Component of Statistics Canada
Catalogue no. 12-001-X Business Survey Methods Division

Article

Evaluating within household selection rules under a multi-stage design

by Tom Krenzke, Lin Li and Keith Rust

June 2010

Statistics Canada Statistique Canada

Canada

Survey Methodology, June 2010
Vol. 36, No. 1, pp. 111-119
Statistics Canada, Catalogue No. 12-001-X

Evaluating within household selection rules under a multi-stage design

Tom Krenzke, Lin Li and Keith Rust [1]

Abstract

The 2003 National Assessment of Adult Literacy (NAAL) and the international Adult Literacy and Lifeskills (ALL) surveys each involved stratified multi-stage area sample designs. During the last stage, a household roster was constructed, the eligibility status of each individual was determined, and the selection procedure was invoked to randomly select one or two eligible persons within the household. The objective of this paper is to evaluate the within-household selection rules under a multi-stage design while improving the procedure in future literacy surveys. The analysis is based on the current US household size distribution and intracluster correlation coefficients using the adult literacy data. In our evaluation, several feasible household selection rules are studied, considering effects from clustering, differential sampling rates, cost per interview, and household burden. In doing so, an evaluation of within-household sampling under a two-stage design is extended to a four-stage design and some generalizations are made to multi-stage samples with different cost ratios.

Key Words: Intracluster correlation; Design effects; Multi-stage sampling.

1. Introduction

The 2003 National Assessment of Adult Literacy (NAAL), conducted for the National Center for Education Statistics, provided an indicator of the nation's progress in English literacy for researchers, practitioners, policymakers, and the general public. As in the 1992 National Adult Literacy Study (NALS), adults were assessed in households in prose, document and quantitative literacy. The booklet designs were based on the 1992 NALS to allow for the measurement of trends between 1992 and 2003.

In order to reduce the cost of interviewers traveling to households, the NAAL involved a stratified four-stage cluster design that resulted in 18,500 completed assessments administered to adults age 16 and older. In the NAAL, counties were grouped to form Primary Sampling Units (PSUs), which were stratified and selected in the first stage. In the second stage, Secondary Sampling Units (SSUs) were formed and selected within the sampled PSUs. The SSUs were individual census blocks, or groups of adjacent blocks with at least 60 households (HHs) formed within tract boundaries. Subsequently, households were selected within SSUs, and one sample person (1 SP) was randomly selected for household sizes up to $3 (B \leq 3)$, and two persons (2 SPs) were selected for household sizes greater than $3 (B > 3)$, where B denotes the number of eligible persons per household. This rule followed the within-household sampling approach used in the first cycle of NAAL (NCES 2001), conducted in 1992. An evaluation of the selection rule was conducted using the current US household size distribution and intraclass correlation coefficients computed from the 2003 survey. In doing so, an evaluation of within-household sampling under a two-stage design (Clark and Steel 2007) is extended to a four-stage design, as used in the NAAL survey and some generalizations are made to multi-stage samples with different cost ratios.

The data used for the evaluation include literacy measures from three scales derived from three types of literacy - prose, document, and quantitative. For more information about the NAAL types of literacy, refer to http://nces.ed.gov/NAAL/fr_tasks.asp. Two types of estimates are used; averages (*e.g.*, average prose literacy score) and percentage of adults at some level of literacy (*e.g.*, percentage *Below Basic* prose literacy). For a discussion of the literacy levels used in NAAL, see http://nces.ed.gov/NAAL/perf_levels.asp. In addition to the NAAL data, the evaluation also uses US sample data from the international Adult Literacy and Lifeskills (ALL), which was conducted by Statistics Canada. The US sample in 2003, sponsored by NCES, was part of a comparative study that measured the skills of adults in several countries. Similar to the NAAL, the ALL was a multi-stage clustered sample survey and measured prose and document literacy, as well as numeracy (OECD 2005). The NAAL sample was much larger (18,500 completes) than the ALL sample (3,400 completes), and the target population for NAAL included ages 16+ while the target population for ALL included 16 to 65 year olds. Table 1 provides a summary of each survey's design and structure.

1. Tom Krenzke, Statistical Group, Westat, Rockville, Maryland 20850. E-mail: tomkrenzke@westat.com; Lin Li, Statistical Group, Westat, Rockville, Maryland 20850. E-mail: linli@westat.com; Keith Rust, Statistical Group, Westat, Rockville, Maryland 20850. E-mail: keithrust@westat.com.

Table 1
Features of the NAAL and ALL surveys

Survey	Area sample	Completes	Data collection	Assessments	Ages	Within-HH sampling rule
NAAL	PSUs, SSUs households, Persons	18,500	Screener Interview Assessment	Prose Document Quantitative	16+	$B \le 3, b = 1$ $B > 3,$ $b = 2$
ALL	PSUs, SSUs, households, Persons	3,400	Screener Interview Assessment	Prose Document Numeracy	16-65	$B \le 3, b = 1$ $B > 3, b = 2$

Note: PSU = Primary Sampling Unit, SSU = Secondary Sampling Unit, b = sample size, B = household size.

A discussion of the design considerations that helped form the evaluation of the within-household sampling rules is provided in Section 2. Section 3 discusses the computation of intra-household correlations under multi-stage sample designs and focuses on incorporating the clustering impact from the initial stages of sample selection when deciding on a within-household selection rule. An evaluation of selection rules was conducted using data from the in-person adult literacy surveys and the results are provided in Section 4. Finally, a brief summary is given in Section 5.

2. Design considerations

There are a number of factors that need to be considered when evaluating the within-households selection rules for surveys such as NAAL and ALL. The remainder of this section will discuss the impact of the following factors on within-household sampling: household burden, clustering persons within households, differential sampling rates, multi-stage sampling, cost considerations, computerized systems, domains of interest and household composition.

Household burden. For the adult literacy surveys, the interview and the assessment take about an hour and a half to administer in total. Therefore, one concern about selecting more than one person per household is the increase of burden to the household and the impact on response rates. However, there is no significant difference (0.05 significance level) in the refusal rates between 1- and 2-SP households in ALL and NAAL as shown in Table 2.

Clustering persons within households. Kish (1965) discusses the benefits of a cluster sample to a simple random sample. A cluster sample typically has a lower cost per person, however the unit variance is higher and it causes greater complexities in statistical analysis. Kish introduced the concept of a design effect (DEFF), which measures the increase in variance due to deviations from a simple random sample, such as clustering persons within households. Many surveys limit the selection to one sample person (SP) per household because of concerns over the increased clustering effect (*i.e.*, increasing effect on variance estimates) associated with multiple SPs per household. The DEFF due to clustering can be expressed as: $\text{DEFF}_{\text{clu}} = 1 + (\bar{b} - 1)$ Rho, where $\bar{b} = \sum (M_B / M) b_B$, $M_B =$ number of households of size B, $M =$ number of households, and $b_B =$ sample size of persons within households of size B (Kish 1965). This DEFF component increases when the sample size within a household increases or when the value of the intracluster correlation (Rho) increases. As given in Cochran (1977), Rho can be approximated as:

$$\text{Rho} = 1 - \frac{\sigma_w^2}{\sigma^2},$$

where

$$\sigma_w^2 = \sum_{i=1}^{a} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.})^2 / (n - a),$$

and

$$\sigma^2 = \sum_{i=1}^{a} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2 / (n - 1),$$

where a is the number of sampled households, and b is the number of sampled persons per household. The DEFF due to clustering is examined further for different within-household sampling rules in the next section.

Differential sampling rates. A clustering effect is not the only factor that increases the variance. Increases in variance are also due to differential sampling rates (resulting in differential weights). Under a 1 SP per household strategy, the increase is directly related to the variation in household size since the sampling rate could vary from 1 out of 1 to 1 out of 7 or more. The DEFF due to differential sampling rates is expressed as: $\text{DEFF}_{\text{wgt}} = \sum (p_B / k_B) \sum (p_B k_B)$, where $p_B = N_B / N$, $N_B =$ number of eligible persons in the population in households of size B, $N =$ number of eligible persons in the population, and $k_B =$ sampling rate within households of size B (Kish 1965). Under certain conditions, the overall DEFF can be expressed as the product of the clustering and differential sampling rate components: $\text{DEFF} = \text{DEFF}_{\text{clu}} \times \text{DEFF}_{\text{wgt}}$. Kalton, Brick and Lê (2005) suggest this product is applicable when the weights are random or approximately random.

Table 2
Refusal rates by 1- and 2-SP households for the adult literacy surveys

Survey	Subgroup	Refusal rate %
NAAL	1-SP households	16.3
	2-SP households	15.7
ALL	1-SP households	17.6
	2-SP households	16.2

Note: SP = sample person.

To arrive at a self-weighting sample, persons within households would need to be selected at a constant rate. However, a rate-based approach is not preferred in most surveys since it would result in walking away from a portion of single-person households and, thus, would increase the cost of the survey. We limit the alternative rules under consideration to those with a minimum of 1 SP per household. Out of concern for burdening households, the maximum sample size was set to two. The sampling rules under consideration are:

1. Take1: 1 SP no matter the household size.
2. Rule2: 1 SP for household sizes up to 2; otherwise 2 SPs are selected.
3. NAAL3: 1SP for household sizes up to 3; otherwise 2 SPs are selected.
4. Rule4: 1 SP for household sizes up to 4; otherwise 2 SPs are selected.
5. Frac5: take at least 1 SP, but no more than 2 SPs and the sample size is a fraction. That is, if the sample size for a household with two eligible persons is 1.6, then two persons are selected 60 percent of the time at random, and one person is selected 40 percent of the time.

While the Take1 approach does not attempt to reduce the DEFF due to differential sampling rates, it is not subject to a clustering impact. However, the other four approaches listed above provide a reduction in the differential sampling rate component while introducing a clustering effect. In the case of Frac5, under the assumption that π-weights are used, as assumed throughout this paper, the approach would result in the most reduction in the differential sampling rate component. The π-weights approach is based on the unconditional selection probability of the person within the household. If the actual sample size within a household is used in the form of ratio weights, the differential sampling rate increases the benefit is less clear and depends on Rho. Figure 1 illustrates the best options under a two-stage household design with fixed effective sample size of persons, without any cost considerations. The US national household size distribution from the 2007 Current Population Survey was used for this illustration. As shown in Figure 1, the fractional approach is the best rule for a wide range of values of Rho. The fractional approach can be programmed into a computerized system when enumerating and selecting household members (more discussion on computerized systems follows). If computerized systems are not available for screening, then the best approach for low values of Rho is the more clustered approach, Rule2; and the NAAL3 rule is best for Rho values greater than about 0.34.

Multi-stage sampling. For multi-stage area designs, the clustering impact of sampling within households is affected by the clustering due to PSUs and SSUs. As pointed out by Kish (1965), the clustering of households and persons within PSUs and SSUs increases the sampling variance (*i.e.*, units within PSUs and SSUs are more similar to each other). The incremental impact of clustering within households may be dampened by the domination of the PSU and SSU variance components (however, the magnitude of the impact will differ depending on the type of estimate and variable). That is, more persons within a household can be selected for surveys with a large amount of clustering due to the first two stages of sampling. Details of this distinction are provided in Section 3.

Cost considerations. The cost of screening a household in a 1 SP per household design versus the cost of interviewing/assessing a second person in a household is investigated in an extensive analysis presented later .

Computerized systems. Computerized systems, such as Computer-Assisted Personal Interview (CAPI), have the capability of handling fractional sample sizes. That is, the random selection of 1 or 2 SPs given a pre-assigned fractional sample size can be programmed. Computerized systems also have the capability of sorting the list of eligible persons and selecting 2 SPs with a systematic random sample. Another benefit is that the selection program can be tested and validated prior to data collection.

Domains of interest. As mentioned earlier, optimal within household sampling depends on the magnitude of the clustering effect associated with the variable of interest. The clustering effect may be much smaller when the variable is associated with a subgroup of the population, rather than the entire population. For example, when a key reporting domain is gender in a survey of the adult population, the reporting category of males is likely to have an average of 1 SP per household and less likely to have 2 male SPs which would introduce a clustering effect. Therefore, when there are multiple domains of interest in a typical household, it is often beneficial to select more than 1 SP within a household. Refer to Mohadjer and Curtin (2008) for an example of design considerations for a survey with focus on multiple subgroups of the population.

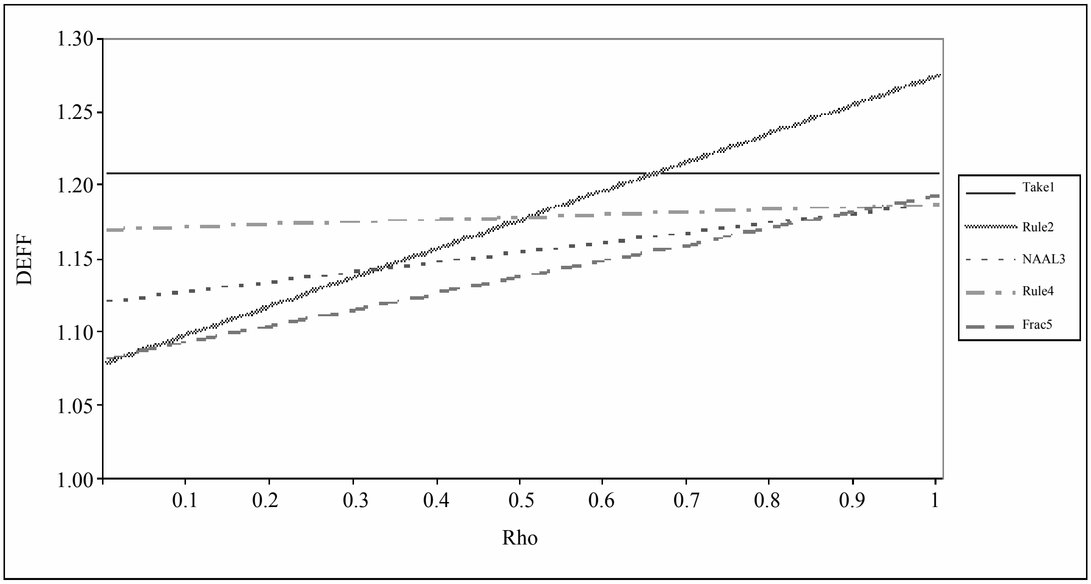

Figure 1 Initial analysis of within-household selection rules

Household composition. Lastly, one may want to consider the household composition and relationships of persons within a household when devising the selection rule. Table 3 displays values of Rho for various relationships between household members, for household with 2 SPs in the NAAL survey. Rho varies greatly by household member relationships. The relationships were derived from gender and age.

3. Estimation of *intra-household* Rho and DEFF under multi-stage sampling

The discussion about Rho thus far has been related to a two-stage design, but both NAAL and ALL have four stages of sampling. The total variance can be decomposed into four between-variance terms attributable to PSUs, SSUs, households and persons, as follows:

$$\sigma_T^2 = \sigma_{\mathrm{PSU}}^2 + \sigma_{\mathrm{SSU(PSU)}}^2 + \sigma_{\mathrm{HH(SSU)}}^2 + \sigma_{\mathrm{PERS(HH)}}^2.$$

As shown below, when applying a two-stage approach to estimate Rho for a four-stage sample design, the numerator not only contains the between household component, but also contains contributions from the between PSU and between SSU components inflating the values of Rho for our purpose.

$$\mathrm{Rho} = 1 - \frac{\sigma_{\mathrm{PERS(HH)}}^2}{\sigma_T^2} = \frac{\sigma_{\mathrm{PSU}}^2 + \sigma_{\mathrm{SSU(PSU)}}^2 + \sigma_{\mathrm{HH(SSU)}}^2}{\sigma_T^2}.$$

Therefore, when evaluating rules for within-household sampling under a multi-stage design, we assume the PSU and SSU design will be the same in the future. This can be accomplished by limiting our focus to within SSU sampling. Therefore, the computation of Rho is contained within SSUs, that is, it is done in a compact manner without effect from the PSU and SSU components. We refer to this as the compact (*i.e.*, within SSU) Rho denoted by Rho^*, expressed as:

$$\mathrm{Rho}^* = \frac{\sigma_{\mathrm{HH(SSU)}}^2}{\sigma_{\mathrm{HH(SSU)}}^2 + \sigma_{\mathrm{PERS(HH)}}^2}.$$

Using the compact Rho^*, we now derive the estimated DEFF under a multi-stage sample design for the purpose of determining optimal within-household sample sizes. The variance of an estimate $(\hat{\theta})$ with b persons per household can be decomposed as:

$$\mathrm{Var}(\hat{\theta}) = \frac{\sigma_{\mathrm{PSU}}^2}{n_{\mathrm{PSU}}} + \frac{\sigma_{\mathrm{SSU(PSU)}}^2}{n_{\mathrm{SSU}}} + \frac{\sigma_{\mathrm{HH(SSU)}}^2}{n_{\mathrm{HH}}} + \frac{\sigma_{\mathrm{PERS(HH)}}^2}{bn_{\mathrm{HH}}}$$

where, n_{PSU}, n_{SSU}, n_{HH} and bn_{HH} are the sample sizes of PSUs, SSUs, households and persons, respectively.

Table 3
Rho for NAAL assessment scores by household member relationships

Estimate	Siblings	Child-guardian	Married	Others
Number of households with 2 SPs	111	205	180	434
Average prose score	0.42	0.35	0.70	0.59
Average document score	0.40	0.27	0.72	0.54
Average quantitative score	0.46	0.36	0.63	0.56
Percentage Below Basic prose	0.52	0.41	0.79	0.67
Percentage Below Basic document	0.54	0.40	0.78	0.60
Percentage Below Basic quantitative	0.51	0.41	0.77	0.65

Then the DEFF due to clustering, relative to taking one person per household and bn_{HH} households is:

$$\text{DEFF}_{\text{clu}}^{\text{HH}} = \frac{\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}} + \dfrac{\sigma^2_{\text{HH(SSU)}}}{n_{\text{HH}}} + \dfrac{\sigma^2_{\text{PERS(HH)}}}{bn_{\text{HH}}}}{\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}} + \dfrac{\sigma^2_{\text{HH(SSU)}}}{bn_{\text{HH}}} + \dfrac{\sigma^2_{\text{PERS(HH)}}}{bn_{\text{HH}}}}$$

$$= \frac{\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}} + \dfrac{1}{bn_{\text{HH}}}(\sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}} + (b-1)\,\sigma^2_{\text{HH(SSU)}})}{\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}} + \dfrac{1}{bn_{\text{HH}}}(\sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}})}$$

$$= \frac{\dfrac{bn_{\text{HH}}\left(\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}}\right)}{\sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}}} + (1 + (b-1)\,\text{Rho}^*)}{\dfrac{bn_{\text{HH}}\left(\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}}\right)}{\sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}}} + 1}$$

$$= \frac{k^* + (1 + (b-1)\,\text{Rho}^*)}{k^* + 1}$$

where,

$$k^* = \frac{bn_{\text{HH}}\left(\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}}\right)}{(\sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}})} = \frac{\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}}}{\dfrac{1}{bn_{\text{HH}}}(\sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}})}$$

Alternatively, $\text{DEFF}_{\text{clu}}^{\text{HH}}$ can be expressed as:

$$\text{DEFF}_{\text{clu}}^{\text{HH}} = 1 + \frac{(b-1)\,\text{Rho}^*}{k^* + 1} = 1 + (b-1)\,\text{Rho}^{**}$$

where,

$$\text{Rho}^{**} = \frac{\text{Rho}^*}{k^* + 1}$$

$$= \frac{\left(\dfrac{\sigma^2_{\text{HH(SSU)}}}{\sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}}}\right)\dfrac{1}{bn_{\text{HH}}}(\sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}})}{\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}} + \dfrac{\sigma^2_{\text{HH(SSU)}}}{bn_{\text{HH}}} + \dfrac{\sigma^2_{\text{PERS(HH)}}}{bn_{\text{HH}}}}$$

$$= \frac{\dfrac{\sigma^2_{\text{HH(SSU)}}}{bn_{\text{HH}}}}{\dfrac{\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}} + \dfrac{\sigma^2_{\text{HH(SSU)}}}{bn_{\text{HH}}} + \dfrac{\sigma^2_{\text{PERS(HH)}}}{bn_{\text{HH}}}}$$

$$= \frac{\sigma^2_{\text{HH(SSU)}}}{\dfrac{bn_{\text{HH}}\sigma^2_{\text{PSU}}}{n_{\text{PSU}}} + \dfrac{bn_{\text{HH}}\sigma^2_{\text{SSU(PSU)}}}{n_{\text{SSU}}} + \sigma^2_{\text{HH(SSU)}} + \sigma^2_{\text{PERS(HH)}}}$$

The Rho^{**} measure is a useful expression for the intra-household correlation under a multi-stage design, which is equal to Rho^* when $\sigma^2_{\text{PSU}} = \sigma^2_{\text{SSU(PSU)}} = 0$. The compact Rho^* measure is useful for evaluating optimal sample sizes while varying the variance ratio k^*. Note, however, that in general Rho^{**} is a function of n_{PSU}, n_{SSU} and the total sample size of persons, whereas Rho^* does not depend on these.

As shown in Table 4, the variance ratio k^*, which is the variance from the first two stages divided by the variance from the last two stages, for a one person per household design, ranges from 0.68 to 1.61 across types of assessments and estimates for the ALL survey.

Table 5 shows estimates for Rho (computed under a two-stage design assumption), the compact Rho^* and Rho^{**} (computed under a multi-stage design assumption where $k^* = 1$) for average NAAL and ALL literacy assessment scores. When including the clustering impact from the first two stages of the four-stage design, the values of the compact Rho^* and Rho^{**} are much smaller than Rho. For example, the two-stage Rho for the NAAL average prose score is 0.57 and the compact Rho^* is equal to 0.33 and Rho^{**} is equal to 0.17. The table also shows that values of the compact Rho^* for average scores are at about the same level for NAAL (range from 0.32 to 0.33) and ALL (range from 0.29 to 0.39). There is some variation by the type of estimate as well; values of Rho^* for ALL are 0 to 0.2 lower for the percentage in Level 1 or 2 than for the average scores. Values of Rho^* can also vary by household size as shown in Figure 2 in Appendix A.

4. Evaluation and results

We compared the current sampling rules with optimal sampling rules by minimizing a variance-cost (VC) function, which is the product of the DEFFs (*i.e.*, variance increase) due to clustering and weighting, and a cost function that is used by Kish (1965):

$$\text{VC} = \text{DEFF}_{\text{clu}}^{\text{HH*}} \times \text{DEFF}_{\text{wgt}} \times n\left(c_p + \frac{c_{\text{HH}}}{\bar{b}}\right),$$

where c_p = cost per added person and c_{HH} = cost per added household. Note that $n/\bar{b}$ represents the number of sampled households. To account for the differential clustering effects for each household size B, we replace $\text{DEFF}_{\text{clu}}^{\text{HH}}$ with:

$$\text{DEFF}_{\text{clu}}^{\text{HH*}} = \frac{k^* + \sum_B \frac{M_B}{M}(1 + (b_B - 1)\,\text{Rho}_B^*)}{k^* + 1}$$

where Rho_B^* is computed as described in Appendix A.

Note that the VC function represents the additional cost of increasing the overall sample size to offset the increase in variance due to the DEFF components. Table 6 provides the results for optimal integer solutions as computed by a computational algorithm which is described in Appendix B. The table shows that as the cost ratio increases from 0.5 to 1 for $k^* = 1$, we would want to take more persons per household, that is, 2 out of 2 instead of 1 out of 2. As the variance ratio goes from 1 to 3 for optimal integer solutions, the only change is for household size of 2 and cost ratio of 0.5. That is, when the variance ratio is equal to 3, it is beneficial to take 2 out of 2 instead of 1 out of 2.

Table 6 also gives the results when fractional sample sizes are allowed. The variance and cost ratios for NAAL and ALL tend to be about 1, where it appears that selecting 1 out of 1, 1.6 out of 2, and 2 otherwise is the best rule. The effects of cost and variance ratios are clearer under the fractional sample sizes when compared to the integer solutions.

If the cost of conducting a screener is small in relation to the cost of interviewing, then variances can be reduced using the fractional walk-away approach. Table 6 shows optimal walk-away sample sizes. Under this approach, for example, a sample size of 0.9 indicates that we walk away from 10 percent of the households where $B = 1$. If the cost of screening is a very small portion of the cost of interviewing, then the optimal design may involve walking away from many more households.

Under the likely NAAL/ALL parameters for cost ratios $(C_{\text{HH}}/C_p = 1)$ and variance ratios $(k^* = 1)$, when compared to the Take1 approach, the VC function can be reduced by about 9 percent by using the NAAL/ALL sampling rule, 19 percent by using the optimal integer solution, 20.4 percent using the optimal fractional solution, and 20.6 using the optimal walk-away approach. In general, the gains from deviating from the Take1 approach grow as the cost per additional households (*i.e.*, screening) increases. The average cluster sizes for each approach are given in Table 7. For the NAAL and optimal integer rule, the average cluster size indicates the percentage of households with 2 SPs. For example about 6 percent of the households would have 2 SPs under the NAAL3 strategy.

Table 4
Values of k^* for the ALL sample

ALL estimate	k^*
Average prose score	0.95
Average document score	1.56
Average quantitative/numeracy score	1.13
Percentage in Level 1 or 2 prose	0.68
Percentage in Level 1 or 2 document	1.61
Percentage in Level 1 or 2 numeracy	1.10

Table 5
Values for Rho, Rho^*, and Rho^{**} for literacy assessment scores

	Rho		Rho^*		Rho^{**}	
Estimate	**NAAL**	**ALL**	**NAAL**	**ALL**	**NAAL**	**ALL**
Number of households with 2 SPs	930	162	930	162	930	162
Average prose score	0.57	0.60	0.33	0.38	0.17	0.19
Average document score	0.53	0.50	0.33	0.29	0.17	0.15
Average quantitative/numeracy score	0.54	0.58	0.32	0.39	0.16	0.20
Percentage Below Basic(NAAL)/Level 1or 2 (ALL) prose	0.65	0.44	0.42	0.28	0.21	0.14
Percentage Below Basic (NAAL)/Level 1 or 2 (ALL) document	0.61	0.37	0.39	0.28	0.20	0.14
Percentage Below Basic quantitative (NAAL)/Level 1 or 2 (ALL) numeracy	0.62	0.36	0.40	0.17	0.20	0.09

Note: Rho^{**} is computed assuming $k^* = 1$.

Table 6
Optimal expected number of persons per *household* by type of person sampling method and *household* size (B)

		Person Sampling Method											
		Integer				Fractional				Walk-away			
k^*	C_{HH} / C_p	$B = 1$	$B = 2$	$B = 3$	$B = 4$	$B = 1$	$B = 2$	$B = 3$	$B = 4$	$B = 1$	$B = 2$	$B = 3$	$B = 4$
1	0.5	1	1	2	2	1	1.4	2	2	0.6	1.3	2	2
1	1	1	2	2	2	1	1.6	2	2	0.9	1.6	2	2
1	2	1	2	2	2	1	1.9	2	2	1	1.9	2	2
3	0.5	1	2	2	2	1	1.6	2	2	0.8	1.5	2	2
3	1	1	2	2	2	1	1.8	2	2	1	1.8	2	2
3	2	1	2	2	2	1	2	2	2	1	2	2	2

Table 7
Percent reduction of NAAL3 and optimal solutions from Take1 strategy and average cluster sizes

		Percentage reduction from Take1 strategy				Average cluster sizes			
k^*	C_{HH} / C_p	NAAL3	Integer	Fractional	Walk- away	NAAL3	Integer	Fractional	Walk- away
1	0.5	8.2	13.0	15.8	18.0	1.06	1.18	1.38	1.21
1	1	9.1	19.2	20.4	20.6	1.06	1.68	1.48	1.45
1	2	9.9	26.1	26.1	26.1	1.06	1.68	1.63	1.63
3	0.5	8.6	17.3	18.7	19.0	1.06	1.68	1.48	1.37
3	1	9.5	23.7	23.9	23.9	1.06	1.68	1.58	1.58
3	2	10.4	30.2	30.2	30.2	1.06	1.68	1.68	1.68

Lastly, a sensitivity analysis was conducted by varying the values of Rho*. A regression model was fit on the percentage reduction from the Take1 strategy of the VC function, with the independent variables being the approach (NAAL3, integer, fractional, walk-away), cost ratio (0.1, 0.5, 1, 2, 10), variance ratio (1, 3, 5) and Rho* (+/- 0.1). For the range of data, Rho* had a limited impact (parameter estimate -7.4 with an associated standard error of 4.5) on the percentage reduction of the VC function, while the other factors had more of an impact.

5. Summary

Several design considerations were taken into account when evaluating the within - household selection rule for the NAAL and ALL surveys, including taking into account clustering effects from initial stages of sampling. To facilitate the evaluation, we formulate a way to incorporate PSU and SSU variance contributions into the computation of the DEFF due to clustering and the intra-household correlation when deciding how many persons and how many households to select in a multi-stage sample design. In doing so, we introduce compact Rho* measure, which is computed within the SSU so it is not impacted by the PSU and SSU variance components. This is useful when determining the DEFF due to clustering within households, while varying the contribution to the total variance from the PSU and SSU stages of selection in multi-stage sample designs. The measure Rho** is introduced as an expression for the intra-household correlation under a multi-stage design, taking into consideration the contribution to total variance from the first two stages of selection.

In addition, a computational algorithm was developed to compute optimal sample size solutions, incorporating the DEFFs due to clustering, differential sampling rates, and costs.

In general, the main factors on the percentage reduction of the VC function from the Take1 approach are the level of dominance from the PSU and SSU variance components in multi-stage sampling, the cost ratio and the rule used. For the range of data evaluated, Rho* had limited impact on the reduction in VC from the Take1 approach. In general, the NAAL rule improves on the widely-used Take1 approach. The optimal integer rule improves on the NAAL rule. However, the optimal fractional rule has limited gains over the optimal integer rule. The optimal walk-away rule has gains over the other rules for lower cost ratios. Lastly, when the first two variance components dominate and cost ratio is high, then the integer, fractional and walk-away rules are essentially the same.

Acknowledgements

The authors acknowledge valuable contributions by Leyla Mohadjer and Bob Fay.

Appendix A

Estimates of Rho* by household size

Survey estimates are not attainable for Rho* by household size since only 1 SP was selected for household size of 3 or less and since the sample size was too small to create estimates for each household size of 4 or more. Therefore, estimates of Rho* by household size are modeled using Census data. Figure 2 shows Rho* on the y-axis and household size on the x-axis. The upper line is from the US Census public-use microdata sample (PUMS) file for education attainment for ages 25+. The upper line shows that education attainment is more similar among households with two adults, perhaps more likely to be married couples. It shows a drop off when going from two to three adults. We captured the variation in households size by computing the ratio of Rho* for the NAAL prose literacy scores to the Rho for the Census PUMS education attainment among households with $B > 3$ and applying the ratio to the PUMS Rho across all household sizes. The resulting values are the estimates of compact Rho_B^* for $B = 1, 2, \ldots 11$.

Appendix B

Computational algorithm

A computational algorithm was developed to arrive at optimal within-household sample sizes for each household size B. The algorithm was constructed to generate optimal integer or fractional solutions that capture the effects of clustering, differential sampling rates and cost, under the constraints of at least one selected person per household and no more than 2. Here are the steps of the algorithm (all processing runs converged within four iterations):

- Initialize by setting $b = 1$ for all values of B (Take1).
- Compute $\text{DEFF}_{\text{clu}}^{\text{HH*}}$, DEFF_{wgt}, c_p, c_{HH}, and VC(0).
- Do $I = 1$ to 5.
  - Do $B = 1$ to 11.
    - Compute $\text{DEFF}_{\text{clu}}^{\text{HH*}}$, DEFF_{wgt}, c_p, c_{HH}, and VC for all $1 \le b_B \le 2$, given the set of b_B, for all $B' \ne B$.
    - Identify the b_B with the smallest value of VC.
  - End.
  - If $\text{VC}(I) = \text{VC}(I-1)$ then stop.
- End.

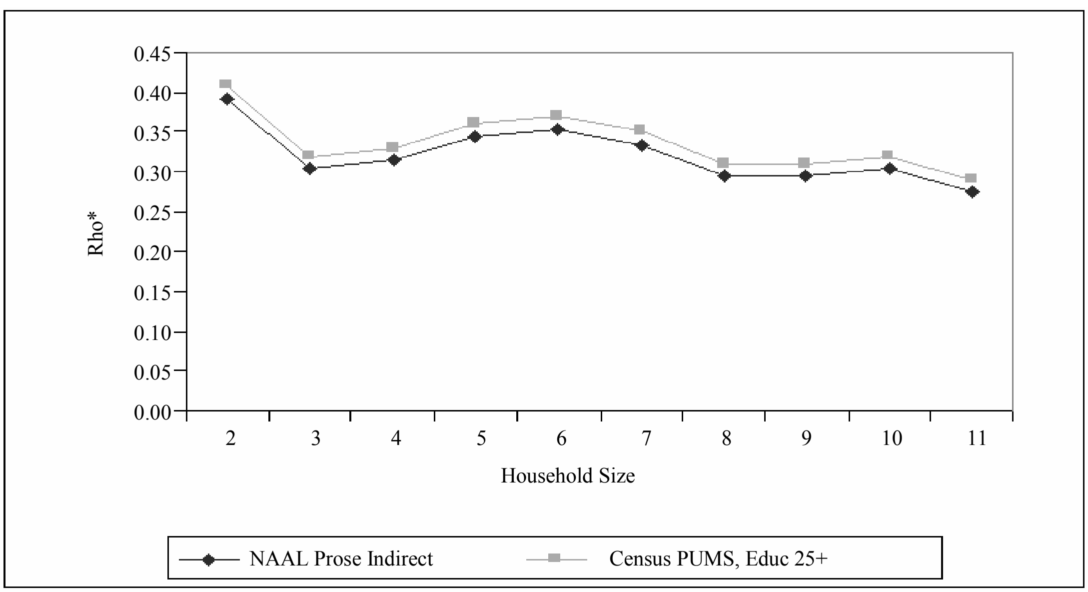

Figure 2 Estimates of Rho* for NAAL by household size

References

Clark, R.G., and Steel, D.G. (2007). Sampling within Households in Household Surveys. *Journal of the Royal Statistical Society*, Series A, 170, 63-82.

Cochran, W.G. (1977). *Sampling Techniques*. 3rd Ed. New York: John Wiley & Sons, Inc.

Kalton, G., Brick, J.M. and Lê, T. (2005). Estimating Components of Design Effects for Use in Sample Design, Household Sample Surveys in Developing and Transition Countries, Chapter VI United Nations, New York, 95-121.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons, Inc.

Mohadjer, L., and Curtin, L.R. (2008). Balancing sample design goals for the National Health and Nutrition Examination Survey. *Survey Methodology*, 34, 1, 119-126.

NCES (2001). Technical Report and Data File User's Manual For the 1992 National Adult Literacy Survey. U.S. Department of Education, National Center for Education Statistics, Washington, DC.

OECD (2005). Learning a Living: First Results of the Adult Literacy and Life Skills Survey. Organisation for Economic Co-operation and Development, Paris. Statistics Canada, Ottawa.