DUAL-FRAME RDD AND AREA SAMPLE FOR HOUSEHOLD SURVEY WITH PARTICULAR FOCUS ON LOW-INCOME POPULATION

Joseph Waksberg, J. Michael Brick, Gary Shapiro, Ismael Flores-Cervantes and Bridgett Bell, Westat, Inc.
Gary Shapiro, Westat, Inc., 1650 Research Boulevard, Rockville, Maryland 20850

1. Introduction

The National Survey of America's Families (NSAF) is part of a multi-year study to assess the New Federalism by tracking ongoing social policy reforms and relating policy changes to the status and well-being of children and adults. The major objective of the study is to assess the effects of the devolution of responsibility for major social programs such as Aid to Families with Dependent Children from the federal to the state level. The NSAF is collecting the information on the economic, health, and social dimensions of well-being of children, non-aged adults, and their families in 13 states that will be intensively studied as part of the project, and in the balance of the nation to permit national estimates. The 13 states were selected to vary in terms of their size and geographic location, the dominant political party, and key baseline indicators of well-being and fiscal capacity. A sample of the balance of the nation is included so that national estimates can also be produced. Low-income families are oversampled because the policy changes of interest are anticipated to affect them most. The initial round of the NSAF is taking place in 1997 and a follow-up round is planned for 1999 or 2000. Two rounds of case studies are occurring in parallel with the survey to provide a detailed understanding of the policy changes occurring in each of the 13 states. This study is being directed by the Urban Institute and Child Trends and is funded by a consortium of foundations, led by the Annie E. Casey Foundation. Westat is responsible for data collection and related activities.

This paper describes the overall sample design of the survey, with emphasis on two innovative features. One feature is the use of a random digit dialing (RDD) sampling frame to cover households with telephones in combination with an area sample to cover households without telephones. A second innovative feature was the exclusion from the area frame of those block groups that have very low rates of households without telephones to improve efficiency. Another interesting aspect of the NSAF sample design involved decisions regarding oversampling and subsampling. We discuss the decision to subsample households determined not to have low income in the screening interview and the decision to not oversample low income geographic areas.

Some important difficulties that were encountered in the conduct of the survey are also presented. The three key results are: 1) Lower than anticipated response rates; 2) Higher than expected misclassification of households by income; and 3) Lower than expected rates of nontelephone households in the area sample.

2. Key Target Groups and Statistics

The survey is designed to provide social and economic characteristics of the total civilian, noninstitional population of the U.S. under the age of 65. Although a wide variety of data content was planned for the survey, particular interest in the following key statistics was expressed:

- Percent of children in families below 200 percent of the poverty threshold;
- Percent of children in families below the poverty threshold;
- Percent of children living in families headed by a single parent; percent of children not covered by health insurance;
- The average annual number of physician contacts per child;
- Percent of adults under 65 below 200 percent of the poverty threshold; and
- Percent of adults under 65 below the poverty threshold.

For each statistic, data were desired for each of 13 selected states, and for the nation as a whole.

As can be seen, the survey is particularly focused on poor persons, and on children under 18 years of age. About eight households need to be contacted to locate each family with children below 200 percent of poverty. Since separate statistics were required for each of 13 states, and the balance of the U.S., a very large and costly screening effort was required.

One obvious way of reducing costs is to use the telephone for screening, using RDD as the sampling procedure. Unfortunately, RDD used alone is likely to be subject to serious biases. It has been estimated that about 20 percent of families in poverty do not have telephones, and that 10 percent of all families with a child 3 years old or under are without telephones. The nontelephone households probably have even lower incomes than other poverty families, and their

economic characteristics are also likely to be different. The first and major problem in the development of the survey design was how to keep costs within reasonable limits without introducing serious biases or large sampling errors. We addressed this problem by the use of a random digit dialing (RDD) frame to cover the approximately 95 percent of U.S. households with telephones, and the use of an area sample to represent households without telephones.

A second problem was how to integrate surveys for two periods of time in an efficient way. A preliminary review of experiences with longitudinal surveys indicated that with RDD there would be serious loss in response rates if the second survey attempted to contact the families and persons in the first survey. In addition to normal attrition, most households that moved in the 2- or 3-year period would be lost, as well as members of households that changed composition. The loss could easily be 40 percent or more.

We will now describe the procedures used to avoid or minimize the effects of these problems.

3. Major Considerations in The Development of the Sample Design

Most of the interviewing was carried out in early 1997, with final interviews being done later in the year. The survey is designed to have approximately equal samples sizes for each of the 13 states. The equal sample sizes refer to the number of low-income families with children rather than the number of households sampled.

We discuss here the allocation of the sample between the RDD and area samples, and the exclusion from the area frame of block groups with low rates of households without telephones.

3.1 Allocation to Telephone and Area Components

Adopting the usual models for the optimal allocation of a sample among strata, the ratio of the telephone and nontelephone sample in a state can be approximated by:

$$\frac{n_t}{n_{nt}} = \frac{W}{1-W}\sqrt{\frac{dc_{nt}}{c_t}},$$

subscripts t and nt refer to telephone and nontelephone; W is the proportion of the population coming from the telephone component; d is the design effect for the area sample due to clustering, and c is the cost of data collection per interviewed unit.

After the allocation of the sample was determined, it was further examined to evaluate how reasonable and stable it was. First, the components of variance of the estimates due to the telephone and nontelephone sample were estimated. If it appeared that one component was dominant or controlling the overall precision of the estimates in a state, then alternatives that increased the allocation to the dominant component were examined.

Table 1 shows the expected numbers of screened households that result from application of this formula. The area sample households are the number when low yield block groups are excluded. (See the next section for discussion of this design feature.) A more interesting comparison is the number of low-income children which is also shown in Table 1. The table shows that the variation between states in the area survey screened sample is not large, but the variation between states in the interviewed area sample of low income households with children is much greater.

3.2 Truncation of Area Frame

In order to reduce the potentially high cost of screening for nontelephone households in parts of the nation with very few nontelephone households, the area sample frame was truncated to exclude block groups with very low proportions of households without telephones. Data from the 1990 Census was used to determine cut-off levels for excluding block groups from the sampling frame. Cut-offs were determined on a state by state basis and chosen so that less than 10 percent of nontelephone households were excluded. Weighting adjustments are planned to compensate for the truncation. The reduction in the sampling frame reduced the screening workload by about 55 percent. The cut-offs for truncation in each state, and their effects on the workload are shown in Table 2.

4. Other Key Design Features

The RDD portion of the survey was carried out through a list-assisted method for sample selection. The sample households were first screened for the presence of children under 18 years of age, and to exclude those containing only person 65 and over. Households without children were subsampled for the adult sample. The households were further asked several simple questions about their 1996 income. All households with reported income under 200 percent of poverty and a subsample of higher income families were retained for the longer, detailed interview. The subsampling rates for higher income families and for households without children varied slightly among the states, and were designed to provide key statistics for the different subdomains with prespecified precision.

The nontelephone households were selected via a stratified, multi-stage, area sample. The PSUs were ones that are commonly used in area samples -- MSAs

and counties or combinations of several counties. Since state samples were required, MSAs that crossed state lines were subdivided. The number of sampled PSUs varied from 4 to 12 among the states, and with 18 PSUs for the "balance of the U.S." Area segments consisting of Census blocks or combinations of blocks constituted the second stage of sampling. Compact clusters were used, that is, all households in the segments were in the sample. The actual effective cluster size was much smaller since the households were screened for presence of telephone and only nontelephone households were eligible for the survey. Detailed interviews using the same questionnaires and interview techniques as in the RDD part of the survey were attempted at all nontelephone households. The overall sampling rates for nontelephone households were about the same as for the higher income households with children in each state.

The second survey is expected to be carried out in 1999 or 2000. It is not planned as the second phase of a longitudinal survey, since this would have produced serious attrition in response rates. However, strong correlations between the samples for the two years will be introduced to reduce the variances of estimates of changes between the two time periods. The future telephone sample will basically consist of the same set of telephone numbers that was used to screen households in 1997. The screening for households with children and to exclude those containing only persons 65 and over will be repeated. Low-income households with children will be oversampled in the same way as in 1997. A minor supplementary sample consisting of telephone numbers in new exchanges and new blocks of listed numbers will be added to the sample. The same principle will be followed in the nontelephone part of the sample; the 1997 segments will be used in 1999/2000.

The crucial sample sizes were those necessary to achieve reasonably reliable estimates for characteristics that are 30 percent of low-income families with children, where low income is defined as below 200 percent of poverty. The key criterion was to obtain a minimal detectable difference of 4.5 percent between the first and second survey years with an alpha of 5 percent and a power of 80 percent for each state. An effective sample size of 800 poor children per state per year is needed for this requirement. Better reliability was desired for the nation as a whole, and a sample of about 1,100 poor children was used for the part of the U.S. outside the 13 states. The effective sample size of low-income children for the nation as a whole was about 5,400.

Smaller effective sample sizes were used for nonpoor children since the intent was to permit analyses of all children, rather than nonpoor by themselves. The effective sample sizes ranged from 450 to 1,000 among the states; the variation was due to state differences in the proportions of low-income children. It should be noted that the effective sample sizes are not equal to the actual number of sample cases. The number of interviewed cases exceeded the effective sample sizes by a wide margin, mostly due to the variability in sampling rates.

The effective sample size for poor adults was about 500 in each state; for all adults, the effective sample ranged from 800 to 1,250. Somewhat larger samples were selected in the balance of the U.S.

5. Oversampling and Subsampling Consideration

5.1 Oversampling Low-Income Geographic Areas

Early in the planning of the sample design, the possibility of defining and oversampling low-income geographic areas for the RDD components was considered but rejected. An analysis of the 1990 Census distribution of low-income families among block groups (Waksberg, Judkins and Massey, 1997) indicated that such oversampling in area samples was not useful, and would be counterproductive in most surveys. It appeared clear that there is even less reason for such an approach in an RDD survey because the units that can be identified in an RDD sample are more heterogeneous than those in an area sample.

Such oversampling was also rejected for the area sample, partly for the same reason as for the RDD component; and partly because the truncation of the sampling frame largely eliminated higher income areas from the frame.

5.2 Subsampling Nonpoor Households

An important feature of the sample design of the RDD component was the identification of low-income households with children. In order to encourage as high a response rate as possible, the screening instrument was kept simple, with only a question on whether the income was above or below a particular level. More intensive probing for income was planned for the detailed interviews. It was recognized that the simple screening would not always provide the correct classification of households as poor or nonpoor, and a major factor in determining the actual sample sizes necessary to achieve the effective sample was the expected extent of error in the screening. There is evidence from other surveys that a simple question on income produces nontrivial response errors. The most recent evidence available to us was from the Continuing Survey of Food Intake Interview (CSFII),

conducted by Westat for the Department of Agriculture in 1994 and 1995. Table 3 shows some CSFII data.

Although the CSFII definition of low income used a threshold of 130 percent of poverty rather than the 200 percent in our survey, it seemed reasonable to assume that error rates would be roughly similar. Thus, based on our CSFII analysis, we estimated that 15 percent of the poor would be classified in the screener as nonpoor and thus subsampled, and 3 percent of the nonpoor would be classified in the screener as poor. Subsampling rates were needed that would keep the design effects from this subsampling at a reasonable low level. The rates were also influenced by the need to get required sample sizes for the poor. Table 4 contains the subsampling rates.

There was no subsampling for the nonpoor in the area sample component. One reason was that most nontelephone households were expected to be poor. A second reason was that the area sample was selected at much lower sampling rates than the RDD part of the survey, and subsampling could have created uncomfortably high design effects.

The sample sizes and effective sample sizes obtained in the survey are different from the target sample sizes for a number of reasons. The targets are based on projections from 1990 Census data and estimated response rates. Some of the reasons are discussed in the next section, and some data from the survey on the rates are presented

6. Preliminary Indications of Effectiveness of Sampling Plan

Data collection activities are still going on so that definitive information is not yet available on the extent to which the achieved sample corresponded to the assumptions in the development of the sample design. However, three aspects of the sample are disappointing. First, the cost of obtaining high cooperation rates was much greater than planned. Special additional efforts were undertaken, such as sending Federal Express letters to initially refusal RDD households with listed addresses. Although these efforts have had positive effects, it appears that final response rates will still be lower than planned in some states.

The rate of misclassification of low-income households is higher than anticipated in most states. Table 5 provides misclassification rates for households with children for the balance of the U.S. and two illustrative states, based on survey data collected so far. (Misclassification rates were generally somewhat higher for households without children.) We believe an important reason misclassification rates are higher than CSFII misclassification rates is that the CSFII rates are with respect to 130% of poverty rather than the 200% of poverty in NSAF. There were also several procedural differences between the two surveys that also contributed to different misclassification rates. For CSFII, both the screener reports and detailed interviews are in terms of household income, whereas for NSAF the screener report is in terms of household income but the detailed interview is in terms of family income. Also, there was an interval of several weeks between the initial determination of household income and the final family income determination for much of the NSAF sample, whereas for CSFII the detailed interview was conducted immediately after the screener report. Finally, CSFII was carried out as a face-to-face interview, and there may be differences in patterns of misreporting in RDD and personal interviews.

A third unexpected outcome was a lower rate of nontelephone households in the truncated frame than was reported in the 1990 census. We believe the major reason for this is the variation in telephone rates in block groups over time. In general, block groups with high rates of households without telephones in 1990 may tend to have lower rates in 1997, while block groups that have low rates in 1990 may tend to have higher rates in 1997. Some analysis, still in progress, indicates that this may be happening, and we are attempting to quantify the effect.

7. Conclusions

We have discussed two innovative features of NSAF. The first of these is the use of RDD in combination with an area sample to cover households without telephones. We discussed how we used a standard model to attempt to optimally allocate the sample between the two components. The second feature was confining the area sample to those block groups with moderate or high rates of households without telephones in 1990. Although this was generally successful in increasing efficiency, it did not result in as large an increase in the rate of nontelephone households as we expected.

Another feature of the survey was subsampling households that reported income over 200% of the poverty level in a brief screening interview. While some misclassification of income was expected based on experiences from another survey, the misclassification rates were greater than we expected. The result is a lower yield of low income households than expected. Fortunately, the subsampling rates were not very high, so that the survey effectiveness was not seriously impaired.

Reference

Waksberg, J., D. Judkins, and J. Massey (1997), "Geographic-Based Oversampling in Demographic Surveys of the U. S.", *Survey Methodology*, 23, #1, p. 61-71.

Table 1. Screened sample sizes and low-income child interviews in RDD and area sample components

	Screened sample			Low-income child sample		
State/area	Telephone sample size	Area sample size	Ratio of area screened sample size to telephone screened sample size	Telephone expected sample size	Nontelephone expected sample size	Ratio of nontelephone to telephone
Alabama	10,344	2,842	0.275	930	151	0.162
California	12,750	1,886	0.148	936	48	0.051
Colorado	15,045	2,110	0.140	988	76	0.077
Florida	16,170	2,350	0.145	1,003	81	0.081
Massachusetts	23,402	1,877	0.080	1,014	39	0.039
Michigan	14,135	1,879	0.133	984	63	0.064
Minnesota	16,439	1,887	0.115	1,006	32	0.032
Mississippi	8,329	2,822	0.339	992	206	0.208
New Jersey	26,918	2,355	0.087	1,068	77	0.072
New York	16,676	2,351	0.141	969	94	0.097
Texas	12,430	2,352	0.189	1,106	146	0.132
Washington	14,886	1,892	0.127	985	40	0.041
Milwaukee	14,884	1,893	0.127	1,000	40	0.040
Balance, Wisconsin	15,226	1,893	0.124	1,008	40	0.040
Balance, U.S.	19,714	3,765	0.191	1,411	143	0.101
Total	222,304	32,044	0.144	14,411	1,199	0.083

Table 2. Cut-off for exclusion of high telephone rate areas and effect on workloads

State/area	Cuff-off for low percent nontelephone (%)	Percent of nontelephone households excluded (%)	Percent of block groups excluded (%)	Percent of all households excluded (%)
Alabama	<5	7.3	37.4	40.9
California	<2	7.3	59.2	59.0
Colorado	<3	8.8	56.6	57.9
Florida	<3	9.1	48.0	54.6
Massachusetts	<2	9.4	70.5	70.4
Michigan	<3	9.8	56.1	59.9
Minnesota	<2	9.1	57.6	60.7
Mississippi	<8	9.7	34.5	35.9
New Jersey	<2	5.6	68.2	66.8
New York	<3	7.5	58.7	58.7
Texas	<5	7.7	42.2	45.3
Washington	<2	6.1	53.3	53.7
Milwaukee	<2	9.2	56.7	59.5
Balance, Wisconsin	<2			
Balance, U.S.	<3	8.0	54.9	57.1

Table 3. Comparison of income of screener and detailed interview in Continuing Survey of Food Intake (1995)

	Detailed interview					
Screened report	Poverty <130%	Poverty >130%	Total	Poverty <130%	Poverty >130%	Total
<130% Poverty	85.3	2.7	19.4	88.8	11.2	100%
>130% Poverty	14.7	97.3	80.6	3.7	96.3	100%
Total	100%	100%	100%	19.4%	80.6%	100%

Table 4. Subsampling rates for higher income and unknown income households

State/area	Proportion retained for sample	
	Higher income	Unknown income
Alabama, Mississippi	.40	.57
California, Colorado, Florida, Michigan, New York, Texas, Washington, Balance U.S.	.33	.50
Massachusetts, Minnesota, Milwaukee, Balance Wisconsin	.29	.44
New Jersey	.25	.40

Table 5. Comparison of income of screener and detailed interview households with children

	Detailed interview				
Screened report	<200% Poverty	>200% Poverty	<200% Poverty	>200% Poverty	Total
Balance of U.S.					
<200% Poverty	73.4%	9.6%	79.6%	20.4%	100%
>200% Poverty	22.3%	87.2%	11.6%	88.4%	100%
Unknown	4.3%	3.2%	40.7%	59.3%	100%
Total	100%	100%	—	—	—
California					
<200% Poverty	75.4%	8.9%	85.1%	14.9%	100%
>200% Poverty	16.7%	88.1%	11.3%	88.7%	100%
Unknown	7.9%	3.0%	64.4%	35.6%	100%
Total	100%	100%	—	—	—
Massachusetts					
<200% Poverty	65.0%	6.9%	75.2%	24.8%	100%
>200% Poverty	27.5%	90.3%	9.0%	91.0%	100%
Unknown	7.4%	2.8%	46.3%	53.7%	100%
Total	100%	100%	—	—	—