

Appendix B

Final Sampling Plans for the Teacher and Parent Surveys

Final Sampling Plan for the Teacher Survey

Basic Sample Design

The basic design for the survey is stratified systematic sampling. The stratification is done by school type and school level. The school type here is the distinctive nature of schools regardless grade level. The school types used for sampling purposes are regular (zoned) schools, city-wide schools (all of which are high schools), alternative schools, and special education centers. Furthermore, schools adopting specific reform models, e.g., Direct Instruction (DI) and Achievement First (AF) programs, are separated out to create their own stratum. The DI/AF stratum is divided into two substrata according to the program. Regular schools are further stratified by level, namely, elementary, elementary/middle, middle, and high. This stratification has created the following categories:

Stratum Number	Description
11	Regular Elementary Schools
12	Regular Elementary/Middle Schools
13	Regular Middle Schools
14	Regular High Schools
21	AF Schools
22	DI Schools
30	City-wide Schools
40	Special Education Centers
50	Alternative Schools

Eligibility Criteria

Teachers in schools with lower grades (Pre-K–1 and Pre-K–2) are not eligible. Teachers from the Upton School (#303) for students recuperating from hospital stays are also excluded, as are those not assigned to a specific building. Moreover, schools to be closed soon are also deleted. In addition, we include only teachers whose primary subject or secondary subject is as follows:

Subject Code	Description
0703	Elementary Education
0801	English
0803	Reading
1200	Mathematics
1601	Biology
1602	Chemistry
1603	Earth/Space
1604	General Science
1605	Geology
1606	Physical Science
1607	Physics
1608	Other Science
1703	History
1706	Social Studies

1801	Special Education Infant – 3
1802	Special Education Grades 1 - 8
1803	Special Education Grades 6 – Adult
1804	Hearing Impaired
1805	Severely Handicapped
1806	Visually Impaired

Only three teachers are eligible because their secondary subjects are in scope, and they are all in middle or elementary/middle schools.

Besides special education teachers included in the above list, speech pathologists or therapists (position type 17) are also included. Their subject code is blank, but those individuals will be classified as special education teachers later.

Sort Variables

Within each stratum (except 30, 40, and 50), the teacher list will be sorted by various school and teacher characteristics, and then a systematic sample will be selected. School-level sort variables are as follows:

School Sort Variable	Description	Number of Categories
RE_C	Reconstitution eligibility	2 (Eligible/Ineligible)
AREA	Administrative area	8 (CA, SE, SO, NE, MC, NW, SW, NO)
SPI00_C	School performance index	2 or 3 (Definition below)
PCFREE_C	Percent free/reduced lunch	2 (Definition below)

All schools in administrative area DI are DI schools and, therefore, the administrative area code DI is redundant for other schools. Note also that 16 of the 19 schools in administrative area CA are AF schools. Seven schools that have newly been designated as reconstitution eligible in year 2001 are treated as reconstitution schools for sampling purposes.

In other strata, where the SPI00 variable is used as a sort variable, the distribution of the variable is very different from stratum to stratum, and categorization was done within each stratum separately as follows:

Stratum	No. of Categories	Boundary Points	Distribution (%)
11	3	30 and 50	34, 43, and 23
12	2	50	62 and 38
13	2	35	41 and 59
14	2	85	57 and 43
21	2	40	58 and 42
22	2	30	54 and 46

Note that the range between the smallest and the largest values of SPI00 in Stratum 11 is much wider than those for other strata. The equality goes to the upper boundary, a principle that applies to other categorization of continuous sort variables as well.

The PCFREE_C categories are defined by the boundary point 75, which roughly divides the teacher population into halves. This means that about 50 percent of teachers serve schools with more than 75 percent of students eligible for free or reduced-price lunch.

Originally, we planned to use percentage of special education students as a school sort variable, but it was deleted because special education teachers will be separately designated as such by a teacher variable. Teacher sort variables are as follows:

Teacher Sort Variable	Description	Number of Categories
SUB_C	Subject	2 (Special education/All others)
CRTYPE_C	Certification type	2 (Certified/Provisional)
YRSEXP_C	Years of experience	3 (Definition below)

Note that subjects are grouped into two categories: special education and all others. Special education includes six special education subjects (1801–1806, inclusively) and speech pathologists or therapists. Nine certification types are collapsed into two categories, which are defined as follows:

CRTYPE_C	Certification Code	Description
1	01	Resident Teacher
1	02	Standard Professional - 1
1	03	Standard Professional - 2
1	04	Advanced Professional
2	05	Provisional Degree
2	06	Provisional Non-Degree
2	07	Professional Eligibility
2	08	SPCI Renewal
2	09	Extended SPC 2

The years of experience variable is categorized by less than or equal to 3 years, between 3 and 20 years, and more than 20 years. Teachers are almost evenly distributed among these categories.

All sort variables will be used for Strata 11, 12, 13, 14, 21, and 22, but only teacher sort variables will be used for Stratum 30. The remaining strata (40 and 50) do not need sorting since all those teachers will be selected.

Sample Allocation

The sample allocation was performed so as to get a good precision level for each stratum. A precision level of 5 percent or less is considered good when a population proportion of 50 percent is estimated. Table 1 shows the sample allocation results.

The target sample size is the sample size needed to achieve the expected precision level. The SE column gives the expected precision level in terms of the standard error (SE) of an estimate for the population proportion of 0.5. For other proportions, the precision levels given in the table are conservative in the sense that the true expected precision level will be higher (i.e., smaller SE). It is easy to see that the sampling fractions are large. When the precision levels were computed, they had the effect of lowering the SE. However, whatever implicit stratification provided by sorting the list by those variables might bring to further lowering of the SE, it was not incorporated in the calculation.

The field sample size is obtained by inflating the target sample size to compensate for survey nonresponse assuming a 80 percent response rate. The field sample sizes thus calculated for Strata 40 and 50 were very close to the strata population sizes and thus were raised to the population sizes so that all teachers from the two strata would be selected with certainty. Beyond these field sample sizes, 20 percent more sample will be selected as reserve wherever possible and used as needed during data collection.

Table 1.—Sample Allocation and Expected Precision Level

Stratum		Teacher counts	Target sample size	Sampling rate (%)	Standard error (se)	Field sample size
11	Regular Elementary	1,355	200	14.8	0.033	250
12	Regular Elementary/Middle.....	410	150	36.6	0.033	187
13	Regular Middle	940	200	21.3	0.031	250
14	Regular High.....	574	170	29.6	0.032	213
21	AF Schools.....	593	100	16.9	0.046	125
22	DI Schools	347	100	28.8	0.042	125
30	City-wide Schools.....	317	120	37.9	0.036	150
40	Special Education Schools.....	109	80	73.4	0.029	109
50	Alternative Schools.....	116	80	69.0	0.031	116
Total/Overall.....		4,761	1,200	24.6	0.014	1,525

Final Sampling Plan for the Parent Survey

Target Population

The target population for this survey consists of parents or primary care givers of all the students currently enrolled in the Baltimore City Public School System, excluding students in the Pre-K–2 schools and excluding Pre-K students in schools that accommodate learning beyond the second grade.

Sampling Frame

The sampling frame will be a list of all the eligible students currently enrolled in the Baltimore City Public School System, with school and contact information (e.g., home address, phone number, etc.).

Stratification

The sampling frame will be stratified first by level of school (elementary, elementary/middle, middle, high, alternative, and special education centers), and then by type of program (AF, DI, other) within the strata of elementary schools and elementary/middle schools, and by city-wide status (city-wide vs. other), within the stratum of high schools (see Table 2).

Sample Size

The target sample size is 750 completes. Assuming that 70 percent of the addresses are valid and that 20 percent of the parents/primary care givers will respond to the survey, we need an initial sample of 5,357 parents or primary care givers to ensure 750 completes. To reach that goal, we will release the sample in batches. The first subsample will be 2,679 parents/care givers, based on the assumption that 70 percent of phone numbers are valid and that 40 percent of the eligible parents/care givers will respond. Considering the limited resources, we will make five attempts for each number at different times of the day and define that as our best effort. If five attempts have been made for each of the 2,679 numbers and we still haven't reached the target number of completes, we will release another subsample of telephone numbers following the same procedure as described above until 750 completes have been obtained.

Sample Allocation

To ensure representation of students in the different schools and/or programs, some of the strata will be oversampled (e.g., students in alternative schools, special education centers, or DI programs) while some other strata will be undersampled (e.g., elementary schools) (see Table 2).

Sample Selection

Within each of the 11 strata, students will be sorted by reconstitution eligibility (eligible/ineligible), administrative area (SE, SO, NE, MC, NW, SW, NO), school performance index (SPI),¹ percentage of students who are in the free/reduced-price lunch programs,² grade, and race/ethnicity. The sample will then be selected systematically within each stratum so that the sample distribution of students across the sorting variables/categories will be identical to that in the target population. This systematic selection will follow the formula $R+I, R+2I, R+3I \dots R+nI$, where R is the random start point on the list of students, and I is the interval of selection. The interval of selection is derived from N_h/n_h , where N_h is the total number of students in Stratum h , and n_h is the number of students to be sampled from Stratum h .

Table 2.—Population and Sample Distribution of Students Across Level of Schools and Programs

School type	Number of students in BCPSS	Percent of students in BCPSS	Estimated number in initial sample (n=5,357)		Estimated number of responses (n=750)	
			Proportionate allocation	Disproportionate allocation	Proportionate allocation	Disproportionate allocation
Elementary schools	41,953	44	2,357	1,679	330	235
AF.....	10,637	11	600	500	84	70
DI.....	3,886	4	214	286	30	40
Other.....	27,430	29	1,543	893	216	125
Elementary/ Middle schools	11,524	12	650	714	91	100
AF.....	592	1	36	29	5	4
DI.....	3,090	3	171	157	24	22
Other.....	7,842	8	443	529	62	74
Middle schools	16,886	18	950	821	133	115
High schools	22,371	23	1,257	1,429	176	200
City-wide	8,459	9	478	500	67	70
Other.....	13,912	15	779	929	109	130
Alternative schools	1,830	2	100	357	14	50
Special ed. centers	806	1	43	357	6	50
Total	95,370	100	5,357	5,357	750	750

NOTE: Details may not add to totals due to rounding.

Design Effect and Effective Sample Size

As shown in Table 2, the designed sample will be disproportionately allocated across the different strata. This process of allocating the sample disproportionately to their distribution in the target population will cause the designed sample to deviate from a simple random sample and thus produce design effect. This type of design effect can be calculated using the following formula:

$$Deff_{\text{disproportionate}} = W_h K_h W_h / K_h$$

¹ The same classification scheme as used in the Teachers Survey.

² The cutting point is 75 percent.

where W_h is the proportion of Stratum h in the total target population, and K_h is the relative weight of the cases in Stratum h . The design effect for estimating population parameters will be approximately 1.18. Taking into consideration this design effect, the effective sample size for the population estimates will be:

$$750/1.18=636.$$

This means that a total number of 750 responses obtained through disproportionate sample allocation will be equivalent to a simple random sample of 636 cases.

Sampling Precision

This design effect will also be taken into account in calculating the sampling precision for estimating population parameters for the total target population. Table 3 shows the 95 percent confidence intervals for the different estimated proportions based on an effective sample of 636. For example, if 50 percent of respondents report involvement in school activities, the 95 percent confidence interval for this estimated proportion will be $.50 \pm .039$. Likewise, if 25 percent or 10 percent of respondents report so, the 95 percent confidence intervals for these estimates will be $.25 \pm .034$ or $.10 \pm .023$. The widest confidence interval occurs when the estimated proportion is .50.

Table 3.—Precision Levels for Estimating Parameters of the Total Target Population (n=636)

Estimated proportion (P)	Sampling error (SE)	95% confidence interval
.10.....	.012	$\pm .023$
.25.....	.017	$\pm .034$
.50.....	.020	$\pm .039$

Table 4 indicates the 95 percent confidence intervals for making subpopulation estimates for the different types of schools. Based on our proposed sampling design, the widest 95 percent confidence intervals will be $.50 \pm .14$. Since the design effect is caused by the disproportionate sample allocation across the strata, it will not apply when making estimates for each of them. Finite population correction was used only for the stratum of special education centers because none of the sampling rates for all other strata exceed 5 percent, and thus the finite population factor can be neglected.

Table 4.—Precision Levels for Making Subpopulation Estimates (P=.50)

Subpopulation	Sampling error (SE)	95% confidence interval
Regular Elementary (n=125)045	$\pm .088$
Regular Elementary/Middle (n=74)....	.058	$\pm .114$
Regular Middle (n=115)047	$\pm .091$
Regular High (n=130).....	.044	$\pm .086$
AF Schools (n=74).....	.058	$\pm .114$
DI Schools (n=62)063	$\pm .091$
City-wide Schools (n=70).....	.060	$\pm .117$
Special Education Schools (n=50).....	.068	$\pm .134$
Alternative Schools (n=50).....	.071	$\pm .138$

