STRIVING FOR COST EFFICIENCY IN THE SAMPLE DESIGN FOR THE SURVEY OF YOUTH IN RESIDENTIAL PLACEMENT

Sheila Krawchuk, Gary Shapiro, and Andrea Sedlak
Sheila Krawchuk, Westat, 1650 Research Boulevard Rockville, Maryland 20850

KEY WORDS: Institutional Survey, Two-Stage Sample Design, Differential Sampling Rates, Geographical Clustering of Sampling Units.

1. Introduction

Although other surveys have collected information about youths in residential placements and the facilities they reside in, the Survey of Youth in Residential Placement (SYRP) is the first survey since 1987 to collect data directly from offender youths aged 10-20. Sponsored by the Office of Juvenile Justice and Delinquency Prevention (OJJDP) and conducted in spring 2003, the first of these national surveys collected data from youths by audio-computer-assisted self-interviewing. Data was collected about their offense histories, service needs, and custody experiences.

The main survey objective was to provide youth data that would complement two biennial facility surveys about youths in residential placement. Inherent in that goal was the aim of providing as much as possible, comprehensive coverage of youths in residential placement, including those in both public and private facilities. Providing reliable estimates by race/ethnicity and gender was also a survey requirement.

One of the challenges in meeting these objectives included the fact that interviewing could be costly given that over 75 percent of the universe facilities had fewer than 30 offenders. This is important because field staff visits each sampled facility, and even very small facilities with only a couple of youths take up most of a day for the staff. An additional challenge to the last objective was that both females and Hispanic males were rather rare but important youth populations. Compounding both of these challenges, smaller facilities differed in important ways from larger facilities. They contained a larger portion of females and tended to be private more often than public compared to larger facilities.

This paper details how these challenges were addressed, while still maintaining a balance between survey cost and data reliability. Section 2 discusses the basic design, while Section 3 discusses unusual aspects of the design which are the focus of this paper. Section 4 provides a brief conclusion.

2. Basic Sample Design

The SYRP sample was selected using a stratified, two-stage, probability-proportional-to-size (PPS) sample design. Residential facilities were selected at the first stage of selection from the same universe of facilities used for the most recent Census of Juveniles in Residential Placement (CJRP) (with updating). Offender youths were sampled within selected facilities at the second selection stage. Along with the cost efficiencies of this type of design, the desire was to have all sampled youths with about the same final probability of selection/weight. Two exceptions to this desirable feature were females and Hispanic males due to the need to oversample these two youth types to achieve the desired reliability of estimates within these subgroups.

Although early research indicated that a sample size of about 10,000 completed youth interviews would yield estimates with the required reliability, as already noted, it was additionally realized that females and Hispanic males would need to be oversampled because of their small population sizes. It was determined in the survey planning stages that females would be oversampled by a factor of 2.5 to 3 and Hispanic males by a factor of close to 1.5, compared to other male youths. Oversampling was accomplished in two ways. For the selection of facilities, facility size was estimated using information about the residents gathered during the most recent CJRP. However, rather than using just the number of youth offenders as the facility size measure for PPS sampling, females were given a weight of three, Hispanic males a weight of 1.5, and other males a weight of one. The facility measure of size then consisted of the sum of its youth weights. Thus, a facility with more females and/or Hispanic males than another facility with the same total number of youths would have a higher probability of selection. The youth sample was selected when sampled facilities were visited during data collection, using current information about the residents. Again, at this youth level, females were given a size measure of three, Hispanic males a size measure of 1.5, and other males a size measure of one. Youths within facilities were then selected with probability proportional to these measures.

Because of the complexity of facility recruitment, there is a substantial cost for each facility

even before any interviewing of youths is done. Thus, the cost per completed youth interview is much higher at small facilities than at larger facilities. To reduce costs, "small" facilities were undersampled by about a factor of three compared to those that had at least 80 youths (enough to provide an expected 60 completed interviews). As noted earlier, small facilities, more often private facilities than public, also have a higher proportion of females—one of the youth types that needed to be oversampled. However, investigation showed that the percentage reductions in sample sizes for both females and youths from private facilities were modest if small facilities were undersampled. Medium facilities were also undersampled, but to a lesser degree. To further reduce travel costs, small and medium facilities also were combined into groups of geographical proximity (see Section 3.2). If small or medium groups were sampled, all facilities within the selected groups were contacted and all eligible youths were sampled to be interviewed.

3. Special Aspects of the Sample Design to Reduce Costs

3.1 Excluding the Very Smallest Facilities

It was decided after the SYRP field test, where the very high costs per facility for recruitment and data collection were documented, that in addition to undersampling small facilities, the very smallest ones should be excluded.

Although it would have been simpler to just exclude from the sampling frame those facilities with fewer than x offender youths (x to be specified) according to their latest CJRP data, we did not recommend this simpler approach. More than three years would have elapsed between the time of the available CJRP data on number of youths in residence (October 1999) and the first national SYRP data collection (March 2003). A facility could have grown considerably during this interval. Furthermore, many facilities have widely fluctuating numbers of offender youths in residence. A facility with fewer than x offender youths on the census date for the CJRP might actually have averaged a larger number of youths during other recent time periods. As a result, the simpler procedure for excluding the very small facilities would run the risk of excluding some somewhat larger facilities from sample.

Instead, another method was decided upon. In this approach, all small facilities would be sampled according to the original plan. Then, all sampled facilities with fewer than y offender youths ($y > x$, y to be specified) according to the CJRP would receive a preliminary telephone call. The call would determine the number of offender youths in residence at the time of the phone call, and the maximum number of offender youths in residence at any point in time during the preceding month. Facilities were dropped from the sample, that reported:

- Having fewer than x offender youths;
- Having had fewer than x offender youths simultaneously in residence over the last month; and
- Expecting that nothing would change to substantially increase the number of their offender youth residents in the near future.

Thus, in order to be deleted from the sample, a facility had to be identified as "very small" by answers to all three questions. By this method, only the very smallest facilities would be unrepresented among the facilities that participate in the national SYRP.

As noted, the variables x and y above had to be determined. In doing so, we looked not only at different sizes of "very small" but also at simultaneously varying the undersampling rate (either two or three) of the small facilities. The values simultaneously considered for the variable x (less than this number of offenders in the facilities) were none, three, five, and six. For this exercise, "small" facilities were those defined as having a weighted size of fewer than 38.

Twelve computer program runs were done, one for each combination of an undersampling rate of the small facilities, and the x parameter. In each run, the number of facilities to be sampled to achieve a youth sample size of at least 10,000 was selected. Keeping in mind that results from these test samples were subject to sampling variability, each sample was examined to compare:

- The total number of facilities that would need to be sampled to achieve the youth sample size;
- The number of facilities and offenders excluded; and
- The differences in exclusion results for three types of public facilities that most often were "small" (halfway houses, foster care, and group homes) (type 1), other public facilities (type 2), and private facilities (type 3).

Table 1 shows the results for the number of facilities that would need to be sampled to achieve a

youth sample size of at least 10,000. The number of facilities needed, and the breakdown in sizes of those sampled indicated that having no undersampling of small facilities was not acceptable from the perspective of cost (about 250 sampled facilities or fewer was the desired sample size). Additionally, excluding no small facilities resulted in an unacceptably large sample size of 294 or more, as did excluding facilities with fewer than three with an undersampling rate of two (274 or more facilities). Thus, further consideration was given only to undersampling rates of two and three, and in particular, exclusions of facilities with three, five, or six offender youths for undersampling by a rate of three, and exclusions of five or six for undersampling by a rate of two.

Table 1. Number of facilities needed to be sampled, and by test sample size category

Under-sampling rate	Exclusion value of *x*	Total number of facilities sampled	Number of large sampled	Number of medium sampled	Number of small sampled
1	0	375	117	34	224
1	3	347	113	24	210
1	5	316	113	28	175
1	6	311	114	36	161
2	0	311	122	38	151
2	3	274	123	38	113
2	5	255	129	28	98
2	6	242	129	28	85
3	0	294	128	36	130
3	3	251	129	36	86
3	5	237	124	40	67
3	6	235	125	46	64

Table 2 shows the number of excluded facilities out of a universe of 2,847[1] and the estimated numbers of excluded youths for the five remaining options (based on average size). Based on these numbers, the (3,3) combination looked the most promising in terms of keeping the exclusions to an acceptable level, as it excluded fewer facilities than the other combinations, and just over 1 percent of offender youths (based on a 1999 CJRP offender youth population of about 106,000). The (3,3) combination had the smallest number of excluded facilities of both types 1 and 3—the types that could have been hurt the most by excluding very small facilities—as well as having the smallest number of excluded facilities overall. Thus, undersampling small facilities by a factor of three, and excluding facilities that answered affirmatively to the three questions noted above where the *x* variable was equal to 3, was the methodology adopted for excluding very small facilities in the field. The *y* variable noted earlier was set a bit above *x*=3 (to 5) so as not to be too restrictive in which sampled facilities underwent this extra screening.

Table 2. Number of excluded facilities and offenders

Under-sampling rate	Exclusion value of *x*	Number of excluded facilities	Estimated number of excluded youths
2	5	914	2,285
2	6	1,014	3,042
3	3	726	1,089
3	5	914	2,285
3	6	1,014	3,042

In the final SYRP sample, of the 25 facilities that were pre-screened for being potentially the very smallest, 15 were dropped from the sample.

3.2 Geographic Clustering of Small and Medium Facilities

As noted earlier, to further reduce travel costs, it was decided that both small and medium facilities would be grouped so that, if sampled, they would be within close traveling distance of each other. To begin, we needed to determine which facilities to classify as small and which ones to classify as medium. Small facilities were denoted as those with weighted size measures of less than 38. Medium facilities were those that had weighted size measures of between 38 and 59. Facilities that had weighted size measures of 60 or more were the large facilities.[2]

[1] Independent living facilities were excluded from this exercise.

[2] Our facility weighted size measures gave each female a weight of three, each Hispanic male a weight of 1.5, and each other male, a weight of one. The only exceptions to this were independent living facilities, and facilities with numbers of youths being either zero, one or two. These facilities were given weighted and actual size measures of 0.25. This was done because it was quite possible that these would be dropped out if sampled, and for grouping we didn't want them to add much to the group's number of youths. Independent living facilities were special facilities that were thought to have only very few youths. These should be caught by the "number of youths <3" rule, but since frames are not perfect, this wasn't always the case. Therefore, they were given the measure of size of 0.25 based on their facility type rather than based on the number and types of youths they had.

The usual FIPS codes consisting of numeric state and county codes could not be used for geographical grouping since the county numbers were assigned alphabetically as opposed to geographically. Thus, counties beginning with the letter "A" had smaller codes sequentially assigned than those beginning with the letter "B", and so on. Instead we used a Westat-developed SAS macro, WESPSU©, for geographical grouping of counties.

3.2.1 Small Facilities

Procedures for grouping were as follows: First, county level records were created which contained counts of small facilities and their youths. The WESPSU© macro was then used to group the 2,792 small facilities within state, so that the maximum distance between facilities within each group was 150 miles. The program was instructed that, in order to form a group, the group had to contain a minimum of 57 youths. We were striving for an average of 75 youths per small group (the number of youth interviews we originally wished to have in each large facility). If the minimum had been set to exactly 75, the groups would usually have been too large and would therefore have to be manually split. Investigation showed that the minimum of 57 would enable us to achieve our overall average of 75 youths per small group and would reduce the frequency with which groups had to be split manually.

Manual work was nevertheless needed to complete the grouping task. After using WESPSU©, there were facilities that still did not belong to any group. Additionally, there were some groups that we felt were too large, either based on the numbers of youths they had or on the numbers of facilities they contained. The latter was a concern, because field staff would have to visit all the facilities in any small group that was sampled.[3] Thus, although a very large part of the grouping work had been done by the program, a fair amount of manual work was still needed to complete the grouping of the small facilities.

The 2,792 small facilities resulted in 311 groups with an overall average number of 75 youths per group. The minimum number of youths in a group was 42 and the maximum was 114.

Including all facilities, the range of the number of facilities in a group was 2 to 29. With the 0.25 sized facilities ignored, the range was 2 to 16.

3.2.2 Medium Facilities

The 337 medium facilities were classed into 148 groups using the same program parameters used for the small facilities. This meant that there were usually only two or three facilities per medium group. The average number of youths per group was about 80, with a minimum of 48 and a maximum of 111. The average number of facilities per group was 2.5, with a range of one to five. Finally, for sample selection, it was noted that both small and medium facilities were undersampled. To accomplish this, each small group was assigned a measure of size of 25 (1/3 of the initial target of 75 sampled youths (later changed to 80). Medium facility groups each had a measure of size of 45. All large facilities had a weighted measure of size based on the total number of youths as well as the youth composition.

Of the 277 facilities selected for the final SYRP sample, 148 were large, 34 in 13 groups were medium, and 95 in 10 groups were small.

3.3 Strategies for Dealing With the Possibility of Sample in Sparse States

Early in the planning stages of SYRP, five states were noted as states of cost concern because of their sparseness of residential facilities and youths. Using the newest frame, several more were identified (see Table 3).

Our task was to determine how likely it was that facilities would be selected in these states. If the chances of this happening were reasonably high, a course of proactive action was suggested.

The proposed strategy for dealing with sparse states was that the small and medium facilities would be grouped as normal so that, if facilities were sampled, the sample would always contain more than one facility. For the larger facilities a different methodology was required to try to ensure that there would be either zero or more than two large sampled facilities from the state. The sample design included stratification, so facilities and facility groups were combined into strata. The number of facilities (groups) to be sampled would vary, depending on the size of the stratum. Consider those strata in which only one facility (group) was to be selected. Using State A as an example, each of these strata would be permitted to have no more than one State A facility (or facility group). In each stratum, any

[3] Complicating the concern of having too many facilities in any one group were the facilities given sizes of 0.25. Often it was the case that many of these tiny-sized facilities ended up in the same group. Since it was not certain that they would be dropped out if sampled, we tried to spread these out, whenever possible. Thus, in assessing group numbers of facilities, we observed the total number, as well as the number of non-0.25 sized facilities.

State A facility (group) would be listed first. Sample selection would be done for all strata together by determining a single random number. If the random number were very low, then the first listed facility (group) from every stratum would be in sample. In that case, there would be a large number of State A sample facilities, since a State A facility would probably be selected from every stratum that contained one. If the random number were not low, then no State A facilities would be sampled. The procedure for strata in which the number of sample facilities was more than one would be similar, although slightly more complicated.

To assess whether or not this methodology would need to be applied, we prepared the frame as if a facility sample of 167 facilities (expected 75 youths within each) would be selected so that probabilities of selection could be determined.[4]

We looked at two probabilities: the probability of selecting any facility in each of the problem states; and the probability of selecting in each state, exactly one large facility but no medium or small facilities. For each of the states, only four probabilities were ever defined as it turned out that the maximum number of large facilities, medium facility groups, and small facility groups was four. The probability of selecting any facility in each state was calculated as:

$$1 - (1 - p_1)(1 - p_2)(1 - p_3)(1 - p_4) \qquad (1)$$

where the p_1, p_2, p_3 and p_4 are the probabilities as defined in Table 3. The calculation for the second probability noted above was calculated by Equation 2 if there was only one large facility, and by Equation 3 if there were two large facilities.

$$p_1(1 - p_2)(1 - p_3)(1 - p_4) \qquad (2)$$

where p_1 is the probability of selecting a large facility, and the other probabilities are for selecting medium or small facilities.

If there are two large facilities with probabilities p_1 and p_2, the probability of selecting exactly one of them and no small or medium is as follows.

$$\left\{1 - \left[(1 - p_1(1 - p_3)(1 - p_4)(1 - p_2)(1 - p_3)(1 - p_4))\right]\right\} - p_1\, p_2\,(1 - p_3)(1 - p_4) \qquad (3)$$

Table 3 shows that for most states, the probability of selecting only one large facility into the sample was reasonably high when no special procedure was applied. Thus, it appeared that it might be worthwhile to apply the special methodology. However, when it came time to apply this strategy, it was decided that it would not be applied so as to keep the methodology simpler and to make it easier to explain and defend the sample design. We could have been unlucky. However, only three of the problem states had sample selected in them.

4. Conclusion

The sample design for the Survey of Youth in Residential Placement began as a straightforward stratified, two-stage, PPS sample. The oversampling of females and Hispanic males to fulfill data analysis objectives increased its complexity. The additional features to reduce costs—excluding the very smallest facilities, and grouping and undersampling small and medium facilities—both were successful in keeping costs down, but also added to the complication of the sample design. Thus, in this survey as in many surveys, necessary cost reducing measures usually contribute to a more intricate sample design than what was originally planned.

5. Acknowledgements

There were a number of people who made important contributions to the sample design for this survey. This includes Joe Moone, former project officer for Office of Juvenile Justice and Delinquency Prevention, members of the SYRP Advisory Board, staff at national Council on Crime and Delinquency, and additional Westat staff.

[4] Note that this exploration was conducted before final decisions on sample design were made. Therefore, in this design, small and medium facilities were grouped using FIPS codes (travel distance was not important for this exercise). Small facilities were undersampled by a factor of three and medium facilities were not undersampled.

Table 3. Estimated probabilities of selection for large facilities, and small and medium facility groups in cost problematic states

State	Number of facilities	Large facilities	Medium facility groups	Small facility groups	*P*(Large)	*P*(Medium)	*P*(Small)	*P*(None)	*P*(Exactly 1 large only)
A	27	1	1	2	P_1=0.291	P_2=0.146	P_3=P_4=0.05	0.454	0.224
B	7	2	1	1	P_1=P_2=0.146	P_3=0.146	P_4=0.05	0.408	0.206
C	17	1	0	1	P_1=0.291		P_2=0.05	0.326	0.277
D	10	1	0	1	P_1=0.146		P_2=0.05	0.189	0.139
E	25	2	0	1	P_1=0.291 P_2= 0.146		P_3=0.05	0.425	0.336
F	9	1	1	1	P_1=0.146	P_2=0.146	P_3=0.05	0.307	0.118
G	25	1	0	1	P_1=0.437		P_2=0.05	0.465	0.415
H	25	2	0	2	P_1=0.728 P_2=0.437		P_3=P_4=0.05	0.862	0.505
I	30	2	0	2	P_1=P_2=0.146		P_3=P_4=0.05	0.342	0.227
J	13	1	1	2	P_1=0.146	P_2=0.146	P_3=P_4=0.05	0.342	0.113
K	13	1	0	2	P_1=0.291		P_2=P_3=0.05	0.360	0.262
L	8	0	0	1			P_1=0.05	0.05	0
M	34	1	0	2	P_1=0.146		P_2=P_3=0.05	0.229	0.132
N	27	2	0	2	P_1=P_2=0.146		P_3=P_4=0.05	0.342	0.227