SSP Data Summary for Researchers
(2009-2014 SSP data file provided by Gary L. Bowen for use by researchers)

Dr. Gary Bowen has provided access to an SSP data file containing records from schools and programs that administered the SSP between the 2009-2010 and 2013-2014 school years. Over 5,000 students (N=5,171) in grades 6 through 9 and from 17 schools in two communities in North Carolina are included in the dataset for researchers. Schools were included in this dataset if at least 50% of their students were completed the SSP. Schools in the first community used the SSP as part of a school district effort to use data to inform school-level interventions addressing student needs. This district used Opscan forms for the data collection and all 6th, 7th, 8th, and 9th graders in participating schools were targeted to take the SSP. Schools in the second community also collected SSP data as part of efforts to obtain data to guide strategies to support students. Participating schools elected to take part. Students in this community took the SSP online.

Information on the Sample

Thirteen of the schools were in one northwestern community in North Carolina (community 1 in the SSP dataset). The community is a rural county with about 10 municipalities. (Here and throughout, approximate figures are given to maintain the confidentiality of the data sources.) In the year of the data collection, the population of the community was approximately 80,000 was about 90% White. Median household income was about $35,000; per capita income was about $20,000; and about 20% of the population lived below the poverty line. Data were collected from schools in Community 1 during the fall of the 2009-2010 school year.

Four of the schools were from a primarily urban community (community 2) of over 230,000 people in the central part of North Carolina. The population in the years of data collection was about 40% White and about 40% African American. Median household income was about $45,000; per capita income was about $25,000; and about 20% of the population lived below the poverty line. Data from 3 schools in this community were collected during the 2012-2013 school year. In the fourth school, data were collected during the 2013-2014 school year.

Table 1 provides demographic features of the overall sample of 5171 cases.

Table 1: Characteristics of the Full SSP Sample Overall

	Numbers (%)
Community 1	3650 (70.6)
Community 2	1521 (29.4)
Girls	2573 (50.0)
Boys	2575 (50.0)
Free Lunch Participation	
(data available for Community 2 only)	897 (59.9)
African American	883 (17.1)
Latino*	730 (14.2)
White	3114 (60.5)
All other	422 (8.2)
6th grade	1425 (27.6)
7th grade	1215 (23.5)
8th grade	1346 (26.1)
9th grade	1177 (22.8)

*See information on coding of race/ethnicity in Notes on the Coding of Variables section below.

Table 2 provides information about cases in each school in the dataset, including the grade levels of students in the dataset, the total number of students who took the SSP at the school; and the percentage of students taking the SSP out of the number of students who completed mandated standardized testing at the school. Addition columns provide some indication of the representativeness of cases in the SSP dataset relative to the characteristics of the school as a whole. For community 1, which had few African American students and in which free lunch data were not obtained from students, information is presented on the percentage of Whites in the sample and the percentage of Whites who took standardized tests at the school in the Spring of 2010. For community 2, which had a more racially/ethnically diverse population, information is presented on the percentage of African Americans in the sample and among test-takers as well as on the percentage of free/reduced lunch program participants among SSP respondents compared to the test-takers. Researchers using the dataset or portions of it will need to make their own defensible arguments about the representativeness of the cases used in their analyses relative to the school-level information.

Table 2
Sample Size and Composition (selected variables)
of SSP Sample and School Population

Schools with * had MS plus 9th graders	SSP (n)	Test-takers at school (n)	SSP n as % of Test takers at school	% White in SSP sample	%White in test-takers at school	Free lunch in SSP sample	% Free lunch in test-takers
Community 1							
School 1	136	142	0.96	94.1	95.1	-	-
School 2	117	121	0.97	94.9	99.0	-	-
School 3*	210	236	0.89	69.9	70.3	-	-
School 5	722	784	0.92	88.6	93.0	-	-
School 6	73	74	0.99	87.7	94.0	-	-
School 7	477	478	1.00	60.0	65.0	-	-
School 8	595	664	0.90	86.2	93.0	-	-
School 9*	215	282	0.76	75.3	78.4	-	-
School 10*	58	72	0.81	86.2	94.4	-	-
School 12*	351	427	0.82	86.6	92.0	-	-
School 14	537	553	0.97	65.0	69.0	-	-
School 15	74	75	0.99	94.6	94.6	-	-
School 17*	85	86	0.99	85.7	84.0	-	-
	SSP (n)	Test-takers at school (n)	SSP n as % of Test takers at school	% Black in SSP sample	% Black in test-takers at school	Free lunch in SSP sample	% Free lunch in test-takers
Community 2							
School 4	526	604	0.87	35.6	47.6	64.5	72.5
School 11*	257	444	0.58	27.3	40.5	26.3	51.4
School 13	389	486	0.80	66.5	77.0	62.3	71.0
School 16	349	621	0.56	53.6	68.3	75.4	79.0

Technical note: The denominator of ratios used to calculate percentages of SSP students out of all students in each school came from publicly available, archived dataset retrieved in March 2017 from http://accrpt.ncpublicschools.org/docs/disag_datasets/. We chose to use the test-taker data because we expected students eligible to take standardized tests to most likely to be included in the SSP data

collection, and because additional information is available and comparable on test-takers in state databases. There are typically only minor differences between enrollment and test-taker counts. Within the statewide dataset for each year SSP data were collected, we determined the number of students in each school with SSP data who took the "Multiple choice" (MC) standardized test in math and reading. At times there were slight discrepancies ($n < 5$) in the number of students who took math versus reading tests; in such cases we used the larger number to ensure we did not overestimate the percentage of test-taking students who were represented in the SSP dataset. In two schools, the number of students with SSP data slightly exceeded the number of test-takers (most likely due to students leaving the schools before standardized testing occurred). In those cases we used the total enrollment data of the schools to obtain percentages below 100%. Effects were minimal. Counts of students taking retests, which are available for some years in North Carolina's test data banks, were *not* included in our calculation of the percentage of students in each school who were represented in SSP data.

Notes on the Coding of Variables

Researchers using the SSP 2009-2014 database should refer to the SSP Questionnaire Booklet for the exact wording of items and response options. The SSP dataset also includes brief labels and response options for all values of variables with Variable View.

The SSP asks students if they are Hispanic/Latino in a separate question before the question about race. In our recoded race/ethnicity variable (Race4), any student who selected Hispanic as his/her ethnicity, is coded as Hispanic in the race variable (regardless of responses to the subsequent race question). Users may prefer to code race/ethnicity differently using the original SSP variables.

All missing values are coded as 99 in the SPSS dataset.