Ricard Giné-Garriga (ricard.gine@upc.edu), School of Civil Engineering, Universitat Politècnica de Catalunya – Barcelona, Campus Nord – Edif C2, C Jordi Girona 1–3, 08034 Barcelona, Spain

[CH]Sampling in surveys with reduced populations: a simplified method for the water, sanitation and hygiene sector

[AU]Agustí Pérez-Foguet and Ricard Giné-Garriga

*[ABS]**Abstract:*** Making decisions efficiently and equitably requires up-to-date and reliable information. In the water and sanitation sector, devolving decision-making to local governments is increasingly promoted to stimulate local development. However, too few data are available at a level of disaggregation that is appropriate to allow decisions to be made about local-level allocations or for monitoring equitable outcomes across communities. Collecting robust data through cost-effective methodologies is therefore a key element of planning and programming, and for this correct sampling methods are of primary importance. Although different sampling strategies are currently being used to support national-level interventions, none have been optimized for data collection at the local level with only small populations, and standard approaches are usually overly expensive and time consuming. Against this background, in this paper we used simplified linear piecewise approximations to calculate the sample size for proportions in terms of given precision, confidence levels and population size. To support the use of this proposed approach by practitioners in the field, easy-to-use pre-calculated tables have been included. For sampling, easy-to-follow practical guidelines for household selection and transect walk planning are also provided. Further, six rural communities in Honduras are presented as an initial case study, with total populations ranging from 11 to 44 HHs. The results illustrate the validity and applicability of this approach for sampling design and sample size determination.

[KEY]**Keywords**: sample size, sample selection, household survey, local planning, developing countries

[A]Introduction

Development strategies typically aim to determine whether or not a population in an area of influence meets certain welfare standards. To this end, decision makers require accurate and up-to-date information to avoid decisions based on false assumptions. In the water, sanitation and hygiene (WaSH) sector, national-level interventions typically perform a baseline to identify the under-performing regions (e.g. districts) and an endline to determine progress and impacts after programme completion. Information is primarily collected at a water point or from households (HHs). A variety of tools and techniques have been developed in recent years to gather this information, such as water-point mapping (WaterAid and ODI, 2005) and household surveys, for instance the UNICEF-supported Multiple Indicator Cluster Survey (United Nations Children's Fund, 2006), Rapid Assessment of Drinking Water Quality (Howard *et al.*, 2012) and Water Safety Plans (Bartram *et al.*, 2009).

Household surveys are, by and large, the most commonly used tools for collecting WaSH data (Joint Monitoring Programme, 2006; Macro International Inc., 1996; United Nations Children's Fund, 2006; Giné-Garriga *et al*., 2013). In practical terms, and usually with only limited resources, one would like to take a representative sample of HHs, i.e. a subset of the population that accurately reflects the members of the entire population. Depending upon how many members are covered by a certain infrastructure, the decision about adequacy can be made; if the infrastructure is deemed inadequate, additional efforts must be planned to improve the level of service delivered. Thus the decision about the size of such samples affects both the precision and the cost of the survey (Bennett *et al.*, 1991; United Nations Children's Fund, 2006).

Whenever covering the overall study area would be practically impossible, such as when conducting national surveys, various sampling methods can be used, including simple random sampling, stratified sampling and cluster sampling (Lwanga and Lemeshow, 1991; Macro International Inc., 1996; United Nations Children's Fund, 2006). However, these methods present significant flaws if they are directly applied at lower administrative scales to support decentralized decision-making. At the local scale, information for performance monitoring or benchmarking comparisons needs to be highly disaggregated, as the number of communities/villages is large (Grosh, 1997). Furthermore, the population size in each administrative subunit might be quite low; for instance, in Latin America, the number of HHs in such a community typically ranges from 20 to 500. With these figures, directly applying the standards and guidelines used for large-scale surveys would require overly large samples, which in practice hinders the implementation of any local (i.e. village-based) surveys. Examples of minor adjustments to the standard proposal of sample size based on a normal distribution for finite populations can be found in the literature (Israel, 2013). These adjustments, however, do not work when variables do not follow a normal distribution, such as for reduced populations or for estimating proportions that are significantly different from $p = 0.5$ (i.e. proportions that are close to the extreme values of 0 and 1). As an alternative to statistically sound proposals, some practical guidelines advocate for a constant minimum value for sample size or a constant percentage of the population. However, this means that if the sample sizes are too low, the survey results in estimates with error bounds (confidence intervals) larger than desired (Chadha, 2006). Thus, simple but statistically sound criteria for selecting reduced sample sizes for small populations are needed.

A further requirement for achieving reliable estimates is to have a valid sampling procedure for the selection of HHs. Mathematically speaking, the ideal procedure would be to have a list of all

HHs in the community, such that a selection can be chosen from the list at random; this could then be easily computed with any standard spreadsheet. If such a list does not exist, it may be created by carrying out a quick census, for example by consulting community leaders, since the total population should be relatively small and thus manageable. If this is not practicable, a complete random exercise cannot be achieved. In order to ensure that the sample is as representative as possible, any method that achieves a random or near-random selection of HHs, preferably spread widely over the community, would be acceptable as long as it is clear and unambiguous and does not give the fieldworker the opportunity to make personal choices that may introduce bias (Bennett *et al.*, 1991). Experience suggests that a full list of HHs is rarely available, and neither creating a census list nor randomly selecting a specific number of HHs – even with a spreadsheet application – is straightforward, and both are overly time-consuming.

Against this background, we present a simplified sampling method for household-based surveys for small populations (e.g. of fewer than 500 HHs), which is the common rural context in some low-income countries. We focus on the estimation of proportions (with true or false answers for questions) and determine sample sizes based on exact confidence intervals. The proposed method is also useful for the inference of multiple or graded answers; however, error bounds should be understood in terms of equivalent binary questions; for example in a four-answer question, the requirements apply separately to each of the four binary questions, to each category and to its complementary answer.

To promote the practical application of this method in the field, we have developed an easy-to-use table (Table 2) to support practitioners when they need to decide the appropriate sample size for a survey, thereby allowing them to strike the right balance between precision and cost. For sampling, we present easy-to-follow practical guidelines for random household selection. We specifically opt for a transect walk, i.e. a systematic walk along a defined path across the community, to ensure as wide a coverage as possible, covering the greatest diversity in terms of water resources and sanitation infrastructure. This walk is usually conducted by the technician or practitioner and community members.

We finally show that, although there are inherent limitations to the accuracy and precision of estimates, the approach adopted produces outputs that are valuable for local decision-making processes.

[A]Methods

For large populations and when a normal distribution of the variable can be assumed (such as in national surveys), the standard representative sample of size n for estimating a proportion p is given by equation (1) (Cochran, 1977):

$$n = \left(\frac{z_{\alpha/2}}{d}\right)^2 p(1-p) \qquad (1)$$

where:

d is the required level of precision (also called sampling error), i.e. the range in which the true value of the population is estimated to be. This range is often expressed in percentage points. A precision rate of ± 5% ($d = \pm\, 0.05$) is typically used in household surveys, based on the argument that lower precision would produce unreliable results, while a higher precision would be too expensive as it would require a very large survey. The value $d = 0.05$ gives a confidence interval of 95%;

α is the confidence (or risk) level. Based on the central limit theorem, the key idea is that when a population is repeatedly sampled, the average value of the attribute obtained by those samples is equal to the true population value, with some samples having a higher value and some a lower value than the true population value. If a 95% confidence level is selected, 95 out of 100 samples will have the true population value within the range of precision specified earlier. This risk is reduced for 99% confidence levels and increased for 90% (or lower) confidence levels;

z is a constant that relates to the normally distributed estimator of the confidence level ($\alpha = 0.05$, $z_{\alpha/2} = 1.96$; $\alpha = 0.1$, $z_{\alpha/2} = 1.64$; $\alpha = 0.2$, $z_{\alpha/2} = 1.28$); and

p is the estimated proportion of HHs giving a particular response for one given question. A p value of 0.5 is chosen to maximize $p(1-p)$, thus providing a conservative estimation of the required sample size (i.e. larger than required).

However, equation (1) cannot be applied if the distribution of estimates differs significantly from that of the normal distribution, which occurs with small populations or if the estimates of proportions are far from 0.5. Alternatively, the size of the sample n may be computed from exact confidence limits for p, i.e. using the Clopper–Pearson interval (Reiczigel, 2003). To obtain accurate confidence limits for small, finite populations, we use the finite population correction that is usually used (Anderson and Burstein, 1967, 1968) but with minor fine-tuning (Burstein, 1975). For practical purposes, a summary table is prepared to support the selection of an optimal sample size on the basis of total population N, the desired precision, and the maximum permissible sampling error. Two different sampling approaches are then tested: a random selection of HHs and a transect walk across the community. In terms of implementation, the advantage of the transect walk is simplicity as compared with the 'difficulty' of undertaking a random selection.

The proposed approaches for determining sample size and selecting HHs were tested in six rural communities in Honduras. For this, a census including all HHs was conducted in each community. Three indicators were assessed: drinking water (water handling and storage practices), sanitation (improved infrastructure) and hygiene (handwashing). The true population proportions varied considerably across the communities and within the indicators: close to 50% (e.g. 24 HHs out of 44) or close to extreme values (e.g. 0%, 2 HHs out of 38; or 2 HHs out of 11). Using reduced sample sizes, two computer-based simulations were performed and compared for all three indicators in each community: i) based on random selection of HHs, and ii) based on simulating a transect walk.

[A]**Results and discussion**

By adopting the previous approach for sample size determination, different sampling plans could be numerically computed in a standard spreadsheet for a variety of α–d value pairs (Table 1). Notably, sampling for a population size N lower than 10 may be to a certain extent meaningless, as significant sacrifices have to be made for either the confidence level (α) or the precision (d). For instance, a sample size n of 7 in a population N of 8 ($n : N$ of 7 : 8) would produce estimates within 33% (±16.6%, $d \leq 1/6$) of the true proportion, with 80% confidence ($\alpha = 0.2$); alternatively, a sample size $n : N$ of 6 : 8 would produce estimates within 25% ($d \leq 1/4$) of the true proportion, with 90% ($\alpha = 0.1$) and 95% ($\alpha = 0.05$) confidence estimates. In contrast, a sample size of 12 from a population of size N less than 100 guarantees estimates with 90% confidence with 25%

precision ($\alpha = 0.1$, $d \leq 1/4$). For higher precision values, the sample size would have to increase significantly, i.e. from 12 to 35 to obtain a precision of ±12.5% ($d \leq 1/8$), and to 46 for a precision of ±10% ($d \leq 1/10$); for a 95% confidence level ($\alpha = 0.05$), the required sample size from N of <100 would be 15 to obtain $d \leq 1/4$, 42 for $d \leq 1/8$, and 53 for $d \leq 1/10$. These figures are comparable to those obtained by applying the standard equation (1). It is worth noting that significant differences exist for large values of α (such as for $\alpha = 0.2$, as shown in Table 1), which indicates the limits of approximation by a normal distribution.

[CAP]**Table 1** Sample size *n* for different values of *N*, α and *d*

N	$\alpha = 0.05$				$\alpha = 0.1$				$\alpha = 0.2$			
	$d \leq 1/10$	$d \leq 1/8$	$d \leq 1/6$	$d \leq 1/4$	$d \leq 1/10$	$d \leq 1/8$	$d \leq 1/6$	$d \leq 1/4$	$d \leq 1/10$	$d \leq 1/8$	$d \leq 1/6$	$d \leq 1/4$
8	8	8	8	6	8	8	8	6	8	8	7	5
10	10	10	9	7	10	10	9	7	10	9	8	6
15	14	14	12	9	14	13	11	8	14	12	10	7
20	18	17	15	10	18	16	13	9	17	15	12	7
25	22	20	17	11	21	19	15	9	19	16	13	8
30	25	23	18	12	24	21	16	10	21	18	13	8
40	31	27	21	13	29	24	18	10	25	20	14	8
50	36	31	23	13	33	27	20	11	28	22	15	8
75	46	37	27	14	40	32	22	12	32	25	16	9
100	53	42	29	15	46	35	23	12	35	26	17	9
150	64	48	31	16	53	39	25	12	39	28	18	9
250	75	54	34	16	60	42	26	12	43	30	18	9
500	87	60	36	17	67	46	27	13	46	31	19	9
Eq. 1	97	62	35	16	68	44	25	11	42	27	15	7

The results from Table 1 can be easily applied in practice in a random selection of HHs. Simple random sampling is, however, unlikely to be the sampling method of choice in an actual field survey (Lwanga and Lemeshow, 1991). In these cases, a transect walk across the community might be an acceptable solution. The figures presented in Table 1 can be further exploited to prepare easy-to-use tables (Table 2) that support practitioners in selecting and implementing the most appropriate sampling plan (X : Y): the total number of *N* HHs is divided into smaller, more manageable sets of Y HHs, and the number of HHs (X) that should be surveyed is determined. Similar to Table 1, Table 2 makes it evident that there is little value in sampling for a population size *N* lower than 10, as almost all HHs should then be targeted for sampling. It can also be easily inferred from Table 2 that a sample size of 24 from a population of size 40 is needed to guarantee estimates with 90% confidence within 12.5 percentage points of precision ($\alpha = 0.1$, $d \leq 1/8$). Specifically, one would need to survey 6 out of every 10 HHs in the community (6 : 10). At a practical level, this sampling plan could be applied by different sampling sequences, such as A-A-A-A-A-A-B-B-B-B (where A represents a surveyed household, and B a non-surveyed household), A-B-A-B-A-A-B-A-B-A or A-A-B-A-B-A-A-B-A-B.

[CAP]**Table 2** Sampling plan for different values of N, α, and d

N	$\alpha = 0.05$			$\alpha = 0.1$			$\alpha = 0.2$		
	$d \leq 1/10$	$d \leq 1/8$	$d \leq 1/6$	$d \leq 1/10$	$d \leq 1/8$	$d \leq 1/6$	$d \leq 1/10$	$d \leq 1/8$	$d \leq 1/6$
1–10	All HHs	All HHs	All HHs	All HHs	All HHs	All HHs	All HHs	All HHs	All HHs
11–20	All HHs	4 : 5	4 : 5	All HHs	4 : 5	4 : 5	All HHs	4 : 5	5 : 7
21–30	7 : 8	4 : 5	5 : 8	5 : 6	4 : 5	5 : 8	4 : 5	4 : 6	4 : 8
31–50	6 : 8	6 : 9	5 : 10	5 : 7	6 : 10	4 : 9	6 : 10	4 : 8	4 : 11
51–70	6 : 9	5 : 9	4 : 10	6 : 10	5 : 10	3 : 9	5 : 10	3 : 8	3 : 11
71–100	4 : 7	5 : 11	3 : 9	5 : 10	4 : 10	3 : 11	4 : 10	3 : 10	2 : 10

The decision about the choice of the sequence has been demonstrated to have an impact on achieved results. Indeed, a sampling plan $n : N$ of 5 : 10 (i.e. sampling 5 HHs sequentially) is preferred over a $n : N$ of 1 : 2 (i.e. sampling every other HH), given that the former promotes the randomness of selection. As a consequence, we opted to consider groups of 5–11 HHs when defining the sampling plans. On the basis of this premise, it should be noted that the enumerator will need to make two decisions in household selection: i) selection of the sampling sequence of HHs (as outlined above), and ii) selection of the first interviewed household. Both decisions may affect the results; any process that reduces the subjective decisions of the interviewer should be therefore promoted. In this study, both the sequence and the first interviewed household were defined randomly.

A comparison of different sampling plans presented in Table 2 is shown in Figure 1. We have not, however, included those sampling plans that would produce statistically meaningless results that would be useless in terms of making informed decisions, i.e. when $\alpha = 0.2$ (which indicates that 20 out of 100 samples will not have the true population value within the specified range of precision) or when $d < 1/4$ (which indicates a precision rate of ± 25%). The straight sloping line reflects a hypothetical census approach in which all HHs are included in the sample ($n = N$). The curved lines reflect different statistical significance levels in which a sampling approach is adopted (each curve line is described by different values of d and α). For instance, a sample size $n = 25$ in a population $N = 30$ (i.e. 5 out of 6 HHs) should be included in a survey to produce estimates within 10% (±5%, $d \leq 1/10$) of the true proportion with 90% confidence (1/10, 0.1 line in the graph). Following this curve along the graph (i.e. retaining the same statistical precision) reveals that a sample size $n = 36$ would be required in a total population $N = 50$ (similar to the sampling plan $n : N$ of 5 : 7 given in Table 2). If the precision of estimates can be lowered (e.g. $\alpha = 0.1$, $d < 1/8$; 1/8, 0.1 line), then a sample size $n = 30$ would suffice to cover same population $N = 50$ ($n : N$ of 6 : 10; Table 2).
[CAP]Figure 1 **Sample size n with respect to the total number of HHs, N, and different values of d and α (the values are indicated, in this order, for all curves). Three different scenarios are represented by horizontal lines: sample n of 24, 36 or 48 HHs.**

Figure 1 also illustrates the reference values that correspond to three different scenarios of effective workload in data collection (with each scenario described by a horizontal line).

In the first scenario, a total sample $n = 24$ HHs is surveyed on a daily basis; for example a team of two enumerators, each one visiting two HHs per hour for six hours. In a population $N = 70$, estimates with 90% confidence ($\alpha = 0.1$) within 16.6 percentage points of precision ($d \leq 1/6$) could be produced in one working day (1/6, 0.1 line in the graph).

The second scenario is based on an overall sample $n = 36$ HHs per day; for example a team of two enumerators, each one visiting three HHs per hour for six hours. The increased productivity would provide a higher level of statistical significance and produce estimates in the same population of size 70 within 25% (±12.5%, $d \leq 1/8$) of the true proportion, with 90% confidence (1/8, 0.1 line).

The last scenario (sample $n = 48$ HHs) could be implemented in one working day by a team of two enumerators, each one visiting four HHs per hour for six hours. This sampling approach would produce estimates with 90% confidence ($\alpha = 0.1$) within 10 percentage points of precision ($d \leq 1/10$) in a population $N = 100$ (1/10, 0.1 line).

These figures are given as examples, as the effective daily workload (i.e. the number of surveyed HHs per field worker) will primarily depend on the type of survey (number of questionnaires and questions included in the household interview) and the distance between surveyed HHs (this might be an issue in widely dispersed communities). (See below for a test case.)

[B]Applying the proposed sampling method

To illustrate the power of this simplified sampling method for household-based surveys, six rural villages with populations ranging from 11 to 42 HHs were surveyed in Honduras. In each community, we visited all HHs and carried out a quick questionnaire (requiring less than 15 minutes per HH on average) related to core sanitation and hygiene indicators. Specifically, three indicators were assessed: Q1) use of handwashing facility with soap and water at home; Q2) access to improved sanitation, as defined by the WHO/UNICEF Joint Monitoring Programme (Joint Monitoring Programme, 2008); and Q3) availability of in-house, clean and safe drinking water storage containers (for clarity purposes, results from this indicator are neither shown nor discussed here).

Table 3 presents the achieved results in the six visited villages. Specifically, we show the total number of HHs (N) and the resulting population proportions for Q1 (p_1) and Q2 (p_2) (based on the actual answers). For each indicator, one ranking is produced for priority setting according to the indicator value: a rank of 1 denotes the highest priority (assigned to the neediest village in relation to sanitation services and hygiene practices), and a rank of 6 denotes lowest priority.

[CAP]Table 3 True proportion values for Q1 and Q2

Sampling		Village A	Village B	Village C	Village D	Village E	Village F
Total number of HHs	N	**44**	**38**	**38**	**42**	**13**	**11**
Q1	y_1	**24**	**29**	**30**	**33**	**5**	**2**
	p_1	0.545	0.763	0.789	0.786	0.385	0.182
	Rank$_1$	3	4	6	5	2	1
Q2	y_2	**32**	**35**	**36**	**36**	**7**	**3**

p_2	0.727	0.921	0.947	0.857	0.538	0.273
Rank$_2$	3	5	6	4	2	1

Notes: y = number of surveyed HHs that verify the selected indicator; p = proportion of HHs that verify the selected indicator (y/N); and colour (red-orange-green) shows the prioritization groups based on the estimated proportions, with red being the highest priority, and green the lowest

Based on these rankings, the six villages can be grouped in three categories. Villages F and E show the lowest level of service for both indicators (shown in red in Table 3, denoting a risky situation). Interestingly, they are the smallest surveyed villages, with 11 and 13 HHs respectively. In contrast, villages B, C, and D have the highest service levels (depicted in green). Note that this large disparity within these six villages justifies obtaining sanitation coverage data for each village rather than focusing on obtaining more in-depth coverage only for a few villages, as the base for a survey. In terms of decision-making, where sanitation coverage is lowest and open defecation is widespread, the coordination of initiatives to support new construction of facilities or the implementation of social sanitation marketing strategies would emerge as potential remedial actions. Similarly, the launch of handwashing campaigns and other hygiene-related initiatives to promote hygiene education might be effective in places were handwashing behaviour is poor.

Previous examples illustrate the need for reliable information at a level of disaggregation that is appropriate to decision-making. Successful poverty-reduction strategies rely on targeting recipient communities and villages based on real hardship; thus up-to-date and reliable data are not only the lifeblood of local-level planning but also provide the raw material for equity-oriented prioritization mechanisms.

[B]Validating the sampling method

In Table 4, we validate the approach adopted in this study for sample size determination and for household selection. We compare the results from the two different sampling approaches, namely random sampling and a transect walk that gives near-random sampling. In both cases, up to 10,000 random repetitions were simulated with RStudio, a free and open-source integrated development environment (IDE) for R (the R code is available at https://doi.org/10.5281/zenodo.1217205).

Basic statistics for each simulation have been computed, including the number of surveyed HHs and the estimated proportion p and its confidence interval. In the random sampling approach, the number of sampled HHs remained the same in all repetitions. However, the number of surveyed HHs varied between two different transect walk simulations, as it depends both on the sampling sequence and the first sampled HH. For this reason, we present the average number of HHs included in the sample, n_{av} (Table 4).

The averages of basic statistics are not shown in Table 4, as they match or are very similar to the expected values for all cases. Rather, other statistical parameters are shown to determine how well each sampling approach performs. In particular, it presents: the largest one-sided confidence interval (see max$|pu' - pl'|/2$); the largest difference between p' (estimated) and p (true proportion – see Table 3) (see max$|p' - p|$); and the number of cases for which the proportion p' does not fall within the expected confidence intervals (see Count_no_CI).

[CAP]Table 4 Estimates for Q1 and Q2 based on the sampling plan of α = 0.1; d = 1/8 or d = 1/6 (depending on the total population *N* of each village) and two sampling approaches, of random versus transect walk.

Sampling		**Village A**	**Village B**	**Village C**	**Village D**	**Village E**	**Village F**
Total number of HHs	***N***	**44**	**38**	**38**	**42**	**13**	**11**
Required precision	***d***	1/8 (0.125)	1/8 (0.125)	1/8 (0.125)	1/8 (0.125)	1/6 (0.167)	1/6 (0.167)
Random approach	***n***	25	24	24	25	10	9
Q1	**max | pu′ – pl′ | / 2**	0.124	0.115	0.112	0.116	0.164	0.136
	Count_no_CI	0.068	0.050	0.031	0.059	0.036	0.021
	max | p′ – p |	0.265	0.195	0.169	0.174	0.185	0.182
Q2	**max | pu′ – pl′ | / 2**	0.124	0.084	0.074	0.105	0.164	0.150
	Count_no_CI	0.043	0.043	0.000	0.063	0.000	0.060
	max | p′ – p|	0.207	0.079	0.053	0.143	0.162	0.162
Transect walk	***n* : *N***	6:10	6:10	6:10	6:10	4:5	4:5
Q1	n_{av}	26.42	22.81	22.81	25.21	10.45	8.87
	max | pu' – pl' | / 2	0.130	0.123	0.127	0.123	0.159	0.173
	Count_no_CI	0.047	0.001	0.051	0.061	0.014	0.059
	max | p' – p|	0.185	0.111	0.169	0.161	0.115	0.182
Q2	n_{av}	26.41	22.82	22.83	25.22	10.46	8.88
	max | pu' – pl' | / 2	0.128	0.096	0.085	0.111	0.164	0.189
	Count_no_CI	0.010	0.029	0.000	0.012	0.000	0.096
	max | p' – p|	0.153	0.079	0.053	0.143	0.138	0.162

Notes: *p* = proportion of HHs that verify the selected indicator (y/N); pl′ = estimated lower confidence limit; pu′ = estimated upper confidence limit; p'_{av} = average *p* value after performing a random/near random selection of HHs up to 1,000 times; pl'_{av} = average of estimated lower confidence limit after performing a random/near random selection of HHs up to 10,000 times; pu'_{av} = average of estimated upper confidence limit after performing a random/near random selection of HHs up to 10,000 times; Count_no_CI = number of cases for which the proportion p′ does not fall within the expected confidence intervals

Statistically, the results behave as expected. Key points to highlight are:

- For both questions, and regardless of the sampling approach, the first parameter (see max│pu′ – pl′│/ 2) is lower than *d* for almost all cases. The few exceptions all refer to the transect walk approach and are not significantly larger than the prescribed *d*; these are produced by particularly low values of real *n*.
- The real proportions *p′* are in the great majority of cases within the confidence interval of the estimate (see the low values of the parameter Count_no_CI), and it happens in all villages. The percentage of cases for which this did not happen is significantly lower than the level of confidence (i.e. 10%; α = 0.1).

- Finally, the maximum difference between the estimated and the real proportions (see max $|p' - p|$) is not considerably higher than *d*, although it doubled in some cases. It should be noted that this is the maximum error obtained considering 10,000 simulated samplings, and that the real proportion is found within the confidence interval in the majority of cases (1 – Count_no_CI).

To summarize, using reduced sample sizes in small populations can produce estimates that are sufficiently precise and accurate to support local decision-making. In addition, no differences are observed in the method employed for household selection between pure random selection and the near-random transect walk selection.

[A]Conclusions

In an era of increasing decentralization of basic services, the need for reliable performance data at the local level is rapidly emerging. In particular, decision-makers are required to identify the neediest areas to allocate resources efficiently, equitably and transparently. Many field-data-collection methodologies with different sampling strategies have been developed recently, but these present significant shortcomings when applied in decentralized contexts with reduced populations. Moreover, as the precision of survey data must be balanced against survey costs, the question of how large the sample should be in order to produce precise confidence intervals for the estimates obtained is of primary importance.

This study aimed at adopting a simplified sampling approach for local surveys. Specifically, it defined the sampling plan required to assess household-related issues in small populations. Thus it helps with selecting the minimum sample size on the basis of available resources, desired precision and the maximum permissible sampling error. From a practitioner's viewpoint, this information should be useful for selecting the sampling plan that best confronts the dilemma between precision and cost. In addition, we show that an easy-to-adopt approach for the selection of HHs, through a transect walk in the community, may also be highly effective. Based on one case study from Honduras, the results indicate that a combination of a reduced sample size with a transect walk for household selection may be valid for producing estimates with sufficient precision to be used in decision-making processes.

[A]Acknowledgements

The authors are grateful to A. Jiménez (Stockholm International Water Institute) and O. Flores for their useful comments and suggestions, and to the MSc students D. Requejo and H. Grau, as well as to Honduran organizations, for their support in undertaking the fieldwork. Financial support from the Catalan Government (Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR), project ref: 2014SGR1545:2014-2016) and the Centre de Cooperació per al Desenvolupament Universitat Politècnica de Catalunya (grant ref: 2015-O034) is gratefully acknowledged.

References

Anderson, T. W. and Burstein, H. (1967) 'Approximating the upper binomial confidence limit', *Journal of the American Statistical Association* 62(319): 857–61 <http://www.jstor.org/stable/2283675> [accessed 30 May 2018]

Anderson, T.W. and Burstein, H. (1968) 'Approximating the lower binomial confidence limit', *Journal of the American Statistical Association* 63(324): 1413–15. Available from http://www.jstor.org/stable/2285894 [accessed 30 May 2018].

Bartram, J. *et al.* (2009) *Water safety plan manual: step-by-step risk management for drinking-water suppliers*, World Health Organization and International Water Association, Geneva.

Bennett, S. *et al.* (1991) 'A simplified general method for cluster-sample surveys of health in developing countries', *World Health Stat Q*. 1991/01/01, 44(3): 98–106. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1949887 [accessed 30 May 2018].

Burstein, H. (1975) 'Finite population correction for binomial confidence limits', *Journal of the American Statistical Association*, 70(349): 67–69 <http://dx.doi.org/10.2307/2285378>.

Chadha, V. K. (2006) 'Sample size determination in health studies', *NTI Bulletin*, 42(3 & 4): 55–62.

Cochran, W. G. (1977) *Sampling Techniques*, 3rd edn, John Wiley and Sons, New York.

Giné-Garriga, R., Jiménez, A. and Pérez-Foguet, A. (2013) 'Water-sanitation-hygiene mapping: An improved approach for data collection at local level', *Science of the Total Environment*, 463–64: 700–11 <http://dx.doi.org/http://dx.doi.org/10.1016/j.scitotenv.2013.06.005>.

Grosh, M. E. (1997) 'The policymaking uses of multitopic household survey data: A primer', *World Bank Research Observer*, 12(2): 137–60.

Howard, G. *et al.* (2012) *Rapid Assessment of Drinking-water Quality: A Handbook for Implementation*. World Health Organization and United Nations Children Fund, Geneva.

Israel, G.D. (2013) *Determining Sample Size*, *Agricultural Education and Communication Department, UF/IFAS Extension*, PEOD6, Gainesville. Available from http://edis.ifas.ufl.edu [accessed 30 May 2018].

Joint Monitoring Programme (2006) *Core questions on drinking-water and sanitation for household surveys*, WHO/UNICEF, Geneva/New York. Available from http://www.wssinfo.org [accessed 30 May 2018].

Joint Monitoring Programme (2008) *Progress on Drinking Water and Sanitation: Special Focus on Sanitation*, *Joint Monitoring Programme for Water Supply and Sanitation*, WHO/UNICEF, Geneva/New York. Available from http://www.wssinfo.org [accessed 30 May 2018].

Lwanga, S. K. and Lemeshow, S. (1991) *Sample Size Determination in Health Studies: A Practical Manual*, World Health Organization, Geneva.

Macro International Inc (1996) *Sampling Manual*. DHS-III, Macro International Inc., Calverton, Maryland, USA. Available from http://www.measuredhs.com [accessed 30 May 2018].

Reiczigel, J. (2003) 'Confidence intervals for the binomial parameter: some new considerations', *Statistics in Medicine*, 22(4): 611–21 <http://dx.doi.org/10.1002/sim.1320>.

United Nations Children's Fund (2006) *Multiple Indicator Cluster Survey Manual 2005*, UNICEF, Division of Policy and Planning, New York.

WaterAid and ODI (2005) *Learning for Advocacy and Good Practice – WaterAid Water Point Mapping*, Overseas Development Institute, London.