Charles D. Palit, University of Wisconsin-Madison

With the growth of the U. S. telephone system, telephone surveying has become a practical means of scientifically collecting data. With growth in the cost of travel substantially increasing the cost of personal interviews, telephone surveying has become increasingly the data collection mode of choice. This, in turn, has led to the development of a class of sampling designs known as random digit dialing (RDD) designs. While most of the RDD designs yield samples for which valid statistical estimators can be constructed, some do not. In addition, innovative designs, notably the Waksberg-Mitofsky type which yields valid statistical estimators, force us to reconsider the basic definition of a probability sample.

In its most primitive form, random digit dialing procedures involve the generation of a random or simple stratified random sample from the set of all possible phone numbers for the target area. For example, suppose we wish to sample from the central office code 231-xxxx, then there is a total of 10,000 possible phone numbers. These are the numbers between 231-0000 and 231-9999. If we select one or more four digit numbers "at random" and attach them to the central office code "231" then we have a random sample. If we do such a selection independently for a number of central office codes using the same sampling rate, or selection probability, then we have a simple stratified RDD sample.

Usually not all of the sample phone numbers will represent actual sample points. Some of the possible phone numbers will have no phones attached and some of the numbers will point to other ineligible phones. For example, in a survey requiring residential numbers, business numbers may be ineligible. This phenomenon means that RDD samples may contain a high proportion of nonsample phone numbers. Ratios of 7 or more to 1 are not only possible, but common. The ratio of total numbers from which interviews are obtained to all numbers in the sample is called the hit rate. Again for straight RDD samples, hit rates of 1 in 8 are quite usual. For convenience, we will call the hit rate for a simple RDD sample the "natural" hit rate.

Improving the hit rate is, of course, one of the concerns of the sampling statistician since it would seem to be a very direct way of increasing the design efficiency. As usual with such matters, we must watch the price paid for this improvement. In fact, the gain or improvement may be pure illusion, and not improve design efficiency at all. The prudent statistician should, therefore, examine with care the consequences of possible improvements before adopting them.

Prescriptions for improvements have been offered by Sudman (1973), Mitofsky and Waksberg (1978), and others. In general, these procedures use clustering mechanisms to improve the hit rate. Clustering unfortunately puts a complex probability structure on the design. This complicates later analysis, and reduces the efficiency of the design in terms of precision of estimation for the given sample size.

In general the larger the cluster size the better the hit rate, but this is not a free lunch, for it is also true that the larger the cluster size, the worse the design efficiency, as measured for example by the deft, where

$$\text{Deft} = \sqrt{\frac{\text{Actual Sampling Variance}}{\text{Variance for a Simple Random Sample of the Same Size.}}}$$

National samples using the Waksberg-Mitofsky design achieve good hit rates at the expense of design efficiency. For example, Groves & Khan (1979) in one national sample report average defts on the order of 1.30.

In spite of this phenomenon, the overall cost-effectiveness of the Waksberg-Mitofsky design in the sense of dollar cost per unit of precision will often be better than a simple stratified RDD design. However, this is not always the case. With surveys with long interviews, it is possible for the simple stratified RDD design to outperform the Waksberg-Mitofsky design. Groves & Khan (1979) perform an analysis which shows a break-even point at around 33 to 1 for a national sample with cluster sizes of 8, a natural hit rate of 1 in 5, and an empirically derived deft of 1.37. That is, if the cost of processing the working number exceeds 33 times the cost of processing the non-working number, the simple stratified RDD sample will outperform the Waksberg-Mitofsky design. If all else is held constant, then a lower natural hit rate, say 1 in 4, moves the break-even point in favor of the simple stratified RDD sample, while an upward movement moves the break-even point in favor of the Waksberg-Mitofsky design. The moral here is that even if our choice of designs were limited to a choice between a simple RDD design and a clustered design such as the Waksberg-Mitofsky design, there will be times when the simple stratified RDD design will be the design of choice. Longer interviews and lower natural hit rates point to the use of a simple stratified RDD instead of the Waksberg-Mitofsky design.

Of course, our choice of designs is not limited in this fashion. There is an alternative sampling strategy that will often yield better results than either of the two classes of designs just mentioned.

The alternative is to use stratification and disproportionate allocation instead of clustering techniques to improve the hit rate. Using stratification, hit rates equivalent to those obtained by the Waksberg-Mitofsky design can often be achieved with less loss in design efficiency. In addition, such designs can be easily used to construct a Master Sample which can be (more or less) continuously updated for changes in the universe, and which—with each succeeding survey—provides additional data for its own improvement. The runtime logistics of this design are considerably cheaper than those of the Waksberg-Mitofsky design.

In its simplest form, this design is implemented by dividing the universe of all possible telephone numbers into two strata—one for phone numbers which are judged likely to be attached to eligible sample points, and the other for phone numbers which are judged unlikely to be attached to eligible sample points. Both strata are sampled, but at different rates. Hit rate improvements are generated by sampling the "likely" stratum at a heavier rate. For example, if the "likely" sampling rate is 1/I, then the sampling rate for the "unlikely" stratum is set at

$$\frac{1}{KI}$$

where $K > 1$.

Table 1 shows the results of using such a design for a statewide cross-sectional sample in Wisconsin.

TABLE 1

Response Results from a Statewide Cross-sectional Sample in Wisconsin Using a Disproportionate Stratified RDD Design*

	Stratum		
	Likely	Unlikely	TOTAL
Sample size	2810	623	3433
Ineligible	1426	619	2045
Never answered	92	2	94
Non-response answered	230	0	230
Completed	1062	2	1064

*Source: Wisconsin Survey Research Laboratory Project 1215.

Three thousand four hundred and thirty-three phone numbers were required to produce 1,064 completed interviews. The hit rate of one in 3.2 is about par with that achieved by cluster designs. If a simple stratified RDD sample had been used, the hit rate would have been one in 6. What about its design efficiency?

If we assume that the variance for the elements in the "unlikely" stratum is the same as the variance in the "likely" stratum, and that we have a random sample of elements from each stratum, then the deft is approximately

$$\text{Deft} = \sqrt{\frac{(1 + K^2\pi)\ (1 + \pi)}{(1 + K\pi)\ (1 + K\pi)}}$$

where $\pi = \dfrac{\text{Number of eligible sample in "unlikely" stratum}}{\text{Number of eligible sample in "likely" stratum}}$

$$K = \frac{\text{Sampling rate in "likely" stratum}}{\text{Sampling rate in "unlikely" stratum}}$$

Using this, we can compute a deft for the sample reported on in Table 1. Since, in practice, proportionate stratified random samples are selected from within the "likely" and "unlikely" strata, and since the assumption of simple random sampling used to derive the expression for deft in (1) ignores any beneficial effect from sub-st tion within the "likely" and "unlikely" can look on this expression as an approx upper bound for the design's deft. For at hand, the deft is about 1.02—lower t average defts reported for the Waksberg-l design where the natural hit rate is lower, i.e., 1 in 5 and very close to that of a random sample.

Naturally the value of deft could have been reduced by choosing lower values of K, but deft choosing a lower value for K would lower the hit rate.

TABLE 2

Synthesized Hit Rates and Defts for the Statewide Sample as a Function of K

Ratio of Sampling Rates in Likely and Unlikely Strata	Hit Rate	Deft (Upper Limit)
1	1 in 6	1
2	1 in 4.4	1.00
3	1 in 4.0	1.00
4	1 in 3.5	1.01
6	1 in 3.2	1.02

Table 2 shows some synthesized hit rates and defts for approximately equivalent samples using different K's. It gives us some idea of what the trade off pattern looks like. As you can see, if K had been 3, the hit rate would have been 1 in 4 and the deft 1.00—essentially that of a random sample.

Not all separations are so successful. As π increases, so will the deft. Even so, there will be many occasions when the disproportionate stratified sample is the design of choice. This is particularly true if repeated surveys are to be performed over the same area; if only because this procedure lends itself well to the construction of a Master Sample from which work samples can be peeled off and used as needed.

The general procedure for constructing a Master Sample is relatively simple. To construct a Master Sample, we select blocks of numbers instead of numbers, usually 100 per block, and peel off numbers from the sample blocks as required for working samples. Three major strata are used in the Master Sample: likely, new-likely, and unlikely. The new-likely stratum is needed for the sequential procedure used to improve the Master Sample after each survey. The likely and unlikely strata are defined as before. The new-likely stratum contains those sample blocks which are initially in the unlikely stratum and for which it has been discovered to contain working or eligible numbers. Blocks in the new-likely stratum are sampled at a rate which lies between the rates used for the likely and unlikely strata, e.g., 2/KI. The price paid here is a moderate amount of clustering in the new-likely stratum.

Periodically, the Master Sample must be updated by adding blocks from newly-created exchanges. Obviously more can be said about the construction and use of master samples, but I believe that this brief description provides a good overview of the general procedure.

Summary

RDD design options have developed to a point where no one design is automatically the design of choice for all occasions. Two of the factors which must be considered are length of interview and the natural hit rate. High natural hit rates favor the simple RDD sample, while lower rates favor the use of more complex clustered designs. Long interview lengths work in favor the simple RDD design. However, disproportionate stratified designs promise: hit rates equivalent to cluster designs, simpler analysis patterns, and simpler administration logistics.

References

Frankel, Martin R. and Lester R. Frankel. (1977) "Some Recent Developments in Sample Survey Design," Journal of Marketing Research, Vol. XIV, pp. 280-293.

Groves, R.L. and R. M. Khan. (1979) "Surveys by Telephone," Academic Press.

Kish, L. (1965) "Survey Sampling," John Wiley & Sons, Inc.

Klecka, William R. and Alfred J. Tuchfarber. (1978) "Random Digit Dialing: A Comparison to Personal Surveys," Public Opinion Quarterly, Vol. 42, pp. 105-114.

Lahiri, D. B. (1951) "A Method of Sample Selection Providing Unbiassed Ratio Estimates," Bulletin, International Statistical Institute, 33.

Sudman, S. (1973) "The Uses of Telephone Directories for Survey Sampling," Journal of Marketing Research, Vol. X, pp. 204-207.

Tyebjee, T. T. (1979) "Telephone Survey Methods: The State of the Art," Journal of Marketing Research, Vol. 16, pp. 96-101.

Waksberg, J. (1978) "Sampling Methods for Random Digit Dialing," Journal of the American Statistical Association, Vol. 73, pp. 40-46.

Acknowledgment

Thanks to Bob Groves for his support and for pointing out at least one of my errors.