

Appendix B

*1999 NHSDA Quarter 1 Subsampling
Procedures*

1. PAPI Selection Probability Corrections

Excessive yields from PAPI screenings were observed and the problem was identified early in quarter 1 of the 1999 NHSDA. The probable cause of the problem was identified to be the Newton using the wrong household type probability selection vectors; this was later confirmed for 417 of the 500 quarter 1 PAPI segments. A plan was initiated to pull all PAPI cases with screening result codes less than 10 (i.e., not finalized) and corrected selection probability vectors would be transmitted back to the Newtons after appropriate corrections had been made. In order to compensate for the excessive yields, the selection probabilities on pending cases were reduced by a factor of 0.50 for the white household type, and by a factor of 0.80 for the black household type. Selection probabilities for Hispanic household types were left at their correct targeted values. For weighting purposes, the actual person selection probabilities used to select the sample persons were maintained in the control system.

2. Yield Reduction 1

Midway through quarter 1, it became apparent that more sample had been put in the field than could possibly be completed. Thus, a further sample reduction was implemented. Only pending screening cases were subsampled and at a rate of 1/3 retained. The process involved trying to get all field interviewers to transmit their current status codes and then removing a sample of pending cases based on the control system status at a set time (time 1). On the next transmit (time 2) cases sampled out were removed from the field interviewer's Newton. Any cases that had achieved a final screening status code (response or nonresponse) between time 1 and time 2 were put back on the field interviewer's Newton at the next transmission (time 3). A special flag was set up to identify cases in the following categories:

- 0 Not subject to subsampling at this round
- 2 Subject to subsampling and retained at time 1
- 3 Subject to subsampling and dropped out at time 1
- 4 Code 2 and code 3 cases, finalized between time 1 and time 2 and put back on the field interviewer's Newton. (Actually only code 3s were recoded to a code 4 when they were put back.)

In order to determine which cases were finalized between times 1 and 2, control system snapshots were taken of all cases eligible to be sampled at time 1, and the screening codes were retained for weighting purposes. **Table 1** shows the results of the sampling process. The time 2 snapshot, however, reflects the status as of the cutoff date and does not reflect the counts at the various times at which the transmissions for removing the sampled out cases occurred. These actually occurred on a flow basis. Of the 238 cases finalized among the sampled out cases, 44 were code 31s (one person selected) and 13 were code 32s (two persons selected). The flagging did not identify the retained cases that changed status from pending to final during the same period, but 640 retained cases that changed screening status from pending to complete by the cutoff date were identified. Rather than drop these finalized screenings, an approximate weight correction for the comparable portions of the dropped and retained sample was recommended. **Table 1a** and **Table 1b** show that use of an intermediate weight (between 1 and 3) allows the

retaining of all cases completed during the subsampling period and still preserves the initial sample representation based on the subsampling weight factors.

**Table 1a. First Subsample of Pending Screening Cases (Preliminary Tabulations):
Quarter 1 CAI**

Time 1 action	Number	Time 2 status	Number	Weight factor	Weighted number
Pending screening dropped from sample	8,791	Still pending (codes 00-09)	8,640	0.00000	0
		Finalized (codes 10-32)	151	1.00000	151
Pending screening retained in sample	4,370	Still pending (codes 00-09)	3,997	3.01167	12,038
		Finalized (codes 10-32)	373	2.60684	972
Subtotal	13,161		13,161		13,161
Screening finalized	40,193	Still finalized	40,193	1.00000	40,193
Overall total	53,354		53,354		53,354

**Table 1b. First Subsample of Pending Screening Cases (Preliminary Tabulations):
Quarter 1 PAPI**

Time 1 action	Number	Time 2 status	Number	Weight factor	Weighted number
Pending screening dropped from sample	5,571	Still pending (codes 00-09)	5,484	0.00000	0
		Finalized (codes 10-32)	87	1.00000	87
Pending screening retained in sample	2,730	Still pending (codes 00-09)	2,463	3.01167	7,418
		Finalized (codes 10-32)	267	2.98223	796
Subtotal	8,301		8,301		8,301
Screening finalized	17,664	Still finalized	17,664	1.00000	17,664
Overall total	25,965		25,965		25,965

This process did not achieve a 1/3 pending sample retention at time 2 because the sampling was done at time 1. The weight approximation allows the retaining of all work completed up to time 2. If an interviewer did not call in until much later (after a cutoff date), then her/his cases were never subsampled and all those initially subsampled out became code 4s and received the intermediate weight.

Since the procedure employed subsampling of pending screenings (not all pending screenings and interviews as erroneously projected), it did not adequately reduce the number of remaining cases to allow completion of quarter 1 assignments prior to the end of the quarter. It also appeared that no reasonable subsampling rate would have allowed us to adequately reduce the remaining quarter 1 workload without resorting to subsampling some screened SDUs with pending interviews. As a result, a second round of yield reductions was required and was implemented a few weeks later.

3. Yield Reduction 2

This step was implemented on March 18 following further discussions with SAMHSA about options remaining to effectively reduce sample yield without adversely affecting the response rates. With a few exceptions, this round of reductions reduced both pending screening cases and pending interview cases by 1/2. Subsampling was implemented at the SDU level, but in some cases SDUs removed from further followup included one completed interview (which would be retained for analysis purposes) and one pending interview (which was subject to the subsampling process). SDUs subject to the subsampling procedure included:

- (1) All SDUs with screening codes 01 through 09
- (2) All SDUs with screening code 31 (one persons selected) and a pending interview code 51-59.
- (3) All SDUs with screening code 32 (two persons selected) and both with pending interview codes in the 51-59 range
- (4) All SDUs with screening code 32, a finalized interview code (70-79) for one person and a pending interview code 51-59 for the other person.

Note that those SDUs with a completed screening code 31 or 32 and a pending appointment code (code 50) for one or both selected persons were not subject to subsampling since we did not want field interviewers to break any appointments as a result of this process.¹ **Table 2** shows the four categories eligible for subsampling and three additional categories that were not eligible for subsampling.

The subsampling occurred at time 1 and retained ½ of the cases in each of the four categories above. At time 2 some cases were put back onto the field interviewer's Newton because they had reached a final interview status during the interim.

¹A total of 97 SDUs had one person with a 50 code and another person with a code in the 51-59 range and all were retained.

Table 2. Sampling Eligibility by Category at First Snapshot

Category	Screening status	Interview status	Interview status	Eligible
1	Pending screening (00-09)	--	--	Yes
2	One person selected (31)	Pending exc. appt. (blank, 51-59)	--	Yes
3	Two persons selected (32)	Pending exc. appt. (blank, 51-59)	Pending exc. appt. (blank, 51-59)	Yes
4	Two persons selected (32)	Final or appt. (50, 70-79)	Pending exc. appt. (blank, 51-59)	Yes
5	One person selected (31)	Final or appt. (50, 70-79)		No
6	Two persons selected (32)	Final or appt. (50, 70-79)	Final or appt. (50, 70-79)	No
7	Final screening, no one selected (10-30)	--	--	No

Table 3 shows the actions and flag setting at time 2.

Table 3. Actions When Picking up Cases (as Field Interviewers Transmit from Newtons) after Sampling (Applies to Sampled Out Cases Only)

Category	Action	Flag
1	None	3
2	None	3
3	None	3
4	Keep final or appt., code other as 76, put back on Newton	5
5	Keep, put back	4
6	Keep, put back	4
7	Keep, put back	4

Table 4a and **Table 4b** show the weighting factors used which preserve the initial sample representation.

Table 4a. Round 2 Subsample: Quarter 1 CAI

Time 1 Action	Number	Time 2 category	Number	Weight Factor		Weighted DUs
				DU	Person	
Dropped from sample	4,044	1	1,827	1.00000	1	1,827
		2 or 3	1,814	0.00000	0	0
		4	317	1.00000	2	317
		5, 6, or 7	86	1.00000	1	86
Retained in sample	4,030	1	1,499	1.00000	1	1,499
		2 or 3	1,672	1.87594	1	3,137
		4	290	1.40679	1	408
		5, 6, or 7	569	1.40679	1	800
Total eligible	8,074		8,074			8,074
Not eligible for subsampling	45,280	5, 6, or 7	45,280	1.00000	1	45,280
Total	53,354		53,354			53,354

Table 4b. Round 2 Subsample: Quarter 1 PAPI

Time 1 Action	Number	Time 2 category	Number	Weight Factor		Weighted DUs
				DU	Person	
Dropped from sample	2,160	1, 2, or 3	1,959	0.00000	0	0
		4	157	1.00000	2	157
		5, 6, or 7	44	1.00000	1	44
Retained in sample	2,143	1, 2, or 3	1,630	2.00793	1	3,273
		4	162	1.61612	1	262
		5, 6, or 7	351	1.61612	1	567
Total eligible	4,303		4,303			4,303
Not eligible for subsampling	21,662	5, 6, or 7	21,662	1.00000	1	21,662
Total	25,965		25,965			25,965

4. Putting Back CAI Cases in Quarter 2

In order to reduce the effect of unequal weights, all pending dwelling units (all units from round 1 and 1,827 units from round 2) for the CAI sample were put back into the sample in quarter 2. This putting back of cases resulted in a total of 1,695 respondents who were assigned to quarter 1 but were fielded in quarter 2. No subsampled PAPI dwelling units were put back in quarter 2. The sample weights were adjusted to reflect the subsampling and putting back of cases for both samples.