PROBABILITY SAMPLING WITH QUOTAS*

SEYMOUR SUDMAN†
National Opinion Research Center, University of Chicago

This paper describes certain quota sampling procedures and attempts to show that they are very close to traditional probability sampling. Quotas are shown to depend on availability for interviewing and evidence is presented to show that sex, age, and employment status are reasonable predictors of availability. Quota sampling methods are not unbiased but data are presented which suggest that the bias is generally of the order of 3 to 5 per cent. It is shown, however, that the cost differentials between these quota samples and call-back samples are small. The major advantage of this new procedure may well be the speed with which interviewing may be completed during crises such as the Kennedy assassination.

1. INTRODUCTION

Two decades ago when the advocates of probability sampling met and defeated the defenders of quota sampling, the doctrine became established that there was an unbridgeable gulf between the two methods. While it was conceded that quota samples were cheaper, most sampling statisticians had no doubts that quota samples were far less accurate than probability samples, and that even worse, there was no way to measure the accuracy of a quota sample.[1]

This remains the general view today, although Stephan and McCarthy have given a justification of the measurement of sampling variability for quota

* This investigation was supported in whole by Research Grant 2-4402 from the National Science Foundation.

† I am most grateful for the comments and suggestions of Norman Bradburn, Morris Hansen, Nathan Keyfitz, William Kruskal, Frederick Mosteller and the referee. They are, of course, not responsible for the views expressed nor for any remaining ambiguities or errors.

[1] An illustration of the typical view held by sampling statisticians is given in Hansen, Hurwitz, and Madow, *Survey Methods and Theory*, Volume I [6, p. 71]:

> The so-called "quota controlled" sampling method, which has been widely used, is essentially a sample of convenience but with certain controls imposed that are intended to avoid some of the more serious biases involved in taking those most conveniently available. . . . The restrictions imposed on the convenience of the interviewer by this method may possibly considerably reduce the biases. However, they may also be completely ineffective. What is worse, there is no way to determine the biases except by a sample properly drawn and executed.

In William Cochran [2, p. 105], a similar, but slightly more favorable view is taken of quota sampling:

> Another method that is used in this situation [stratified sampling where the strata cannot be identified in advance] is to decide in advance the n_h that are wanted from each stratum and to instruct the enumerator to continue sampling until the necessary "quota" has been obtained in each stratum. If the enumerator initially chooses units at random, rejecting those that are not needed, this method is equivalent to stratified random sampling. . . .
>
> As this method is used in practice by a number of agencies, the enumerator does not select units at random. Instead, he takes advantage of any information which enables the quota to be filled quickly (such as that rich people seldom live in slums). The object is to gain the benefits of stratification without the high field costs that might be incurred in an attempt to select units at random. Varying amounts of latitude are permitted to the enumerators.
>
> Sampling theory cannot be applied to quota methods which contain no element of probability sampling. Information about the precision of such methods is obtained only when a comparison is possible with a census or with another sample for which confidence limits can be computed.

W. Edwards Deming [3, p. 31]:

> There is another kind of judgment sample called a quota sample. The instructions in a quota sample ask the interviewers to talk to a specified number of people of each sex and age, perhaps by section of the city, perhaps

samples [14, Chap. 10, pp. 211–34]. Meanwhile, there has been a major change in quota sampling methods, particularly since the failure of the polls in 1948 [12]. The major change has been the establishment of tight geographical controls which the interviewer must follow. That is, the interviewer in her search to fill her quotas follows a specified travel pattern, visiting pre-designated households. While this quota procedure is now widely used and has produced meaningful data, no one has yet had the audacity to justify this procedure on theoretical grounds.

It is the heretical intent of this paper to attempt a rationalization of this procedure which indicates that it is very close to traditional probability sampling. To differentiate it from older quota sampling methods which do not specify a travel pattern, the procedure will be referred to as Probability Sampling with Quotas. This procedure is not unbiased, but typically the bias is small. On the other hand, a careful cost analysis indicates that differences in direct interviewer costs between probability sampling with call-backs and probability sampling with quotas is also small. The major advantage of this new procedure may well be the speed with which interviewing can be completed. Thus, when speed is critical to obtain immediate public reaction to a crisis such as the Kennedy assassination, probability sampling with quotas can be most useful. The National Opinion Research Center completed the field work on a national study of public reactions to the President's assassination in about ten days using a probability sample with quotas. Quota studies with less urgency are finished in two or three weeks. On the other hand, regular probability samples usually take six weeks or longer.

The next section describes the theoretical foundation of the argument. Sections three and four present empirical data from various NORC studies which confirm the theory. Section five discusses procedures for computing sampling errors, while section six discusses the costs of sampling with call-backs and quotas. The final section compares the results of three almost identical studies of which two were done with call-backs and one with quotas.

2. ASSUMPTIONS UNDERLYING PROBABILITY SAMPLING WITH QUOTAS

In probability sampling with call-backs, the interviewer is given a specific household or individual who is to be interviewed. If the individual is not available on the first call, repeated call-backs are made until the interview is obtained or the respondent refuses to grant an interview.

In probability sampling with quotas the basic assumption made is that it is possible to divide the respondents into strata in which the probability of being available for interviewing is known and is the same for all individuals within the stratum, although varying between strata. Any respondent's probability of being interviewed is the product of his initial selection probability times his

---

by economic level. The report of the results usually boasts of good agreement between the sample and the census in respect to the classes specified, but what does this mean? It means that the interviewers reported what they were supposed to report concerning these classes; it proves little or nothing with respect to the accuracy of the data that constitute the purpose of the study.

There is no way to compare the cost of a probability sample with the cost of a judgment sample, because the two types of sample are used for different purposes. Cost has no meaning without a measure of quality, and there is no way to appraise objectively the quality of a judgment sample as there is with a probability sample.

probability of being available for interviewing. While these probabilities will not be identical for all respondents, they are known and so the sample is a probability sample. There is an implicit assumption that an interviewer in a sample segment follows the same time pattern over repeated surveys and that the respondent has a pattern of availability depending on his characteristics.

The quotas which are then used clearly must be associated with the probability of being available for interviewing. Essentially, the quotas should be based on the reciprocals of the probabilities of being found available. If the probability of the individuals in Stratum A being available is twice as large as the probability of individuals in Stratum B, then the sampling rate for Stratum A should be one-half of that for Stratum B when doing probability sampling with quotas.

In the usual situation, quotas are set for a given stratum based on the sampling rate and universe estimates of the size of the stratum. These quotas are normally determined for the smallest geographic area for which information is available. Thus, in metropolitan areas, Census Tract information is used while in non-tracted areas the quotas are based on the characteristics of the locality or of the rural portion of the county. This method introduces the possibility of error due to inadequate universe estimates, but generally is almost like the method which uses sampling rates directly.

This procedure is wasteful from a sampling viewpoint since households with no one at home are skipped as well as households where the respondent is not available for interviewing at the time the interviewer calls or households with respondents who do not fit the quota. The field cost savings, however, considerably exceed the increase in internal sampling costs.

Probability sampling with quotas has been used primarily for sampling of individual respondents. Where household behavior or opinions are wanted it would be possible to use the same procedure, but since size of household is highly correlated with availability it would be necessary to make it a major quota control. Since any knowledgeable adult is acceptable as the respondent in a household survey, probability sampling with call-backs of households is less costly than sampling of designated respondents in households. Generally, cost and time savings of probability samples with quotas of households will not be great enough to make this method very useful considering the possible biases.

The rationalization of probability sampling with quotas depends on a major assumption while probability sampling with call-backs does not require this assumption. Fortunately, there is strong evidence to be presented in the next section that this assumption is almost true for the kinds of surveys which are generally conducted in the United States. To the extent that the assumption is not true, small biases are introduced but the method still remains a probability sample.

Even in the usual probability sample with call-backs, biases exist due to noncooperators. These same biases exist in probability sampling with quotas. We have not observed any major difference in the over-all cooperation rates achieved by interviewers on probability samples with call-backs as compared to probability samples with quotas. These cooperation rates depend on both respondents and interviewers. Since respondents cannot be aware of the type of sampling, any difference would have to be due to the fact that the interviewer did not try as hard to convert a refusal into a cooperator. At least, with NORC interviewers who do both types of interviews, there is no evidence that this is occurring.

It may be useful for the reader to compare the rationale for samples described above with the Politz-Simmons weighting method sometimes used to adjust for not-at-home bias [13]. In this procedure no call-backs are made and no quotas are used. Typically the respondent is asked whether or not he was home on the preceding five nights, and his answers to this question determine the weight which he receives. Thus a respondent who was home all nights would get a weight of one while a respondent who had not been home on any of the preceding five nights would get a weight of six since only one-sixth of respondents of this type would be found at home on a random night.

The Politz-Simmons weighting has three disadvantages. First, the weighting depends on the respondent's memory of how he spent the last five nights, and, in general, respondents will tend to over-state their availability. Second, the use of weights increases the sampling variability substantially. Finally, the weighting method introduces the need for careful controls when the data are processed to insure that the weighting is done properly. It is my impression that not very many surveys currently use the Politz-Simmons weighting procedure because of these difficulties.

If one were willing to accept the answer to the at-home question as being reliable, and did not worry about the cost of weighting, then it would be possible to develop a combined sampling method which used probability sampling with quotas to keep sampling variability low and used the answer to the at-home question to eliminate remaining sample biases.

3. RESPONDENT CHARACTERISTICS RELATED TO AVAILABILITY FOR INTERVIEWING

How does one go about establishing strata within which individuals have the same probability of being available for interviewing, and how are these strata tested for homogeneity? Since direct data are unavailable one must use past experience on probability samples. Many earlier studies have shown that women are generally more readily available for interviewing than are men. Primarily this is due to the fact that more men than women are employed. When one imposes the additional control of employment status, one sees a substantial difference between employed and unemployed women, but the difference between men and employed women shrinks. In addition, age of men is of some importance. Thus, NORC developed a four stratum system for its probability sample with quotas:

Men under 30	Unemployed Women
Men 30 and Over	Employed Women

Another major factor determining availability is the size of the community in which the respondent lives. The basic probability sample design which is a multi-stage sample drawn with probabilities proportionate to size of the block or enumeration district controls for this. No claim is made that the four strata are optimum. NORC plans additional research to find characteristics which

TABLE 1. AVERAGE CALLS REQUIRED TO COMPLETE AN INTERVIEW ON VARIOUS PROBABILITY WITH CALL-BACK SAMPLES

Sample	All Respondents	Males	Under 30	30+	Females	Employed	Unemployed
NORC							
All places N	2.7 *906*	3.0 *387*	3.2 *76*	2.9 *311*	2.5 *519*	3.0 *212*	2.2 *307*
10 Largest SMA's N	3.2 *206*	3.2 *96*	3.4 *27*	3.2 *69*	3.3 *110*	3.9 *48*	2.8 *62*
Other SMA's N	2.9 *357*	3.3 *155*	3.5 *30*	3.2 *125*	2.5 *202*	2.9 *91*	2.2 *111*
Non-metro counties N	2.3 *343*	2.4 *136*	2.8 *19*	2.4 *117*	2.1 *207*	2.6 *73*	1.9 *134*
Survey Research Center							
All places N	2.2 *7,528*	2.3 *3,658*			2.1 *4,031*		
Large metro N	2.5 *2,299*						
Other Urban N	2.1 *3,717*						
Rural N	2.8 *1,512*						
Britain N	2.3 *1,443*	2.4 *938*			2.0 *505*	2.3 *55*	2.0 *450*
Elmira N	1.9 *1,029*	2.1 *452*			1.7 *577*		
Madison N	2.0 *743*	2.2 *313*			1.8 *430*		

may be more highly correlated with availability. One can certainly make the strata more homogeneous by splitting-off additional strata from those which already exist, but this makes the search procedure more costly. Under some circumstances it becomes cheaper to make call-backs than to continue the search for a respondent with rare characteristics.

Tables 1–4 present in summary form the relationship between characteristics and availability for interviewing. While these results may also be valuable in planning new sample surveys, they are primarily presented to show that the NORC strata are reasonable, if not optimum. Table 1 presents the average calls required to complete an interview by age, sex, and employment status for a typical NORC probability with call-back sample, and several other samples for which data are available.

The results obtained by NORC interviewers agree reasonably well with other published data on the number of calls required to complete an interview. Durbin and Stuart's data when re-computed to put it into the format of Table 1, show the same relationship between availability of men, employed women and

TABLE 2. PROBABILITY OF COMPLETING INTERVIEW ON FIRST CALL BY AGE, SEX, AND EMPLOYMENT STATUS WITH CALL-BACK SAMPLES

Sample		All Respondents	Males	Under 30	30+	Females	Employed	Un-employed
NORC								
All places	Prob. N	.28 *906*	.23 *387*	.24 *76*	.22 *311*	.31 *519*	.19 *212*	.40 *307*
10 Largest SMA's	Prob. N	.19 *206*	.18 *96*	.26 *27*	.14 *69*	.20 *110*	.10 *48*	.27 *62*
Other SMA's	Prob. N	.26 *357*	.21 *155*	.30 *30*	.18 *125*	.30 *202*	.16 *91*	.41 *111*
Non-metro counties	Prob. N	.35 *343*	.28 *136*	— *19*	.31 *117*	.40 *207*	.30 *73*	.45 *134*
Survey Research Center								
All places	Prob. N	.32 *2,963*	.26 *1,340*			.36 *1,623*		
Large metro	Prob. N	.21 *1,724*	.15 *323*			.26 *401*		
Other urban	Prob. N	.32 *1,501*	.20 *659*			.36 *842*		
Rural	Prob. N	.42 *738*	.37 *358*			.47 *380*		
Britain	Prob. N	.44 *1,443*	.40 *938*			.51 *505*	.35 *55*	.53 *450*
Elmira	Prob. N	.38 *1,029*	.24 *452*	.18 *108*	.26 *344*	.49 *577*		
Madison	Prob. N	.40 *743*	.27 *313*	.21 *57*	.29 *256*	.49 *430*		

TABLE 3. PROBABILITY OF COMPLETING INTERVIEW WITH ANY HOUSEHOLD MEMBER ON FIRST CALL BY SIZE OF HOUSEHOLD WITH CALL-BACK SAMPLES

Sample		Household Size						
		All Respondents	1	2	3	4	5	6 or more
NORC								
All places	Prob. *N*	.56 *1,257*	.46 *1,465*	.52 *3,219*	.56 *2,074*	.58 *1,979*	.63 *1,242*	.67 *1,287*
10 Largest SMA's	Prob. *N*	.44 *2,437*	.29 *351*	.36 *675*	.44 *478*	.50 *441*	.58 *260*	.59 *232*
Other SMA's	Prob. *N*	.56 *4,504*	.47 *548*	.51 *1,274*	.55 *835*	.58 *830*	.62 *484*	.67 *533*
Non-metro counties	Prob. *N*	.63 *4,316*	.55 *566*	.62 *1,270*	.64 *761*	.61 *699*	.66 *498*	.70 *522*
Elmira	Prob. *N*	*1,029*	.36 *129*	.38 *662*	.37 *166*	.50 *50*	— *13*	— *9*

unemployed women [4, pp. 395–97]. Mayer presents separate tables of availability by size of community and sex, and a cross-classification of these variables with first calls [9, p. 19–33]. His results are also in agreement with the NORC and British results. Two local studies in Madison and Elmira are also in general agreement with the other studies [8, 15].

For a better understanding of probability samples with quotas, however, it is also useful to consider the probability of completing an interview on the first call as well as the average number of calls required. There is a very substantial increase in the probability of finding a respondent after the first call. Thus, using average calls required over estimates the probability of a respondent being available on a probability sample with quotas. Table 2 shows these first-call probabilities for the NORC study and the other surveys of Table 1. Naturally, these results are somewhat more variable since they utilize only a fraction of the data, but they show exactly the same relationships.

If household information is required, it is not necessary to specify which individual in the household should furnish it. Generally, any knowledgeable adult would be qualified. While this does not bias a probability sample, it could cause a substantial bias in a probability sample with quotas since, as indicated in Table 3, and as one would expect, larger families would be more likely to be

TABLE 4. PROBABILITY OF COMPLETING CALL ON VARIOUS SURVEYS BY NUMBER OF CALL WITH CALL-BACK SAMPLES

Sample	*N*	1	2	3	4	5
National Opinion Research Center 1	2,211	.36	.65	.56	.54	.50
National Opinion Research Center 2	2,866	.42	.44	.52	.48	.48
National Opinion Research Center 3	5,083	.77	.53	.53	.39	.49
National Opinion Research Center 4*	12,441	.59	.50	.46	.47	.47
Largest SMA's	3,035	.48	.40	.37	.40	.41
Other SMA's	4,873	.58	.51	.48	.50	.49
Non-Metro	4,533	.67	.58	.56	.57	.60
SRC Total		.34	.39	.41	.39	.43
Large Metro		.22	.30	.32	.36	.42
Other Urban		.35	.41	.45	.42	.45
Rural		.43	.49	.50	.40	—

* Interview with any household member.

found at home than would smaller ones. For this reason, any such sample of households would clearly need to control for household size.

Table 4 gives the probabilities of completing a call by number of calls for four NORC studies and compares these results to those at the Survey Research Center [9, p. 24]. Typically there is a substantial rise from the first to the second call and then a leveling off on subsequent calls. This is true for the first two NORC studies and for the SRC data.

The third NORC study shown in Table 4 is an example of a survey where a screening call has been made first and an interview is then conducted with a special subsample of those screened. In that case, the interviewer has already obtained information from someone in the household and her probability of completing a call is greatest on the first call. Looking at Table 8, one may see that the costs for this study are the lowest of the four studies presented. As an aside, another NORC experiment discovered that substantial reductions in travel costs were achieved when interviewers phoned for appointments before going to the segment.

The final NORC study was the short screening questionnaire which was the prelude to the survey described in the previous paragraph. Here, the interviewer was permitted to interview any adult in the household instead of a specified family member. Again the probability of completing an interview is greatest on the first call and drops slightly on the second and subsequent calls.

4. TESTING FOR HOMOGENEITY

Having selected these strata (or some other grouping) one must then test for homogeneity for probability samples with quotas. For this purpose it seems reasonable to use the geometric distribution as the theoretical distribution to which our empirical distributions are compared.

Assume that the interviewer for a probability sample with quotas is conducting a random search, and that her probability of completing an interview at any random household is equal to the probability of completing an interview on the first call for a probability sample with call-backs in the same area. Let us pause for a moment to examine these two assumptions. While it is clear that the interviewer is not actually searching at random, the establishment of quotas makes it necessary for her to interview not only during the day, but also in the evening and on week-ends so that her searching times approximate a random procedure. The second assumption is realistic since the same interviewers conduct both types of samples in their areas, and generally have the same time periods available for interviewing.

Using these assumptions, the number of calls required to complete an interview for probability samples with quota is a random variable which has the geometric distribution (a special case of the negative binomial or Pascal). The expected mean and variance of this distribution are known to be:

$$E(X) = 1/p \qquad \text{Variance } (X) = q/p^2 [5, \text{p. } 174, 217\text{–}18].$$

The fit of this model to actual interviewing behavior can be seen in Tables 5 and 6. The first of these, Table 5, compares the means obtained on actual probability samples with quotas with the expected values obtained from the re-

TABLE 5. ACTUAL AND EXPECTED D.U. APPROACHES REQUIRED TO COMPLETE INTERVIEW FOR NORC PROBABILITY SAMPLE WITH QUOTAS*

Sample	All Respondents	Males	Under 30	30+	Females	Employed	Unemployed
All Places							
Actual $\bar{x}$	3.6	3.8	3.6	3.9	3.7	4.4	3.4
Expected $\bar{x}$	3.6	4.3	4.2	4.5	3.2	5.3	2.5
N	(1,916)	(919)	(171)	(748)	(997)	(338)	(659)
10 Largest S.M.A.s							
Actual $\bar{x}$	5.8	5.6	5.6	5.5	6.0	7.0	5.3
Expected $\bar{x}$	5.3	5.6	3.8	7.1	5.0	10.0	3.7
N	(480)	(230)	(36)	(194)	(250)	(96)	(154)
Other S.M.A.s							
Actual $\bar{x}$	3.4	3.8	3.9	3.8	3.4	3.6	3.3
Expected $\bar{x}$	3.8	4.8	3.3	5.6	3.3	6.3	2.4
N	(763)	(359)	(73)	(286)	(404)	(152)	(252)
Non-Metropolitan Counties							
Actual $\bar{x}$	2.5	2.5	2.0	2.7	2.5	2.9	2.3
Expected $\bar{x}$	2.9	3.6	—	3.2	2.5	3.3	2.2
N	(673)	(330)	(62)	(268)	(343)	(90)	(253)

* Actual values derived from NORC probability samples with quotas. Expected values are reciprocals of values in Table 2 which are based on NORC call-back samples.

ciprocal of the probabilities of being home on the first call which were given in Table 2, and which are based on call-back samples. The actual means and expected means are in close agreement, generally, which lends substantial credibility to the model.

It may be useful in considering these results to keep in mind how they were obtained. For NORC probability samples with quotas a detailed listing sheet is kept by all interviewers. This sheet shows every house visited, and the results obtained. While vacant dwelling units and commercial establishments are listed, they were not included in the counts, nor were households which had no respondent who fitted the quota. It should be noted that quota refusals are included since generally these were not final refusals, but rather respondents who were unavailable for interviewing at the time the interviewer called. The expected probabilities from the call-back samples are based on completed cases, and do not account for hard-core refusals, or vacant or commercial units.

Table 6 compares the actual and expected variances for the probability sample with quotas. Since, unfortunately, the expected variances are sensitive to small changes in the mean, they are based on the combined means of both the samples in Table 5. For example, consider males under 30 in the ten largest SMA's. If the actual value of Table 5 is used, $p=1/5.6=.18$ and expected $\sigma^2=(1-p)/p^2=22.2$. If the expected value is used based on call-backs $p=.26$ and $\sigma^2=10.9$. Combining these estimates gives p a value of .213 and $\sigma^2=17.4$ which is the value shown in the table. The comparisons show close agreement for employed females and men over 30 everywhere, and generally good agreement for all groups in the non-metropolitan counties. Men under 30 and unemployed women have higher than expected variances in metropolitan areas.

While this is probably a weak test, the ratios in a few of the cells of Table 6

TABLE 6. ACTUAL AND EXPECTED VARIANCE OF D.U. APPROACHES REQUIRED TO COMPLETE INTERVIEW FOR NORC PROBABILITY SAMPLE WITH QUOTAS*

Sample	All Respondents	Males	Under 30	30+	Females	Employed	Unemployed
All Places							
Actual σ^2	24.4	22.9	23.2	22.8	25.8	32.9	21.9
Estimated σ^2	10.6	12.4	11.3	13.4	8.5	18.7	5.8
Ratio	2.3	1.8	2.1	1.7	3.0	1.8	3.8
10 Largest SMA's							
Actual σ^2	57.4	45.4	50.3	43.9	71.2	76.9	62.8
Expected σ^2	25.3	25.8	17.4	33.4	24.8	63.8	15.8
Ratio	2.3	1.8	2.9	1.3	2.9	1.2	4.0
Other SMA's							
Actual σ^2	16.3	20.4	23.3	19.7	12.5	14.5	11.4
Expected σ^2	9.4	14.2	9.4	17.4	7.9	19.6	5.3
Ratio	1.7	1.4	2.5	1.1	1.6	0.7	2.2
Non-Metropolitan Counties							
Actual σ^2	5.6	6.2	2.0	7.2	5.0	7.0	4.2
Expected σ^2	4.6	6.3	2.0	5.8	3.8	6.5	2.8
Ratio	1.2	1.0	1.0	1.2	1.3	1.1	1.5

* Actual values derived from NORC probability samples with quotas. Expected values are $(1-p)/p^2$ where p is the reciprocal of the average of the actual and expected values given in Table 5, and is based on both quota and call-back samples.

clearly suggest that there are biases still remaining in the groups with the higher than expected variances, but do not insure the absence of biases in the other cells. That is, it is possible that some groups within a stratum have a substantially higher or lower probability of being found at home than the average for that stratum, and due to the large variance of the estimate of the variance of the mean, this cannot be detected.

It is possible to make some estimates of the maximum magnitude of the bias in the final results based on the data shown in Table 6. Suppose that each stratum is not homogeneous, but consists of two equal sized sub-strata which are homogeneous within themselves. (It can be shown that the bias is maximized with two equal sized sub-strata rather than with a larger number of strata or unequal sized strata.)

The estimated means of the sub-strata are shown in Table 7. These values are derived from the formula

$$\sigma^2(X_s) = \frac{1}{2}\left(\frac{q_1}{p_1{}^2} + \frac{q_2}{p_2{}^2}\right) + \frac{1}{4}\left(\frac{1}{p_1} - \frac{1}{p_2}\right)^2$$

Using the values for the observed variances from Table 6, the values for p_1 and p_2 can be easily derived. The bias in the estimate also depends on differential behavior between the two sub-strata. If the proportion of sub-stratum 1 with a given characteristic is R and the proportion of sub-stratum 2 with the characteristic is αR then the bias in the stratum estimate is

$$\frac{p_1 - p_2}{p_1 + p_2}\left(\frac{\alpha - 1}{\alpha + 1}\right)$$

TABLE 7. ESTIMATED BIASES IN QUOTA SAMPLE RESULTS WHEN STRATA CONSIST OF TWO SUB-STRATA

Stratum	Probability of Completing Interview on First Call			Estimated Per Cent Bias in Results			
	Total	Sub-Stratum 1	Sub-Stratum 2	$\alpha=1.3$	$\alpha=1.5$	$\alpha=2$	$\alpha=3$
10 Largest SMA's							
Men Under 30	.22	.11	.33	9.4	10.0	16.7	25.0
Men Over 30	.16	.12	.30	4.7	5.0	8.3	12.5
Women-Employed	.13	.09	.17	5.8	6.2	10.3	15.5
Women-Unemployed	.23	.10	.36	10.7	11.4	19.0	28.5
Other SMA's							
Men Under 30	.28	.16	.40	8.1	8.6	14.3	21.5
Men Over 30	.22	.17	.27	4.3	4.6	7.7	11.5
Women-Employed	.22	.22	.22	0	0	0	0
Women-Unemployed	.36	.21	.51	4.9	5.2	8.7	13.0
Non-Metropolitan Counties							
Men Under 30	.50	.50	.50	0	0	0	0
Men Over 30	.34	.27	.41	3.9	4.2	7.0	10.5
Women-Employed	.32	.30	.34	1.1	1.2	2.0	3.0
Women-Unemployed	.44	.33	.55	4.7	5.0	8.3	12.5
All-Strata	—	—	—	3.2	4.9	8.2	12.3

While α will not generally be known, past experience would suggest that α would normally be about 1.5 or less, although it might be as high as 2 or 3. Table 7 also shows estimated biases for each stratum for α values of 1.3, 1.5, 2, and 3. It can be seen that the over-all estimates of bias for an estimate range from 3 per cent to 12 per cent. These estimates do, of course, depend on the estimates of Table 6 as well as the estimates of α and so are subject to large error, particularly in the individual strata. Nevertheless, Table 7 indicates that typically differences between call-back and quota samples will not exceed 8 per cent and that most differences will be of the order of 3 to 5 per cent. Empirical verification of this is given in the final section of this paper and in Table 10.

Summary: This section has presented empirical results which support the reasonableness of treating sampling with quotas as a form of probability sampling. The agreement in Table 5 between the number of dwelling unit approaches on a probability sample with quotas and the reciprocal of the probability of completing an interview on the first call for probability sampling with call-backs strongly supports the notion that these two kinds of sampling methods have much in common.

The results of Tables 6 and 7, however, indicate that some biases still exist in the probability sample with quotas, at least for the strata considered. That is, within some of the strata there are probably individuals belonging to substrata whose probabilities of being available for interviewing differ from the stratum mean. If there were no considerations of time or cost there would never be a reason not to use call-backs instead of quotas. When these time and cost factors are considered more important than the biases (generally small) introduced by the use of sampling with quotas, this method becomes preferred.

5. SAMPLING VARIABILITY OF PROBABILITY SAMPLES WITH QUOTAS

There are no basic differences between methods for computing sampling errors for probability samples with quotas and the usual sample error computations. Of course, the naive use of simple random sampling error formulas for complex national samples is never appropriate for either quota or call-back samples.

Since stratification of samples of individuals generally does not significantly reduce sampling variability, the major difference between complex and simple samples is due to clustering. A probability sample with quotas is a cluster sample and will have about the same variability as a similar cluster sample with call-backs.

Direct methods of computing variances of complex samples are tedious and almost all statisticians currently use short-cut procedures. These are very well described by Kish [7] and Hansen, Hurwitz and Madow [6] and need not be repeated here. The suggested procedure of Stephan and McCarty [14, Chap. 10] is one of several possible appropriate methods. These procedures do not measure uncertainty about biases or response errors and so are minimum estimates of error.

6. COSTS OF PROBABILITY SAMPLES WITH CALL-BACKS AND WITH QUOTAS

The chief argument made for the old quota samples was that they were cheap. The costs of probability sampling with quotas are still less than the costs of sampling with call-backs, but the differences are much narrower. This section compares the costs of various NORC probability call-back and quota samples, and indicates that a substantial portion of the cost differential between them is not due to field activities but rather to other aspects of the study unrelated to sampling.

Let us first compare the total costs of six NORC probability call-back and four quota studies. It can be seen from Table 8 that call-back sample costs per case are typically three times as high as the costs per case of quota samples; probability with call-back samples have a median cost of about \$52.00 per case as compared to a median of about \$19.00 for probability samples with quotas.

A brief examination of the table, however, reveals that a substantial part of this difference is due to differences in the planning, processing, and analysis between the two surveys. Almost always, the planning and analysis of call-back samples is costlier, and takes a larger part of the total cost of the study. It seems clear that it is not the sample design which determines the cost of a study, but rather that the cost determines the sample design.

To be more explicit, where survey results will receive very sophisticated

TABLE 8. FIELD, AND OTHER COSTS FOR NORC SURVEYS

Costs	Probability With Call-Backs						Probability With Quotas			
	1	2	3	4	5	6	1	2	3	4
Direct Field Costs	$ 31,800	$ 21,000	$ 19,500	$ 5,000	$22,000	$16,900	$ 8,900	$ 9,900	$ 8,500	$ 9,000
Field Supervision	8,100	29,500	4,900	2,500	9,500	6,000	1,900	1,700	1,200	1,900
Other Survey Costs	173,100	106,200	93,400	31,400	38,500	26,500	16,000	18,600	14,100	14,800
Total Costs	213,000	156,700	117,800	38,900	70,000	49,400	26,800	30,200	23,800	25,700
Total Cases	2,380	2,810	2,200	760	2,500	1,500	1,200	1,500	1,300	1,500
Cost/Case	89.50	55.80	53.50	51.20	28.00	32.90	22.30	20.20	18.30	17.10
Direct Field Cost/Case	13.40	7.50	8.90	6.60	8.80	11.30	7.40	6.60	6.50	6.00
Total Field Cost/Case	18.70	18.00	11.10	9.90	12.60	15.30	9.00	7.70	7.50	7.30

analysis or when critical decisions will be based on them, it will be worthwhile to pay a substantial cost to achieve high standards of sampling, processing, and control. Thus, the Census Bureau rightly has very high standards on their Current Population Surveys. On the other hand, many exploratory studies do not require such high standards since the analysis may be more limited and the questionnaire may itself be a major source of error. Here quota sampling would be justified.

On the other hand, the relationship goes both ways. One reason why analysis costs are higher on call-back samples is that the analysts spend more time waiting for results to become available. Very often the field data collection period is extended for several weeks which delays the processing an additional period. While this waiting time may sometimes be useful in developing codes and modes of analysis, it is frequently wasted.

If one looks only at total field costs, which include both direct and supervisory costs, the ratio of the costs of probability samples with call-backs to probability samples with quotas drops from 3 to 1 to 2 to 1. The median field cost per case for call-back samples of the six studies is $14.00 as compared to $7.50 per case for the quota samples.

This comparison can be carried still one step further. The major difference between the two types of samples is the cost of supervision. When one examines only the direct cost of interviewing, the difference shrinks to about $2.50 between the median direct field cost per case of $9.00 for call-back samples and $6.50 for quota samples.

This difference represents the one extra hour per case which the average interviewer on a call-back sample must spend to find her respondent, as well as some additional travel expenses. It is clear that this difference is less important than the difference in the cost of supervision. Thus, it would seem possible to reduce the cost of a call-back sample almost to the level of a quota sample if one could reduce supervisory costs.

Again it should be noted that not all of the difference in supervisory cost is due to the sampling method. Some of this difference can be attributed to the greater quality control checks generally used in the more expensive samples. The additional effort generally made in training interviewers on call-back samples requires more supervisory time. There are, however, some characteristics of a call-back sample which do generate greater supervisory costs. Typically, the interviewer is told to make three call-backs and then to check with the supervisor for further instructions if she has not completed an interview. The decision process whereby this occurs is quite costly in supervisory time, as are the letters and long-distance calls which accompany the revised instructions. If methods can be found for standaridizing the follow-up procedures and eliminating most of the ad hoc decisions now made, a substantial part of field supervisory costs could be eliminated.

Another aspect of call-back sampling which leads to higher supervisory costs is the time period required to complete the interviewing. Since some respondents will be temporarily unavailable at the interviewer's initial call, she will need to return at a later time. If the respondent is on vacation or in the hospital it may take several weeks before the interviewing is completed. Proper allow-

TABLE 9. AVERAGE TRAVEL COST AND MARGINAL COST TO COMPLETE INTERVIEW BY NUMBER OF CALLS FOR NORC CALL-BACK SAMPLES

Calls Required to Complete Interviews	1			2			3			4		
	N	Aver. Total Cost	Marg. Cost	N	Aver. Total Cost	Marg. Cost	N	Aver. Total Cost	Short Marg. Cost	N	Aver. Total Cost	Long Marg. Cost
1	792	$3.23		1,202	$2.89		7,285	$ .89		3,894	$1.13	
2	791	4.14	.91	738	3.50	$ .61	2,562	1.34	$ .45	631	2.55	$1.42
3	349	5.30	1.16	480	3.72	.22	1,187	1.98	.54	293	3.68	1.13
4	152	6.98	1.68	215	4.43	.71	661	2.50	.52	103	4.33	.65
5	64	8.46	1.48	112	5.24	.81	351	3.13	.63	79	4.59	.26
6	34	8.67	1.21	42	7.06	1.82	176	4.22	1.09	35	6.12	1.53
7 or more	29	9.22	.55	77	8.13	1.07	219	5.97	.6	48	9.26	1.12
Total N	2,211			2,866			12,441			5,083		

ances should be made for this when scheduling, and the flow of completed questionnaires should be watched carefully. Nevertheless, some added cost due to the stretching out of the time period cannot be avoided with call-back samples.

Summary: Table 8 shows that the direct interviewer cost difference between call-back samples and quota samples is about $2.50 or one hour of interviewer time per case. Much larger differences are due to added supervisory costs, some of which might be eliminated by the standardization of call-back procedures. The greater length of time required to complete a sample with call-backs is another factor in increasing supervisory and other overhead costs and there is no way to avoid this increase.

7. MARGINAL COST OF CALL-BACKS

The small difference between direct field costs of call-back sampling and quota sampling is due both to the fact that quota sampling as described in this paper is more costly than uncontrolled sampling, and that the marginal cost of call-backs is not as large as is generally believed. This is not a new finding [1, pp. 189–90], but it seems worthwhile to present additional evidence to support this result.

The data in Table 9 show that the marginal travel cost of additional calls remains fairly constant although sampling variability causes differences as the number of cases becomes small. That is, it is generally less expensive to continue call-backs than to draw an additional sample. Table 10 shows the same relationship when the costs are separated by size of place.

One would think that travel costs per case would have to rise on repeated call-backs since relatively fixed costs are divided among fewer cases, but this is compensated by the increased probability of finding a respondent at home and the interviewer's greater familiarity with the area she is sampling.

The allocation of travel costs requires some arbitrariness. While the NORC procedure adopted is certainly a valid one, there are other allocation methods which might be equally valid and which would probably also lead to similar results. Travel costs are divided into two parts—travel time within a segment and travel time to and between segments.

TABLE 10. AVERAGE TRAVEL COSTS AND MARGINAL COSTS TO COMPLETE INTERVIEW BY SIZE OF P.S.U. FOR FOUR NORC CALL-BACK SAMPLES

Calls Required to Complete Interviews	10 Largest Metro			Other Metro			Non-Metro					
	N	Aver. Travel Cost	Marg. Cost	*N*	Aver. Travel Cost	Marg. Cost	*N*	Aver. Travel Cost	Marg. Cost	*N*	Aver. Travel Cost	Marg. Cost
A. 1	1,442	$1.09	—	2,811	$ .89	—	3,032	$ .80				
2	643	1.78	.69	1,045	1.30	.41	874	1.06	.26			
3	351	2.46	.68	485	1.79	.49	351	1.66	.40			
4	237	3.10	.76	268	2.45	.66	156	2.06	.40			
5	149	3.67	.57	130	2.90	.45	72	2.44	.38			
6	86	4.60	.93	68	3.25	.35	—	—	—			
7	59	5.33	.93	27	4.16	.91	—	—	—			
8 or more (av = 10)	68	8.04	.90	39	7.10	.98	—	—	—			
B. 1	724	1.55	—	1,402	1.12	—	1,768	.97	—			
2	158	3.60	2.05	263	2.46	1.34	210	1.86	.89			
3	88	4.95	1.35	127	3.72	1.26	78	2.18	.38			
4	32	6.19	1.24	49	4.26	.54	22	3.27	1.09			
5	20	6.95	.76	34	5.47	1.21	25	4.05	78			
C.	10 Largest Metro			Prince Georges			Warren				Detroit	
1	150	3.14	—	697	2.97	—	147	2.42	—	208	2.79	—
2	132	3.50	.36	294	3.41	.44	121	3.62	1.20	191	3.56	.77
3	135	3.58	.08	148	3.84	.43	76	4.44	.82	121	3.73	.17
4	59	3.86	.28	55	4.86	1.02	37	4.70	.26	64	4.98	1.25
5	38	4.50	.74	22	4.98	.12	23	6.44	1.74	29	5.34	.36
D.	Metro			Non-Metro								
1	345	3.80	—	370	2.88	—						
2	425	5.04	1.24	336	3.20	.32						
3	229	5.69	.65	117	4.25	1.05						
4	120	6.03	.34	34	4.62	.37						

Travel to Segment.—is derived from the Travel column on the Time Report. It includes time to the segment from the interviewer's home and return. It also includes any travel time from one segment to another. It is generally not too difficult to separate this time from the time spent by the interviewer within the segment.

Travel costs to a segment are allocated equally, but only to the completed interviews made on that trip. That is, not-at-home calls were not charged any travel time to a segment unless there were no completed interviews on that trip. The rationale behind this allocation is that if the interviewer has been in the segment anyway (either with or without an appointment) and has completed one or more interviews, then any additional calls she makes in the segment are gravy, as far as travel time to the segment.

Travel in Segment.—is defined as all time in a segment not spent on the actual interview. Travel in segment includes all waiting time, and time in a respondent's home spent in conversation not part of the interview, as well as time spent locating the proper house in the segment and knocking on doors. Also included here is the time the interviewer spends on the telephone making appointments for interviews. This type of travel time is not always directly noted by interviewers filling out the present time sheet. It is sometimes included under travel time, sometimes under interviewing time and sometimes under other time. In coding the Time Reports, cross-checks are made with questionnaires. If the interviewer combines waiting time or other time within the segment with the interview, the length of the interview as obtained from the questionnaire is subtracted from the total time shown and the balance is called "travel in segment." Even where the interviewer has separated her time, cross-checks still are made to the questionnaire to insure that dates and times agree. If not, the normal procedure is to adjust the Time Report to the questionnaire since times in the questionnaire were presumably filled out immediately while the Time Report is generally filled out later.

Travel costs within the segment were allocated equally to all the calls made in the segment on that trip. That is, all calls, whether or not they resulted in a completed interview were charged with the same fraction of the travel costs within the segment.

8. COMPARISON OF SURVEY RESULTS ON PROBABILITY CALL-BACK VS. QUOTA SAMPLES

This final section compares the results of three NORC studies conducted for Professor Jiri Nehnevajsa of the University of Pittsburgh all of which dealt with attitudes of the American public on questions dealing with world tensions [11]. On each of these surveys five questions were asked in exactly the same way, and the responses to these questions are presented in Table 11. In addition, a large number of demographic characteristics which were obtained on all three surveys are also compared.

As is to be expected from the previous discussion, the differences between the three surveys are small, and can mostly be accounted for by sampling variability, the different times at which the three studies were conducted, and small differences in the ages of eligible respondents. The first sample with call-backs was conducted in June, 1963, the second in June, 1964, while the sample with quotas was conducted in December, 1963, midway between the other two. On the first call-back study, respondents over 65 were excluded while on the last two studies all adult respondents were eligible.

These results do not prove the lack of biases in the probability sample with quotas, but they support the view that for most items these biases cannot be large. The reader may also wish to compare these results to those of Moser

TABLE 11. COMPARISON OF SURVEY RESULTS FOR NORC U. S. NATIONAL SAMPLES: PROBABILITY WITH CALL-BACK AND QUOTA

SAMPLE SIZE	1,416	1,452	1,393
Opinions on World Affairs	Call-Back		Quota
1. *The amount of world tensions just about now*			
0 No tensions at all	0.3%	0.8%	0.3%
1	0.1	0.5	0.4
2	0.8	0.9	1.2
3	2.4	2.7	3.4
4	4.9	6.7	4.1
5	16.3	16.6	20.0
6	14.0	13.2	11.6
7	21.2	16.5	14.9
8	19.1	17.8	17.6
9	8.8	8.5	6.3
10—Extremely high tensions	12.1	15.8	20.2
	100.0%	100.0%	100.0%
$\bar{x}$	6.95	6.92	7.03
$s_{\bar{x}}$	.08	.08	.09
2. *World tensions just about two years from now*			
0 No tensions at all	0.5%	0.7%	0.5%
1	0.6	0.7	0.8
2	1.7	3.1	3.2
3	4.1	4.6	6.5
4	6.7	7.2	7.8
5	12.5	13.6	17.1
6	11.0	9.3	12.3
7	14.8	14.8	13.8
8	19.8	17.8	15.7
9	13.2	11.5	8.3
10—Extremely high tensions	15.1	16.7	14.0
	100.0%	100.0%	100.0%
$\bar{x}$	7.03	6.90	6.54
$s_{\bar{x}}$	.09	.09	.09

TABLE 11 (*Continued*)

Opinions on World Affairs	Call-Back		Quota
3. *World tensions five years from now*			
0 No tensions at all	1.0%	1.7%	1.6%
1	1.1	1.5	2.2
2	4.4	4.1	5.7
3	7.2	7.6	8.0
4	7.9	10.0	9.5
5	15.1	17.8	23.1
6	11.3	11.1	9.5
7	11.9	11.4	9.9
8	14.7	12.9	12.2
9	10.8	8.4	6.7
10—Extremely high tensions	14.6	13.5	11.6
	100.0%	100.0%	100.0%
$\bar{x}$	6.51	6.22	5.91
$s_{\bar{x}}$	.10	.10	.11
Opinions on World Affairs	**Probability**		**Quota**
4. *World tensions two years ago*			
0 No tensions at all	0.6%	1.1%	0.4%
1	0.6	1.7	0.8
2	3.1	3.5	2.7
3	6.0	5.7	4.1
4	8.5	8.3	5.6
5	14.4	12.9	11.8
6	14.0	11.8	7.9
7	14.1	13.5	14.6
8	19.3	17.8	19.2
9	11.2	11.2	14.2
10—Extremely high tensions	8.2	12.5	18.7
$\bar{x}$	6.51	6.57	7.18
$s_{\bar{x}}$	.09	.09	.09
5. *Average world tensions*			
Two years ago	6.51	6.57	7.18
Now	6.95	6.92	7.03
Two years from now	7.03	6.90	6.54
Five years from now	6.51	6.22	5.91
6. *When will the cold war end?*			
Within two years	4.6%	7.0%	3.4%
Within five years	22.4	21.9	20.4
Within ten years	22.8	23.6	23.2
10 to 20 years	16.5	14.8	18.6
21 to 50 years	5.0	6.1	5.9
Over 50 years	5.5	5.7	7.0
Never	13.5	14.0	15.4
Don't know	9.7	6.9	6.1
	100%	100%	100%
Median years	10.0	9.5	11.6

TABLE 11 (*Continued*)

Demographic Characteristics	Probability		Quota	
7. *Sex* (N)	1,434	1,464	1,557	1,482
Male	45.6	44.8	48.1	48.3
Female	54.4	55.2	51.9	51.7
8. *Race*				
White	87.9	85.6	86.8	89.8
Other	12.1	14.4	13.2	10.2
9. *Service (or spouse's service) in armed forces*				
Yes	55.9	48.3	50.1	—
No	44.1	51.7	49.9	—
10. *Service in combat if served in armed forces*				
Yes	37.1	38.3	39.0	—
No	62.9	61.7	61.0	—
N	(726)	(645)	(715)	
11. *Political affiliation*				
Democrat	52.9	54.7	58.5	53.6
Republican	31.6	28.0	28.0	27.7
Other	4.9	0.9	5.6	2.3
None or independent	10.6	16.4	7.9	16.4
12. *Religion*				
Protestant	68.6%	69.2%	67.6%	66.2%
Catholic	24.4	23.3	23.4	25.7
Jewish	2.7	3.4	2.3	3.2
Other	2.0	1.5	2.8	2.5
None	2.3	2.6	3.9	2.4
13. *How strongly do you feel about your religious beliefs?*				
Very strongly	37.9%	43.6%	38.4%	—
Strongly	27.6	22.8	28.4	—
Moderately	29.1	28.3	27.1	—
Not so strongly	3.9	3.3	3.5	—
Not strongly at all	1.5	2.0	2.6	
14. *Marital Status*				
Single	8.7%	7.4%	9.4%	—
Married	79.1	75.9	77.4	—
Divorced	3.6	3.8	3.0	—
Widowed	5.9	10.4	7.6	—
Separated	2.7	2.5	2.6	—
15. *Social class perception*				
Upper	2.2%	3.8%	1.9%	—
Middle	44.4	42.5	44.5	—
Working	47.3	46.7	49.0	—
Lower	3.9	2.8	2.6	—
There are no classes	0.8	2.0	—	—
Don't know	1.4	2.2	2.0	—

TABLE 11 (*Continued*)

Demographic Characteristics	Probability		Quota	
16. *Own or rent*				
Own	63.6	64.5	61.2	—
Rent	36.4	35.5	38.8	—
17. *Household Size*				
1	8.1%	12.0%	6.8%	—
2	24.9	29.2	24.5	—
3	18.1	15.5	20.6	—
4	21.5	16.7	19.4	—
5	13.7	13.1	12.8	—
6	6.8	7.8	8.0	—
7	3.4	2.8	4.0	—
8	1.9	1.7	1.9	—
9 or more	1.6	1.2	2.0	—
$\bar{x}$	3.62	3.42	3.70	—
18. *Education of respondent*				
0–8 years	21.7%	26.5%	26.0%	25.4%
9–11 years	22.8	20.3	22.9	22.5
12 years	30.2	29.6	28.2	28.7
13–15	13.3	13.2	13.7	13.6
16	7.5	6.0	6.0	5.8
17 or more	4.5	4.4	3.2	4.0
Median years	12	12	12	12
19. *Education of spouse of respondent*				
0–8 years	20.1%	25.7%	26.5%	—
9–11 years	23.5	20.1	21.3	—
12 years	32.8	32.2	31.7	—
13–15 years	11.7	11.8	9.8	—
16	8.3	6.9	6.5	—
17 or more	3.6	3.3	4.2	—
Median years	12	12	12	—
20. *Main wage earner*				
Respondent	53.5%	—	51.9%	—
Spouse	38.7	—	37.8	—
Others	7.8	—	10.3	—
21. *Income*				
Under 3,000	17.0%	21.1%	18.6%	23.1%
3,000–4,999	20.0	19.4	22.4	22.1
5,000–7,499	29.2	26.3	29.1	26.6
7,500–9,999	16.4	15.9	15.4	15.4
10,000–14,999	11.9	11.9	10.8	8.9
15,000–24,999	4.4	4.1	2.6	3.9
25,000 and over	1.1	1.3	1.2	
Median	$6,100	$5,900	$5,800	$5,500

TABLE 11 (*Continued*)

Demographic Characteristics	Probability		Quota
22. *Occupation of main earner*			
Professional	13.3%	12.6%	13.6%
Farmers, farm managers and farm laborers	7.2	8.7	8.6
Managers, officials, proprietors	12.4	13.4	12.8
Clerical workers	7.4	7.5	6.5
Sales workers	5.4	4.7	4.6
Craftsmen, foremen	19.7	18.5	16.9
Operatives	16.5	15.4	19.4
Service workers	7.7	9.4	9.4
Laborers	10.4	9.8	8.2
23. *Children under 13*			
0	51.7%	58.8%	53.5%
1	18.5	13.8	17.6
2	15.0	13.1	13.2
3	7.5	8.2	7.7
4	4.7	4.1	5.0
5 or more	2.6	2.0	3.0
$\bar{x}$	1.05	.93	1.04
Children 13–21			
0	66.9	69.3	73.1
1	18.8	15.3	13.9
2	10.3	9.6	9.0
3	2.6	4.2	2.7
4	1.0	1.1	0.7
5 or more	0.4	0.5	0.6
$\bar{x}$	.54	.54	.46
Children over 21			
0	75.4	67.1	69.9
1	11.2	12.5	10.8
2	6.8	9.4	8.2
3	3.4	5.1	4.5
4	1.4	2.3	2.3
5 or more	1.8	3.6	4.3
$\bar{x}$	.51	.76	.76

and Stuart who found only small differences on most items between probability samples with call-backs and uncontrolled quota samples [10]. For uncontrolled quota samples, the best explanation of close agreement with Census data and probability samples is that the interviewing biases are not highly related to responses. For the probability samples with quotas described here, some of the agreement with the call-back samples may also be attributed to the low correlation between sample bias and response, but, most important, the sample biases themselves are small.

The questions dealing with world tensions show mixed results when comparisons are made between the three surveys. On the first question dealing with "world tensions at this time" there are no differences between the three samples. On the second question dealing with "anticipated world tensions two years from now" the sample with quotas is significantly more optimistic than the call-back samples, and this greater optimism is also seen in the question dealing with "world tensions five years from now." Conversely, respondents on the probability sample with call-backs perceived "tensions two years ago" as being higher than did respondents on the call-back samples. Looking at item 5 which summarizes the first four questions, one can conclude that quota respondents believe that world tensions are being and will be steadily reduced. The call-back respondents show no clear pattern, but there is no reason that they should. The final item on tensions asks "when will the cold war end?" Here the call-back samples are more optimistic than the quota sample. It is possible that the assassination of President Kennedy a month earlier may have affected the answers of the probability sample with quotas, but if one disregards this, the possible biases due to differential availability range from 0 to 10 per cent on the four items, averaging about 6 per cent, which agrees with the data in Table 7.

Obviously, it would be desirable to have additional substantive questions for comparison, but unfortunately minor or major wording changes in the remaining items as well as different items make such comparisons impossible. There remain then the demographic items which have the virtue of being asked in the same way on all three studies.

Of the 23 demographic comparisons, there are no differences except for sex and household size. Naturally since there is a quota on sex, the probability sample with quotas matches Census data. The call-back samples were both deficient in males because of non-cooperation. That is, since the cooperation rate among men is lower than among women in the ordinary call-back sample, the sample with quotas is superior on this characteristic.

The comparisons of household size suggest that the quota sample is deficient in 1 and 2 member households. These results suggest that some of the remaining availability bias in the quota sample could be eliminated by imposing a household size control in addition to the controls now in effect. NORC plans to experiment with this control in future surveys.

REFERENCES

[1] Birbaum, Z. W. and Sirken, Monroe, "On the Total Error Due to Non-Interview and to Random Sampling," *International Journal of Opinion and Attitude Research*, IV (Summer, 1950) pp. 179–91.

[2] Cochran, William, *Sampling Techniques* (New York: John Wiley and Sons, 1953).

[3] Deming, W. Edwards, *Sample Design in Business Research* (New York: John Wiley and Sons, 1960).

[4] Durbin, J. and Stuart, A., "Call-backs and Clustering in Sample Surveys: An Experimental Study," *Journal of the Royal Statistical Society A*, CXVII (Part 4, 1954), pp. 387–428.

[5] Feller, William, *An Introduction to Probability Theory and Its Applications:* I (New York: John Wiley and Sons, 1950).

[6] Hansen, Morris, Hurwitz, William, and Madow, William, *Sample Survey Methods and Theory, Volume I* (New York: John Wiley and Sons, 1953).
[7] Kish, Leslie, *Survey Sampling* (New York: John Wiley and Sons, 1965).
[8] Low, Francis and McCormick, Thomas, "Some Survey Sampling Biases," *The Public Opinion Quarterly* XIX (Fall, 1955), pp. 349–405.
[9] Mayer, Charles, *Interviewing Costs in Survey Research* (Ann Arbor, Michigan: Bureau of Business Research, University of Michigan, 1964).
[10] Moser, C. A. and Stuart, A., "An Experimental Study of Quota Sampling," *Journal of the Royal Statistical Society A*, CXVI (Part 4, 1953).
[11] Nehnevajsa, Jiri, "The Near System: A Study in Public Acceptance," Office of Civil Defense, Department of Defense, February, 1964.
[12] Perry, Paul. "Election Survey Procedures of the Gallup Poll," *The Public Opinion Quarterly*, XXIV (Fall, 1960), pp. 531–42.
[13] Politz, Alfred and Simmons, Willard, "An Attempt to Get the 'Not at homes' into the Sample without Call-backs," *Journal of the American Statistical Association*, XLIV (1949), pp. 9–31.
[14] Stephan, Fredrick and McCarthy, Philip, *Sampling Opinions* (New York: John Wiley and Sons, 1958).
[15] Williams, Robert, "Probability Sampling in the Field," *The Public Opinion Quarterly*, XIV (Summer, 1950), pp. 316–30.