Sampling Methodology and
ASDE Survey Sampler

By Michel Rochon

The following general statement broadly defines the methodology used by ASDE to draw telephone samples.

> "ASDE Survey Sampler is a geographically stratified, general, phone population, random sampling by RDD methodology. Stratification is based on each geographical level where sample is requested. Within each geography, area codes and prefix combinations are assigned, based on their weight to one or more geography, therefore providing inclusion and avoiding any bias of exclusion that could result from using a plurality decision rule."

Survey Sampler has been perfected in Canada since 1995 where it is now used by a large majority of the best known and most sophisticated survey research organizations.

According to the definitions currently used in the USA our samples are "Random B" samples. The following definitions will help:

1. Random A samples are samples of random numbers systematically selected with equal probability across all eligible telephone blocks.
2. Random B samples are samples of random numbers distributed across all eligible blocks in proportion to their density of listed telephone households.

The difference between Random A and Random B is that:

Random A samples use every eligible block even when the blocks have no numbers associated with them in phone books, while the Random B samples use only those blocks which have a least one listed number.

Random B samples are more efficient, have a very high projectability and are sufficient for most research studies. Because these samples are selected from blocks according to their density of listed telephone households, (not from all blocks equally, as for a Random Digit A sample), there is a possibility that highly unlisted areas may be underrepresented. However, the tradeoff is in efficiency. The working phone rate is much higher in Random B samples.

ÉCHANTILLONNEUR
asde
SURVEY SAMPLER

729 St-Joseph, Suite 201
Gatineau, QC
J8Y 4B6

819-770-3651
888-323-3651
Fax: 819-770-3688
www.surveysampler.com
www.echantillonneur.com

ASDE RDD samples

The RDD frame of ASDE Survey Sampler is based on a complete analysis of 120 million phone numbers listed in North American phone books, and the geocoding of all telephone exchanges (area code + three first digits of phone number xxx-nnn) in each country. This analysis is redone every 6 months with each issue of the Acxiom Infobases electronic database of phone numbers. The electronic directories are supplemented by the Telcordia list of all working exchanges in North America which is issued every month.

The phone numbers are analyzed at the 100 block level: xxx-xxx-xnnn, each block is counted for residential and business numbers, the lowest and highest number allocated in each block is found. Business and residential numbers are the object of separate counts.

For every geography chosen by the user, ASDE Survey Sampler assembles all blocks belonging to that geography and totals up the number of phones in service. (N) For the required sample size (n) in each geography, ASDE Survey Sampler uses a systematic sampling procedure to select a starting block with probability proportional to the number of DL phone numbers

Within each such block it then randomly generates phone numbers between the smallest and largest known numbers until exactly one number is found. The program then advances (N/n) to the next block, (or if N/n < block size within the same block), and repeats the selection process. Each number generated is checked against a licensed database of directory listed phone numbers, and the number is flagged as directory listed (DL) or directory not listed (DNL) and marked as either business4 or residential.

For **residential samples**, business numbers are rejected and replaced from the same working block. For **business samples**, residential numbers are rejected and replaced from the same working block. The appropriate information (name, address, and postal code) is retained when the number has been found to be directory listed.

For **not listed numbers**, a proximate zip/postal code is added. All numbers are then geocoded for all pertinent census and electoral boundaries. The DL and DNL together constitute the RDD sample.

Weights are applied to the DNL sample based on the density of the listed phones in the block where the DNL number has been generated.
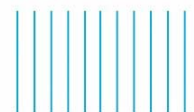
Evaluations

The program and all its underlying assumptions and outputs have been thoroughly evaluated by an independent outside expert, Mr. Alan Sunter a fellow of the American Statistical Association and one of Canada's foremost authorities on sampling methodology and sampling frame evaluation. Mr. Sunter is regularly provided with all specifics concerning the workings of ASDE Survey Sampler, its internal data base structures and its algorithms. He has full access to the operational versions of the program as well as to samples drawn with full audit trail of the program's internal workings during the sampling.

He has confirmed without reservations that the samples provided by Survey Sampler are statistically equivalent to a RDD sample while providing important cost advantages in the survey process.

In his **1996 evaluation** he noted:

> "You have asked me to compare your sampling procedure with the more usual Random Digit Dialling (RDD) procedure often specified as "required" in government "requests for proposal. Your procedure, as I have demonstrated in an earlier analysis, selects both listed and unlisted numbers with equal probability and without replacement within each strata (usually an exchange or a block of exchanges). The listed number sample is of fixed size, as determined by the sample allocation option chosen by the user; the unlisted number sample has some sampling variance in the sample size actually realized but this variance (which appears, in any event, to be unavoidable) has negligible impact in my view on the efficiency of the overall design, the sampling variances of stratum estimates will be equivalent to, or less than, those obtained from simple random samples of equivalent sizes."

In his **2001 evaluation** Mr. Sunter stated:

> "As in my (previous) reviews, I can confirm that the system selects members of the listed/unlisted categories in each stratum with equal probability and without replacement and that the sampling variances of stratum estimates will be equivalent to, or less than, those obtained from simple random samples of equivalent size."

Background notes on Mr. Alan Sunter

Alan Sunter is a consulting statistician whose experience includes five years as director, Business Survey Methods, Statistics Canada, two years as a consultant to the national statistical agencies of Sweden and the U.K., two years with the World Fertility Survey, and 12 years as a consultant to both private and public agencies in Canada as well as to statistical agencies in less developed countries. He has designed and managed many business surveys in Canada as well as household surveys in Bangladesh, Ivory Coast, Nigeria, Benin, Ethiopia, Trinidad and Tobago, and Colombia.
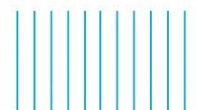
729 St-Joseph, Suite 201
Gatineau, QC
J8Y 4B6

819-770-3651
888-323-3651
Fax: 819-770-3688
www.surveysampler.com
www.echantillonneur.com

He specializes in survey design, survey data processing, and the measurement of both sampling and non-sampling errors in surveys, and has a considerable number of published theoretical statistical papers to his credit. He has been elected to the International Statistical Institute and as a fellow of the American Statistical Association for significant contributions to the theory and practice of survey design and development.