.677)	
English Language Arts Z-Score -	0.167	-0.013	0.000
2011 Weighted Mean	(0.387)	(0.387)	

Notes:
"Survey Respondents" includes all schools that fully responded to the survey, even if they were eventually excluded from the sample (e.g., if the school changed a location during the sample period). Weighted means are weighted by the number of students in each grade-school level. T-tests for differences between weighted means are conducted by running a weighted regression with the group's mean as the dependent variable, and running a t-test for whether the coefficient for being in one group is different from zero.

A Comparing High Schools which Completed the Survey and the Target High Schools

Table A.1 compares several key characteristics/variables between high schools which completed the survey and high schools in the target school list to which the survey was sent. Using figures mostly as of 2011, it can be seen that the average enrollment is very similar between the survey respondents and the target high schools. The average free and reduced price meal eligibility is lower among the survey respondents, in comparison to the target high schools, and charter schools were less likely to respond to the survey, compared to their representativeness in the target school list. Also, 2011 English Language Arts CST scores and dropout rates show that Survey Respondent high schools had higher academic performance.

B Determining the Sample of Tests and Grades for the Test Score Analysis

The CST exams in California high schools are conducted in several subject areas. The sample of tests and grade-levels used for the test score analysis needed to be determined such that: a) a sufficient amount of students took the exam (so that the pool of grade-school observations is sufficiently large); and b) the pool of test-takers is not responding to different open/closed campus policies in place. For the former condition, data from the CDE of the average percent of test-takers out of total enrollment at the grade-level in California for the period 2006-2011 was collected for each test-grade combination. Based on this, an initial pool of possible test-grade combinations was selected based on a minimal test-taker threshold of 30%. For the latter condition, regressions were run separately for each relevant test and grade-level, using the following specification:

$$\begin{aligned} PercentTestTakers_{ys} &= \alpha_0 + \alpha_1 YrOpenNoCond_{ys} + \alpha_2 YrOpenCond_{ys} \\ &+ \gamma_1 X_{ys} + yr_y + scl_s + \varepsilon_{ys} \end{aligned} \tag{1}$$

In equation (1), the dependent variable is the percent of students out of total grade-level enrollment within a school taking the a specific CST exam (one regression for each exam-grade combination). Our variables of interest in the regression are the indicator variables $YrOpenNoCond_{ys}$ and $YrOpenCond_{ys}$, which are equal to 1 when a given grade-school unit experienced either an unconditional or conditional open campus policy, respectively. Statistical significant coefficient estimates for these variables indicates that the percent of test takers for the specific exam-grade being analyzed is changing in response to either an unconditional or conditional open campus policy, in comparison to a closed campus policy (the omitted policy variable). A statistically significant difference between the two coefficient estimates is also an indication that the percent of student test-takers may be responding to different open/closed campus policies. Thus, for an exam-grade combination to be included in the test score analysis, we need the two coefficient estimates to not be statistically significant and the p-value for the test of differences between these coefficient estimates to be relatively large.

The results of the specification in equation (1) for the ten exam-grade combinations satisfying the 30% minimal threshold for overall percent of test-takers within California are presented in Table A.2. As can be seen, the test for statistical insignificance is passed for the following: ELA (grades 9-11), Science (10^{th} grade), World History (10^{th} grade), U.S. History (11^{th} grade), Geometry (10^{th} grade), and Biology (10^{th} grade).[1]

[1]While Algebra I coefficient estimates for the 9^{th} grade are not statistically significant, the large magnitude of the coefficient estimates, along with the relatively low p-value for the test of statistical significance between the two coefficient estimates for the conditional and unconditional open campus policy, deemed this exam to be excluded from the sample of tests. The results presented were very similar when this test was also included in the analysis.

Table A.2: Relation between Percent of Test-Takers at the Grade-School Level and Open/Closed Campus Policies

	ELA - 9th Grade	ELA - 10th Grade	ELA - 11th Grade	Algebra I - 9th Grade	Geometry - 10th Grade	World History - 10th Grade	U.S. History - 11th Grade	Biology - 9th Grade	Biology - 10th Grade	Science - 10th Grade
Unconditional Open Campus	-4.274	1.820	-1.161	-5.739	0.667	-0.122	-1.055	-11.814*	-1.240	0.496
	(3.712)	(5.212)	(1.602)	(4.594)	(2.803)	(3.305)	(1.785)	(6.639)	(4.390)	(11.237)
Conditional Open Campus	-8.366	3.674	-1.712	10.084	-0.727	-1.941	-0.857	-26.316	-2.591	4.252
	(8.889)	(6.760)	(1.976)	(10.571)	(5.472)	(7.407)	(2.043)	(16.364)	(7.049)	(11.357)
P-Value for Difference between Coefficient Estimates	0.478	0.821	0.769	0.142	0.801	0.809	0.919	0.240	0.807	0.661
Mean of Dependent Variable	95.14	94.20	90.45	48.42	32.92	90.44	89.21	40.56	50.23	93.62
Observations	3,823	3,730	3,517	3,751	3,602	3,716	3,517	3,031	3,584	2,533
R^2	0.48	0.31	0.32	0.63	0.44	0.31	0.35	0.77	0.71	0.34

Notes:
The dependent variable is the percent of students out of total enrollment at the grade-level taking the test for the relevant test-grade-level combination stated in each column. Regressions are run at the school-year level (for specific grade-exam combinations). All regressions include school or school-grade level characteristics (racial composition, free and reduced price meal population, total enrollment at school and grade level), school fixed effects and year fixed effects. Data is for the period 2003-2011. Standard errors (in parenthesis) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

C Open Campus Policies and Student Behavior Outcomes

An effort was made to uncover some of the potential underlying mechanisms through which students improve or worsen their academic performance under different open/closed campus policies. The California Healthy Kids Survey (CHKS) provides a unique opportunity to assess this question. It is the largest statewide survey covering students' risky behavior in the United States. At the high school level, $9^{th}$and 11^{th}grade students throughout California answer the survey anonymously on a bi-annual basis.[2] The survey includes a myriad of questions regarding the student's physical and emotional health, perception of the school environment, relationships with peers and adults, and risky behavior. The data available are at the individual student level for the period 2005-2011, with students represented in 405 of the schools in the sample.[3] By default, all students in the schools selected to take part in the the survey take the survey. Parents can request their child to opt out of the survey by providing a letter. I omit from the analysis students not answering or answering negatively a question at the end of the survey as to whether they answered the questions in the survey honestly (~75,000 observations, or a little under 11% of the total sample, but generally approximately 5% of the sample of students who provided answers to the questions used as dependent variables in the analysis) and students who reported an age which is 2 or more years older/younger than the required age for their grade-level (<5,000 observations).

I evaluate changes in student behavior under a conditional or unconditional open campus policy, in comparison to a closed campus policy, using six outcomes in the CHKS: whether the student reported skipping or cutting classes in the last 12 months; whether the student reported smoking marijuana at least once in the last 30 days; whether the student reported being afraid of getting beaten up or actually beaten

[2]The survey is conducted every year among altering school districts, such that each school district only runs the survey every other year.

[3]California school districts are required to participate in the survey, according the California Department of Education. However, large school districts can select a subset of their schools to participate in the survey, so survey coverage may not be complete across all California schools.

up in school in the last 12 months (two separate outcomes); whether the student reported feeling "very safe" or "safe" in their school; and whether the student reported being a member in a gang. The estimating equation is the following:

$$\begin{aligned} CHKSOutcome_{iysg} &= \alpha_0 + \alpha_1 YrOpenNoCond_{ysg} + \alpha_2 YrOpenCond_{ysg} \\ &+ \gamma_1 X_{ysg} + \gamma_2 W_{ys} + V_{iysg} + yrgr_{gy} + grscl_{sg} + \varepsilon_{iysg} \end{aligned} \tag{2}$$

Equation 2 is at the individual student level (subscript *i*). $Outcome_{iysg}$ is an indicator variable equal to 1 if the student positively responded to select questions on the CHKS. All variables are as defined for the equations in the paper and standard errors are clustered at the school level. The regressions are not weighted, due to data being at the individual student level. V_{iysg} represents variables at the individual student level - whether the student ever got drunk, whether the student ever got high, whether the student reports being a gang member (when this is not a dependent variable), and whether the student defines herself as receiving mostly A's and B's or mostly C's and D's (2 separate variables).[4] Although the dependent variables in all regressions are dummy variables, I run linear regressions, as my independent variables of interest are similar in their nature to interaction terms (being observed in a grade-school unit which experienced a specific policy and after that policy is in place), and as such, nonlinear regressions (such as probit) would bias the coefficient estimates (Ai and Norton (2003)). The regression specified in equation (2) was run, while controlling for fourth-order polynomials in time for students in grade-school units which experienced a change in their policies during the sample period and based on the grade-school unit's demographic characteristics during a base year, as detailed in the paper. The regression results are presented separately for the male and female student population.

Table A.3 presents results for the CHKS analysis. The most evident result is the link between an open campus policy - regardless of whether it is conditional or unconditional - and a decrease in students' reporting being beaten up in school or afraid of being beaten up in school. This is particularly strong in magnitude and significant for the male student population. In terms of changes in risky behavior in response to a conditional/unconditional open campus policy: we do not see any changes in skipping/cutting classes, there is some evidence of a decrease in the likelihood of a female student to smoke marijuana in response to a conditional open campus, and for males, there is evidence of a decrease in gang membership in response to a conditional open campus policy. Table A.3 also suggests that a conditional open campus policy reduces the sense of safety in school for both male and female students. This may be consistent with a conditional open campus policy maintaining the more problematic students within school premises.

[4]Race or ethnicity variables are only available for a subset of survey respondents in the CHKS data and thus could not be included as student-level control variables.

Table A.3: High School Student Behavioral Outcomes and Open/Closed Campus Policies - CHKS Analysis

Dependent Variable	Skip/Cut Classes - last 12 mos	Smoked Marijuana - last 30 days	Afraid of Being Beaten Up - last 12 mos	Beat Up in School - last 12 mos	School is "Safe" or "Very Safe"	Gang Membership
			Males			
Unconditional Open Campus	0.021	-0.008	-0.030***	-0.049*	0.012	-0.010
	(0.024)	(0.009)	(0.011)	(0.025)	(0.021)	(0.010)
Conditional Open Campus	0.001	0.007	-0.036**	-0.089**	-0.039	-0.036***
	(0.042)	(0.016)	(0.014)	(0.036)	(0.037)	(0.013)
P-Value for Difference between	0.567	0.265	0.560	0.106	0.084	0.004
Mean of Dependent Variable	0.367	0.157	0.163	0.303	0.632	0.097
Coefficient Estimates						
Observations	242,502	241,645	241,686	241,247	242,172	243,977
			Females			
Unconditional Open Campus	0.000	-0.005	-0.006	-0.032*	0.048***	-0.012
	(0.020)	(0.010)	(0.013)	(0.017)	(0.011)	(0.009)
Conditional Open Campus	-0.018	-0.023*	-0.027	-0.033*	-0.007	-0.003
	(0.032)	(0.014)	(0.020)	(0.018)	(0.026)	(0.010)
P-Value for Difference between	0.457	0.0592	0.217	0.908	0.017	0.039
Mean of Dependent Variable	0.407	0.128	0.160	0.203	0.610	0.055
Coefficient Estimates						
Observations	276,078	275,555	275,458	275,157	275,818	277,422

Notes:

Dependent variable is as stated on the top row. The sample is student CHKS respondents in the 9^{th} and 11^{th} grade for 2005-2011, excluding those not providing a positive answer for whether they answered the survey honestly and those who were 2 or more years away from the appropriate age for their grade-level. Top panel is only male respondents, bottom panel is only female respondents. Regression specifications are as in equation (2) with the following student-level controls (indicator variables): ever drunk, ever high, whether the student characterized herself as receiving "mostly B's and C's" or "mostly C's and D's", and whether the student characterized herself as receiving mostly C's or worse. When the dependent variable is not gang membership, a student-level control variable for gang membership is also included. Number of schools in all regressions is 405. Standard errors clustered at the school level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

D The Synthetic Control Method

The synthetic control method is most commonly applied in a difference-in-differences setting, where some observational units experience some treatment and are observed both before and after the treatment, while other observational units serve as controls, due to no treatment experienced by them during the sample period. As a first step, this method splits the sample of observations into treated units and control units, and then it constructs a control group for each treated unit by assigning weights to each control unit such that its pre-intervention characteristics and outcome variable resemble those of the treated unit. The objective is to choose a vector of weights which minimizes the distance between the pre-intervention characteristics of the treated unit and the characteristics of the (weighted) synthetic control group. This method is most useful when there are considerably more control units than treated units. While the analysis presented in for test scores and dropout rates is not a straightforward difference-in-differences analysis, I can still apply the synthetic control method by defining my treated units as those grade-school units which experienced a change in their open/closed campus policy during my sample period and my control units as those grade-school units which had the same open/closed campus policy throughout the entire sample period.[5] While in Abadie and Gardeazabal (2003) and Abadie et al. (2010) the synthetic control method is used for a single treatment unit with one treated period, the method can be extended to interventions covering multiple treatment units with multiple intervention periods, as is the case when the intervention is defined as a grade-school unit's change in an open/closed campus policy.

E Correlation between Changes in High School Administration and Changes in Open/Closed Campus Policies

Probit regressions were run with the dependent variable being a dummy receiving the value one if a change took place in a school/grade-school's open campus policy that school year and the independent variable being a dummy indicating whether there was an administrative change in the school that same year, or the previous year. Regressions controlling for administrative changes from the previous school-year, rather than the current school-year, were specified to account for the possibility that it may take a year to implement drastic changes within the school, such as changing its open/closed campus policy. The results, presented in Table A.4, show that while there is no correlation between experiencing an administrative change in a specific year and experiencing a change in an open/closed campus policy that same year, there is a very strong and positive correlation between a school experiencing an administration change the pre-

[5] Unlike in the standard difference-in-differences analysis, the policy changes in this analysis are not in one uniform direction. In particular, the grade-school units in the sample experience the following possible policy transitions: closed campus to unconditional open; closed campus to conditional open; unconditional open to conditional open; and unconditional open to closed campus. If treatment is defined as experiencing a change in the open/closed campus policy, then in this analysis, the coefficients of interest do not estimate the actual treatment (i.e. changing the policy during the sample period) but rather the effect of a specific policy which is mutual to some treated units, as well as some control units.

Table A.4: Correlation between Changes in Open/Closed Campus Policy and Administrative Changes

Dependent Variable: Change in Open/Closed Campus Policy	School Level	School Level	School-Grade Level	School-Grade Level
Administrative Change This Year	-0.002		-0.004	
	(0.008)		(0.003)	
Administrative Change Last Year		0.022**		0.015***
		(0.011)		(0.004)
Fixed Effects	Year	Year	Year	Year
Number of Observations	2290	1832	9160	7328

Notes:
The dependent variable is a dummy variable for there being a change in the open/closed policy. "Administrative Change This Year" represents a dummy variable receiving the value of 1 if an administrative change occurred during the year of the observation. "Administrative Change Last Year" represents a dummy variable receiving the value of 1 if an administrative change occurred during the previous year. The first two columns present results from regressions run at the school level and the last two columns present regressions run at the school-grade level. The point estimates are from probit regression, and they represent the marginal change in the probability of changing the open/closed campus policy when a new administration is introduced. *** p<0.01, ** p<0.05, * p<0.1

vious school-year and experiencing a change in its open/closed campus policy, compared to schools not experiencing an administrative change.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association 105*(490), 493–505.

Abadie, A. and J. Gardeazabal (2003, March). The economic costs of conflict: A case study of the basque country. *American Economic Review 93*(1), 113–132.

Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters 80*(1), 123–129.