RESEARCH ARTICLE

Attenuating the nonresponse bias in hunting bag surveys: The multiphase sampling strategy

Philippe Aubry[1]*, Matthieu Guillemain[2]

1 Cellule d'appui méthodologique, Direction de la Recherche et de l'Expertise, Office National de la Chasse et de la Faune Sauvage, Saint Benoist, BP 20, 78612 Le Perray-en-Yvelines, France, **2** Unité Avifaune Migratrice, Direction de la Recherche et de l'Expertise, Office National de la Chasse et de la Faune Sauvage, La Tour du Valat, Le Sambuc, 13200 Arles, France

* philippe.aubry@oncfs.gouv.fr

Abstract

Reliable hunting bag statistics are a prerequisite for sustainable harvest management based on quantitative modeling. Estimating the total hunting bag for a given game species is faced with a multiplicity of error sources. Of particular concern is the nonresponse error. We consider that the major cause of nonresponse bias is when the reluctance to respond is related to a null harvest, which leads to a potentially important overestimation. For tackling the nonresponse bias issue, we advocate the repeated subsampling of nonrespondents, with a final phase of personal interview by phone, intended to be without nonresponse. When a 100% response rate is actually reached at the last phase, both total and sampling variance can be estimated without bias, whatever the response rates at the previous phases. The actual case of imperfect response at the last phase is studied using Monte Carlo simulations. For imperfect response at the last phase, we show that the estimators we advocate are biased downwards but that these bias remain very moderate if the response rate at the last phase is high enough, depending on the circumstances. Furthermore, we illustrate that increasing the number of phases improves the nonresponse bias attenuation. In case of a hunting bag collecting scheme prone to a high nonresponse rate, for obtaining a very satisfying nonresponse bias attenuation we advocate relying on the multiphase sampling strategy with two- or three-phases, and a response rate in the last phase of at least 90%.

Introduction

Management of harvested wildlife populations increasingly moves towards a science-based approach (but see [1–3]) where the sustainability of the populations and the hunting activity itself are ensured by adequate data collection (e.g. [4] for waterfowl in North America). Adaptive harvest management (see [5]) is increasingly used in this context, and relies on continuous monitoring of the populations and hunting bags as minimum required information [6]. Such management is often based on estimates of the population parameters, such as population size

estimates based on counts of animals, or estimates of the total harvest. Until it is mandatory for hunters to report the total number of animals they hunt, even strong policy enforcement cannot lead to an absolute number of animals taken, and total harvest size is only estimated, often through questionnaires and deliberate will of the hunters to fill those.

In the context of hunting, the number of animals killed by a legal hunter (denoted k) is called a bag. Conditionally to given game species, spatial domain and time period (typically the hunting season), let y_k denote the bag for hunter $k \in U$, where U is the population of active hunters (an active hunter is one who participates in hunting, whether he/she is successful or not). We consider situations in which the parameter of interest is the total hunting bag $t = \sum_{k \in U} y_k = N\bar{y}$, with N the size of U. Knowing the total hunting bag at several geographical scales is needed for wildlife management, according to species biology (migratory or sedentary) and population status (threatened, of no concern, invasive, overabundant). Whatever the geographical scale considered, collecting hunting bag data may be achieved only imperfectly. Several reasons may be responsible for such a situation.

1.1 Response error

Hunting bag reporting may be affected by a response error, that is, a discrepancy between the correct number of animals killed for a given species, and that reported by the hunter. This error may be volunteer or not [7, 8]. The response error can be split into several components according to the origins of the error, for instance the prestige or pride to report higher than real bag [7, 9–11], or the inability to recall the exact value of the bag, leading to omission or digit preference [7, 9, 12–19]. Within a group of game animals, the direction of the response error may differ depending on the species, with overreporting for some of the common species, and underreporting for those that are less common [20]. Another possible component of the response error is the misclassification error, that is, attributing the bag to the wrong species, either because of misidentification [7, 21] or because of a name confusion due to regionalisms (see for instance the example mentioned in [22]). This type of error results in bias whose magnitude and direction depends on the species and the region of the country in which the hunters live (see [23], Table 10, and [24], p. 11). In addition, at moderate spatial scale, the hunter can make a location error by attributing the bag to the wrong spatial domain. Collecting hunting bag data is usually achieved by self-reporting, on paper or online questionnaires. In that case, a response error can simply arise through miss filling the correct line or column in the questionnaire (*reporting error*, or *mechanical error* in the sense of MacDonald & Dillman [10]). In practice, for a set of hunting bag data, these sources of response error are non-mutually exclusive from each other.

1.2 Sampling error

We assume that the total hunting bag is estimated by an estimator $\hat{t}_s$ calculated on the basis of a sample of hunters (denoted s), leading to a sampling error $(\hat{t}_s - t)$. We consider that the sampling error is under the control of wildlife statisticians through the use of a probability sampling design $p(\cdot)$ (see for instance [25], Chapter 1). We note expectation and variance under the sampling design by using the subscript p, that is, $E_p(\cdot)$ and $V_p(\cdot)$. We refer in particular to *simple random sampling without replacement* (SRSWOR). Under SRSWOR, a design-unbiased estimator of t is $\hat{t}_s = N\bar{y}_s$, with $\bar{y}_s = n_s^{-1} \sum_{k \in s} y_k$, and n_s is the size of s. The sampling error may be alternatively expressed in terms of means, omitting the factor N, that is $(\bar{y}_s - \bar{y})$. In what follows, we define the operator SRSWOR(N, n) for an SRSWOR sampling design involving a sample of size n drawn from a population of size N.

1.3 Coverage error

In the framework of probability sampling, for estimating the total hunting bag, a first concern is the coverage of the target population (that is, the active hunters). As expected, a poor coverage makes hunting bag surveys inefficient [26] and imperfect coverage results in surveys prone to biased estimations [27, 28]. It may actually be difficult to adequately cover the active hunter population for a given hunting season, and a given group of game species. Indeed, there may not be a full register of the hunters, especially in the absence of a hunting permit such as in the United Kingdom or Ireland for instance. When a register of hunters exists, generally a sub-register of potentially active hunters can be obtained, because the existence of a hunting permit is linked to a system of hunting licences, for all or part of the hunting season. Sometimes there exist legal provisions targeted at a group of game species upon which to rely to obtain a good coverage of the target population. For instance, in the U.S., the Migratory Bird Hunting and Conservation Stamp Act (in short, Duck Stamp Act) requires each waterfowl hunter 16 years of age or older to possess a valid Federal hunting stamp [4, 24, 29]. In this case, duck stamp purchasers form the population sampled (in practice, indirectly, i.e. through the duck stamp dealers). In the past, this frame have provided a rather good coverage of the population of active waterfowl hunters, with only 1% of stamp purchasers having no intention of hunting (see [24], p. 10, and [30]).

1.4 Nonresponse error

A major concern is the fact that only a subset of the sampled hunters respond to the survey. Such nonresponse leads to the partition $s = r \cup m$, $r \cap m = \emptyset$, where r of size n_r is the subset of respondents and m of size n_m is the subset of nonrespondents (m stands for *missing*). Consequently, under SRSWOR, the total estimator is now $\hat{t}_r = N\bar{y}_r$ with $\bar{y}_r = n_r^{-1}\sum_{k\in r}y_k$. In addition to sampling error, the nonresponse introduces another error:

$$\underbrace{(\bar{y}_r - \bar{y})}_{\text{total error}} = \underbrace{(\bar{y}_s - \bar{y})}_{\text{sampling error}} + \underbrace{(\bar{y}_r - \bar{y}_s)}_{\text{nonresponse error}} \tag{1}$$

Whatever the nonresponse mechanism, the nonresponse bias can be written as:

$$\text{NRBias}(\bar{y}_r) \quad = \text{E}(\bar{y}_r - \bar{y}) \tag{2}$$

$$= \text{E}(\bar{y}_r - \bar{y}_s) \tag{3}$$

$$= \text{E}(n_m/n(\bar{y}_r - \bar{y}_m)) \tag{4}$$

with $\bar{y}_m = n_m^{-1}\sum_{k\in m}y_k$. The expression (4) shows that if the nonresponse rate is not zero, then the bias depends on the difference between the means among respondents and nonrespondents. If the means are very close between respondents and nonrespondents, then the nonresponse bias may be neglected, even in case of a high nonresponse rate.

The estimator based on r may be unbiased only in case of ignorable missingness—for a thorough discussion about terminology, see [31] (pp. 103-106)—i.e. when the data are *missing completely at random* (MCAR) or *missing at random* (MAR) (see [32], p. 133). In the case of hunting bag surveys, given the difficulty to implement a proper sampling frame, generally gathering relevant auxiliary variables related to the hunting bag (or to the response propensity) is almost hopeless. Consequently, ignorable missingness is generally limited to the cases where the values taken by variable y are not related to the fact of being respondent or nonrespondent, i.e. MCAR mechanism (see [33], p. 7 and p. 12, or [34], p. 475). For $k \in s$, let $R_k = 1$ if $k \in r$ and

$R_k = 0$ otherwise. Under MCAR mechanism we have:

$$\Pr[R_k = 1] = \phi \quad \text{for all } k \in s \tag{5}$$

with $0 < \phi < 1$ the response propensity. Thus, as we suppose that the hunters respond independently from each other, the MCAR mechanism is a Bernoulli sampling (see [35], Section 3.2). It follows that, conditionally on n_r, the sample r results from the application of a SRSWOR(n_s, n_r) (see [36], p. 44), or equivalently a SRSWOR(N, n_r) (see [25], Theorem 4.1, p. 69). Thereby, under the MCAR mechanism of missingness, $\bar{y}_r$ is an unbiased estimator of $\bar{y}$.

A self-administered questionnaire can be paper- or web-based. A mail survey is potentially the most useful and inexpensive technique (by respect to interview-based surveys), and does not require access to the web nor computer skills, two things unequally shared by hunters between countries (and also within the same country). Accordingly, in what follows we will refer to mail surveys only.

In hunting bag mail surveys, the causes for nonresponse are partly common to any other type of mail survey: questionnaire never received, lost questionnaire, negligence, lack of time available, lack of interest. All these causes are not necessarily related to the hunting bag, hence several of these may be viewed as MCAR. For instance, nondelivery of the questionnaire is typically treated as an ignorable nonresponse cause [37, 38]. On the other hand, mail questionnaires are answered more often by people who, due to their educational and occupational background, more easily express themselves in writing. Writing facility is roughly correlated with educational level or socioeconomic status [39]. It cannot be taken for granted that this factor for responding is not related to the hunting bag. Another source of potential bias, at least at the regional scale, is related to the auspices, a conscious or unconscious slanting of responses because of attitudes toward the agency or organization sponsoring the survey [40]. This is for example a cause of bias that we can perfectly imagine in the case of France, where the hunters may behave differently towards various stakeholders within the hunting community (e.g. hunting NGOs versus national body). The demographic status of a certain game species may also influence the response. For instance, if the species is declining, some hunters may be afraid of publicizing their hunting bags just because they do not want to give clues to restrict their hunting activities even further. Anyway, there is a widespread nonresponse cause which is specific to hunting bag surveys, namely the tendency for nonrespondents to be less active or less successful hunters than are respondents [10, 13, 37, 38, 41, 42] (see also [43], Figure 6). Being related to the hunting bag, this nonresponse cause alone precludes ignoring the nonresponse as a source of (upward) bias. This is a well-documented and cogent argument that will be put at the heart of the present study.

1.5 Multiphase sampling approach

In this paper we only deal with sampling and nonresponse errors (we do not consider response and coverage errors). Several techniques for handling nonresponse problems in sample surveys are available in the literature. It is out of the scope of the present paper to review them in detail and the reader is referred to [44] (Chapter 8), [35] (Chapter 15), and [31, 45–47]. Basically, we may distinguish between, (i) methods applied at the design stage by ensuring that a subsample of the nonrespondents is followed up—a method pioneered by Hansen & Hurwitz [48]—and (ii), those applied at the estimation stage. These two types of methods can be combined as in [49] or [50]. All techniques in category (ii) use auxiliary information related to the variable of interest, or to the response propensity, in one way or another. For instance, if we had such variables for post-stratifying the sampling frame in strata homogeneous with respect to the hunting bag, or with respect to the propensity to respond, then the nonresponse bias

could be greatly attenuated. Unfortunately, most of the time, relevant auxiliary information—in the sense that we have just specified—are not available in the context of hunting bag surveys. The mailing address, age and sex of the hunters usually available in the sampling frame are not such as allowing nonresponse bias attenuation, because they are not enough to give account for hunting bag or for response propensity. In principle, an auxiliary variable which could be very useful at the estimation stage would be the number of ammunitions fired during the hunting season under consideration. Indeed, such an information would allow identifying the least active hunters in the sample, who more likely had a null harvest. We could use this information for reweighting the respondents whose hunting bag was zero, and thus compensate for the deficit of null harvests among them. In practice, it is very unlikely to be able to gather relevant information about nonrespondents without contacting them. Accordingly, methods relying on auxiliary variables are generally not in use in our context (but see [51] for an example of imputation). Moreover, these methods may need assumptions which are difficult or impossible to verify. Lastly, when the hunting bag survey deals with a great number of game species (for instance, about 90 species in France), it is inconceivable to deal with the problem of nonresponse bias separately for each species. Therefore, we argue that the most practical solution in our context is design-based. Indeed, with a design-based approach, we avoid relying on uncheckable assumptions, and we are not limited in practice by the number of game species.

The first aim of this paper is to gather statistical elements scattered through the literature, and secondly to provide an unbiased estimator for the sampling variance (for any number of phases) which, to our knowledge, is still lacking. Although we consider nonrespondent sub-sampling designs because they are free from assumptions and do not require auxiliary variables, a practical requirement of major importance remains. Actually, the total estimator is unbiased only if response rate at the last phase of the sampling design is 100%. The same holds for the sampling variance estimator. It is obvious that, in practice, the response will never reach 100% at the last phase (it was for instance only 75% in [37]), and theoretically the nonresponse bias issue hence remains [49]. Therefore, after describing the theory related to the sampling strategy that we advocate in this paper, the question still is whether or not the estimators are practically useful when some nonresponse remains at the last phase. In addition, it is necessary to provide some indications about the threshold response rate at the last phase under which the whole sampling strategy becomes useless, according to circumstances. To document this topic of utmost practical importance, we rely on Monte Carlo simulations. For this, we propose a nonresponse mechanism generating upward bias, which rely on the essential source of nonresponse error, namely the propensity of nonrespondents to have, on average, a lower hunting bag than respondents.

Two-phase sampling design

We begin with the simplest case, which corresponds to the pioneering work of Hansen & Hurwitz [48]. Informally, their technique is applied as follows: (i) select a sample of hunters and mail a questionnaire to all of them, (ii) after the deadline has passed, identify the nonrespondents and select a subsample among them, (iii) collect the bags from the nonrespondents in the subsample by personal interview and (iv) combine data from the two sets of respondents for estimating the total hunting bag.

2.1 Design

Let s_1 be the first-phase sample of size n_{s_1} drawn from U by SRSWOR with sampling fraction $v = n_{s_1}/N$. A self-administered questionnaire is mailed to each surveyed person $k \in s_1$. After the deadline to reply, the sample s_1 can be partitioned into a subset of respondents r_1 of size

n_{r_1}, and a subset of nonrespondents m_1 of size $n_{m_1} = n_{s_1} - n_{r_1}$. In the second phase, m_1 is sampled by SRSWOR to obtain a subsample s_2 of $n_{s_2} = n_{m_1} v_m$ $(0 < v_m \leq 1)$ persons interviewed in face-to-face mode or by phone. In this phase, the response rate is assumed to be 100% ($n_{r_2} = n_{s_2}$). The design can thus be summarized by the scheme:

$$U \xrightarrow{\text{SRSWOR}} s_1 \rightarrow r_1$$
$$\searrow$$
$$m_1 \xrightarrow{\text{SRSWOR}} s_2 = r_2$$

(6)

Conditionally to the nonresponse, U may be viewed as poststratified into a strata of respondents R of size N_R, with weight $W_R = N_R/N$, and a strata of nonrespondents M of size N_M, with weight $W_M = N_M/N = 1 - W_R$. Denoting $\bar{y}_R$ and $\bar{y}_M$ the mean in stratum R and M, respectively, the nonresponse bias can also be written as:

$$\text{NRBias}(\bar{y}_r) = \text{E}(N_M/N(\bar{y}_R - \bar{y}_M))$$

(7)

2.2 Mean and total estimators

The mean $\bar{y}$ in the population can be written as the linear combination:

$$\bar{y} = W_R \bar{y}_R + W_M \bar{y}_M$$

(8)

The sample s_1 allows estimating W_R without bias by $w_r = n_{r_1}/n_{s_1}$. Similarly, W_M is estimated without bias by $w_m = n_{m_1}/n_{s_1}$. Unbiased mean and total estimators are, respectively:

$$\bar{y}_{\text{HH}} = w_r \bar{y}_{r_1} + w_m \bar{y}_{r_2}$$
$$= \frac{1}{n_{s_1}} (n_{r_1} \bar{y}_{r_1} + n_{m_1} \bar{y}_{r_2})$$

(9)

$$\hat{t}_{\text{HH}} = N \bar{y}_{\text{HH}}$$
$$= \frac{N}{n_{s_1}} (n_{r_1} \bar{y}_{r_1} + n_{m_1} \bar{y}_{r_2})$$

(10)

with $\bar{y}_{r_1} = n_{r_1}^{-1} \sum_{k \in r_1} y_k$ and $\bar{y}_{r_2} = n_{r_2}^{-1} \sum_{k \in r_2} y_k$. Estimators (9) and (10) are unbiased only when the response rate at the second phase is 100%.

2.3 Sampling variance

The sampling variance of $\bar{y}_{\text{HH}}$ may be written as:

$$\text{V}_p(\bar{y}_{\text{HH}}) = \left(\frac{1}{n_{s_1}} - \frac{1}{N} \right) S^2 + \frac{1}{n_{s_1}} \left(\frac{1}{v_m} - 1 \right) W_M S_M^2$$

(11)

with $S^2 = (N-1)^{-1} \sum_{k \in U} (y_k - \bar{y})^2$ and $S_M^2 = (N_M - 1)^{-1} \sum_{k \in M} (y_k - \bar{y}_M)^2$.

The first term corresponds to the first-phase SRSWOR variance, whereas the second term corresponds to the variance due to subsampling (second phase).

Note that the expression given by Hansen & Hurwitz [48] (Equation 2) involves "1/N" variances and not "1/(N – 1)" variances according to the current convention in the field of finite

sampling theory (see for instance [52], p. 23). A demonstration of the variance expression is given by Hansen & Hurwitz [48] (Appendix), but also in [53] (p. 977) or [54] (pp. 204-205). Note also that the expression printed in [44] (p. 178, Equation 8.6) is erroneous because of the factorisation of the finite population correction.

The adaptation of two-phase sampling in the context of nonresponse leads to a specific instance of two-phase sampling for stratification (see [52], p. 371, or [55]). The theory for this latter design may be found in [56], [52] (pp. 327-335), [57] (pp. 90-92) or [58]. In the second phase, the stratified sample on which the estimation is based is composed of:

1. the totality of r_1, that is subsampling is performed by taking $v_r = 1$ (exhaustive "subsampling"),

2. a subsample s_2 drawn from m_1 by SRSWOR, with $0 < v_m \leq 1$.

2.4 Sampling variance estimator

A sampling variance estimator was not given by Hansen & Hurwitz [48]. To obtain a nonnegative unbiased variance estimator, just start for instance from the formula of the variance estimator given by Rao [56]. After some algebraic simplifications we obtain:

$$
\hat{V}_p(\bar{y}_{HH}) = \frac{(n_{r_1} - 1)(N - n_{s_1})}{N(n_{s_1} - 1)n_{s_1}} S_{r_1}^2 + \frac{N-1}{N}\left(\frac{n_{m_1} - 1}{n_{s_1} - 1} - \frac{n_{r_2} - 1}{N - 1}\right)\frac{w_{m_1}}{n_{r_2}} S_{r_2}^2
$$
$$
+ \frac{N - n_{s_1}}{N(n_{s_1} - 1)}\left[w_{r_1}(\bar{y}_{r_1} - \bar{y}_{HH})^2 + w_{m_1}(\bar{y}_{r_2} - \bar{y}_{HH})^2\right]
$$

(12)

with $S_{r_1}^2 = (n_{r_1} - 1)^{-1}\sum_{k \in r_1}(y_k - \bar{y}_{r_1})^2$ and $S_{r_2}^2 = (n_{r_2} - 1)^{-1}\sum_{k \in r_2}(y_k - \bar{y}_{r_2})^2$.

Expression (12) is algebraically equivalent to those provided in [59] (Equation 11), [60] (p. 304), [61] (p. 332, Equation 13.5), and [62] (Equation 9) or [34] (p. 473). Another expression is given in S1 Appendix, in line with our generalized estimator for any number of phases (see next section). Lohr [63] (p. 338) also provides a simplified expression which assumes the finite population corrections can be neglected.

Multiphase sampling for nonresponse

El-Badry [64] generalized the method of Hansen & Hurwitz [48] to any number ℓ of mailing waves, followed by a last phase $L = \ell + 1$ for personal interview. The latter phase has a supposed response rate of 100%.

In extending the two-phase case, now the population U is stratified into ℓ strata R_i containing N_{R_i} persons who respond to the i-th mailing wave, plus a strata R_L with N_{R_L} persons who not yet responded after ℓ mailing waves but are assumed to respond to an interviewer, in face-to-face mode or by phone.

To each stratum R_i with weight $W_{R_i} = N_{R_i}/N$ is associated a nonrespondent strata M_i. Letting $M_0 = U$ and $R_0 = \emptyset$, the partition of U (for $1 \leq i \leq \ell$) may be written as:

$$
U = \bigcup_{j=0}^{i-1} R_j \cup R_i \cup \underbrace{\bigcup_{j=i+1}^{L} R_j}_{M_i}
$$

(13)

with, in particular, $M_\ell = R_L$. For instance, for $L = 5$ ($\ell = 4$) we get the scheme:

$$
\begin{array}{llll}
U = M_0 & \to & R_1 & \\
& \searrow & & \\
& M_1 & \to & R_2 \\
& & \searrow & \\
& & M_2 & \to & R_3 \\
& & & \searrow & \\
& & & M_3 & \to & R_4 \\
& & & & \searrow & \\
& & & & M_4 & \to & R_5
\end{array}
\tag{14}
$$

With $W_{M_\ell} = W_{R_L}$, the weights W_{M_i} ($0 \le i < \ell$) are defined by the recurrence relation $W_{M_i} = W_{R_{i+1}} + W_{M_{i+1}}$.

3.1 Design

Considering ℓ mailing waves, the design is the following:

- the first mailing wave ($i = 1$) is an SRSWOR from population U,

- if $\ell > 1$, each following mailing wave $1 < i \le \ell$ is addressed to a subsample drawn by SRSWOR from the nonrespondents of the previous mailing wave ($i - 1$),

- the last subsample drawn by SRSWOR ($i = \ell + 1 = L$) concerns the nonrespondents of the wave ℓ to whom we resort to personal interview for ensuring a 100% response rate.

For instance, for $L = 3$ ($\ell = 2$) we have the scheme:

$$
\begin{array}{llll}
U & \xrightarrow{\text{SRSWOR}} & s_1 & \to & r_1 \\
& & \searrow & & \\
& & m_1 & \xrightarrow{\text{SRSWOR}} & s_2 & \to & r_2 \\
& & & & \searrow & & \\
& & & & m_2 & \xrightarrow{\text{SRSWOR}} & s_3 & = & r_3
\end{array}
\tag{15}
$$

Letting $n_{m_0} = N$, $0 < v_i \le 1$, the size of each successive sample s_i is defined as $n_{s_i} = n_{m_{i-1}} v_i$, and therefore $\mathrm{E}_p(n_{s_i}) = v_i \mathrm{E}_p(n_{m_{i-1}})$, for $1 \le i \le L$.

3.2 Mean and total estimators

The population mean can be written as a linear combination of the respondent strata means:

$$\bar{y} = \sum_{i=1}^{L} W_{R_i} \, \bar{y}_{R_i} \tag{16}$$

We have (see for instance [65], p. 122):

$$E_p\left(\frac{n_{r_1}}{n_{s_1}}\right) = E_p\left(\frac{n_{r_1}}{N v_1}\right) = W_{R_1} \tag{17}$$

$$E_p\left(\frac{n_{r_2}}{n_{s_1} v_2}\right) = E_p\left(\frac{n_{r_2}}{N v_1 v_2}\right) = W_{R_2} \tag{18}$$

$$E_p\left(\frac{n_{r_3}}{n_{s_1} v_2 v_3}\right) = E_p\left(\frac{n_{r_3}}{N v_1 v_2 v_3}\right) = W_{R_3} \tag{19}$$

$$\vdots$$

$$E_p\left(n_{r_\ell} \Big/ n_{s_1} \prod_{j=2}^{\ell} v_j \right) = E_p\left(n_{r_\ell} \Big/ N \prod_{j=1}^{\ell} v_j \right) = W_{R_\ell} \tag{20}$$

Letting $\Pi_i = \prod_{j=1}^{i} v_j$ for $1 \leq i \leq \ell$, we obtain the general term:

$$E_p\left(\frac{n_{r_i}}{N\Pi_i}\right) = W_{R_i} \tag{21}$$

Accordingly, we have the unbiased estimators ($1 \leq i \leq \ell$):

$$\widehat{W}_{R_i} = \frac{n_{r_i}}{N\Pi_i} \tag{22}$$

and for $i = L$ we get:

$$E_p\left(\frac{n_{m_\ell}}{N\Pi_\ell}\right) = W_{R_L} \tag{23}$$

which leads to the unbiased estimator:

$$\widehat{W}_{R_L} = \frac{n_{m_\ell}}{N\Pi_\ell} \tag{24}$$

The mean $\bar{y}$ can be estimated without bias using $\bar{y}_{\mathrm{EB}}$ (e.g. [64], Equation 3), which can be written with our notations as:

$$
\begin{aligned}
\bar{y}_{\mathrm{EB}} &= \sum_{i=1}^{L} \widehat{W}_{R_i}\, \bar{y}_{r_i} \\
&= \sum_{i=1}^{\ell} \frac{n_{r_i}}{N\Pi_i}\, \bar{y}_{r_i} + \frac{n_{m_\ell}}{N\Pi_\ell}\, \bar{y}_{r_L} \\
&= \frac{1}{N}\left(\sum_{i=1}^{\ell} \frac{n_{r_i}}{\Pi_i}\, \bar{y}_{r_i} + \frac{n_{m_\ell}}{\Pi_\ell}\, \bar{y}_{r_L} \right)
\end{aligned}
\tag{25}
$$

Of course, the total estimator is:

$$
\begin{aligned}
\hat{t}_{\mathrm{EB}} &= N\, \bar{y}_{\mathrm{EB}} \\
&= \sum_{i=1}^{\ell} \frac{n_{r_i}}{\Pi_i}\, \bar{y}_{r_i} + \frac{n_{m_\ell}}{\Pi_\ell}\, \bar{y}_{r_L}
\end{aligned}
\tag{26}
$$

In practice, due to the rounding necessary to obtain integer sample sizes, in place of the sampling fractions provided by the design, we prefer to write the estimator by explicitly showing the sample sizes used:

$$
\hat{t}_{\mathrm{EB}} = \sum_{i=1}^{\ell} \frac{n_{r_i}}{\prod_{j=1}^{i} \frac{n_{s_j}}{n_{m_{j-1}}}}\, \bar{y}_{r_i} + \frac{n_{m_\ell}}{\prod_{j=1}^{\ell} \frac{n_{s_j}}{n_{m_{j-1}}}}\, \bar{y}_{r_L}
\tag{27}
$$

To ensure, on the average, the sampling fractions provided by the design, it is necessary that the sampling sizes n_{s_i} be rounded by randomizing between $\lfloor n_{s_i} \rfloor$ and $\lceil n_{s_i} \rceil$ with respective probabilities $1 - n_{s_i} + \lfloor n_{s_i} \rfloor$ and $n_{s_i} - \lfloor n_{s_i} \rfloor$ [66]. This point is important in case of Monte Carlo simulation (see Section 4.2.1).

Again, estimators (25) and (26) are unbiased if response rate is really 100% at the last phase. For $L = 2$, the estimator $\bar{y}_{\mathrm{HH}}$ (9) is obtained as a particular instance of the estimator $\bar{y}_{\mathrm{EB}}$ (25). Taking $L = 3$, we obtain:

$$
\bar{y}_{\mathrm{EB}} = \frac{1}{n_{s_1}} \left(n_{r_1} \bar{y}_{r_1} + \frac{n_{r_2}}{v_2} \bar{y}_{r_2} + \frac{n_{m_2}}{v_2} \bar{y}_{r_3} \right)
\tag{28}
$$

in accordance with the expression given by Siripornpibul [67] (p. 66, Equation 3.1), but with a different notation.

3.3 Sampling variance

The sampling variance for $\bar{y}_{\mathrm{EB}}$ was given by El-Badry [64] (Equation 4) (see also [68], pp. 407-409). As Rao [62] (p. 105, Equation 36), we prefer the variance expression given by Srinath

[69] (Equation 2.16), that is, with our notations:

$$V_p(\bar{y}_{\text{EB}}) = \left(\frac{1}{n_{s_1}} - \frac{1}{N}\right)S^2 + \frac{1}{n_{s_1}}\left(\frac{1}{v_2} - 1\right)W_{M_1}S^2_{M_1}$$

$$+\frac{1}{n_{s_1}}\sum_{i=2}^{\ell}\left(\prod_{j=2}^{i}\frac{1}{v_j}\right)\left(\frac{1}{v_{i+1}} - 1\right)W_{M_i}S^2_{M_i} \tag{29}$$

with $S^2_{M_i} = (N_{M_i} - 1)^{-1}\sum_{k \in M_i}(y_k - \bar{y}_{M_i})^2$. For $L = 2$, the third term in (29) is not defined and we obtain the variance (11) as a special case. For $L = 3$, we obtain:

$$V_p(\bar{y}_{\text{EB}}) = \left(\frac{1}{n_{s_1}} - \frac{1}{N}\right)S^2 + \frac{1}{n_{s_1}}\left(\frac{1}{v_2} - 1\right)W_{M_1}S^2_{M_1} \tag{30}$$

$$+\frac{1}{n_{s_1}}\frac{1}{v_2}\left(\frac{1}{v_3} - 1\right)W_{M_2}S^2_{M_2} \tag{31}$$

in accordance with the expression given by Siripornpibul [67] (p. 66, Equation 3.2), but with a different notation.

Letting:

$$\Pi_i = \begin{cases} 1 & \text{if } i = 1 \\ \displaystyle\prod_{j=1}^{i-1}\frac{1}{v_j} & \text{if } i > 1 \end{cases}$$

the variance (29) can be rewritten in a more compact way as:

$$V_p(\bar{y}_{\text{EB}}) = \frac{1}{N}\sum_{i=1}^{L}\Pi_i\left(\frac{1}{v_i} - 1\right)W_{M_{i-1}}S^2_{M_{i-1}} \tag{32}$$

and likewise the variance for the total estimator can be written as:

$$V_p(\hat{t}_{\text{EB}}) = N^2V_p(\bar{y}_{\text{EB}})$$

$$= \sum_{i=1}^{L}\Pi_i\left(\frac{1}{v_i} - 1\right)N_{M_{i-1}}S^2_{M_{i-1}} \tag{33}$$

3.4 Sampling variance estimator

Again, a sampling variance estimator was not given by El-Badry [64]. After generalizing the sampling variance estimator for multiphase sampling for stratification to any number of phases (for two-, three-, and four-phase sampling for stratification, see [57], pp. 81-118, and [58]), and after some algebraic simplifications, we obtain the general expression:

$$\hat{V}_p(\hat{t}_{\text{EB}}) = \sum_{i=1}^{\ell}\Pi_i\left[\frac{n_{m_{i-1}}(n_{m_{i-1}} - n_{s_i})}{n_{s_i}(n_{s_i} - 1)}\right]\left(z_i - \frac{1}{n_{s_i}}t_i^2\right) +$$

$$\Pi_L\left[n_{m_\ell}^2\left(\frac{1}{n_{s_L}} - \frac{1}{n_{m_\ell}}\right)S^2_{r_L}\right] \tag{34}$$

with:

$$\Pi_i = \begin{cases} 1 & \text{if } i = 1 \\ \prod_{j=1}^{i-1} \dfrac{n_{m_{j-1}}(n_{m_{j-1}} - 1)}{n_{s_j}(n_{s_j} - 1)} & \text{if } i > 1 \end{cases}$$

and for $1 \leq i \leq \ell$:

$$t_L = \sum_{k \in r_L} y_k \tag{35}$$

$$t_i = \sum_{k \in r_i} y_k + \frac{n_{m_i}}{n_{s_{i+1}}} t_{i+1} \tag{36}$$

$$z_L = \sum_{k \in r_L} y_k^2 \tag{37}$$

$$z_i = \sum_{k \in r_i} y_k^2 + \frac{n_{m_i}}{n_{s_{i+1}}} z_{i+1} \tag{38}$$

For $L = 2$ and $L = 3$ we obtain as particular instances the sampling variance estimators given in S1 Appendix.

Simulating the nonresponse bias

Although the nonresponse bias elimination strategy we presented (through multiphase sampling) is not restricted to hunting bag surveys, the nonresponse mechanism we propose in this section is very specific to the matter at hand.

4.1 Nonresponse mechanism

We separate ignorable causes of nonresponse from nonignorable ones (i.e. related to values taken by y). For the sake of simplicity, among the nonrespondents, we consider the propensity to not respond when the hunting bag is zero (nonactive hunter or unsuccessful hunter) as the only cause of nonignorable nonresponse.

Within U we distinguish the stratum $U_0 \ni k$ such as $y_k = 0$, from the stratum $U_1 \ni k$ such as $y_k > 0$. A sample s of size n_s is drawn by SRSWOR from U. We define $s_0 = s \cap U_0$ of size n_0 and $s_1 = s \cap U_1$ of size n_1. The set-size n_0 is an outcome of a random variable because of the replication of the random draw by SRSWOR. This size follows a hypergeometric distribution whose probability mass function (pmf) is:

$$P_H(n_0|n_s, N_0, N) = \begin{cases} \dbinom{N_0}{n_0}\dbinom{N - N_0}{n_s - n_0}\dbinom{N}{n_s}^{-1} & \text{if } n_0 \in \mathcal{D}_H(n_s, N_0, N) \\ 0 & \text{otherwise} \end{cases} \tag{39}$$

with $N \in \mathbb{N}^*$, $n_s \in \mathbb{N}^*$, $N_0 \in \mathbb{N}^*$ and domain $\mathcal{D}_H(n_s, N_0, N) = (\max(0, n_s + N_0 - N), \cdots, \min(n_s, N_0))$ [70] (p. 251). The mean and variance of n_0 are, respectively:

$$E_H(n_0) = n_s \frac{N_0}{N} = n_s W_0 \tag{40}$$

$$
\begin{aligned}
V_H(n_0) \quad &= n_s \frac{N_0}{N}\left(1 - \frac{N_0}{N}\right)\frac{N - n_s}{N - 1} \\
&= n_s W_0(1 - W_0)\frac{N - n_s}{N - 1}
\end{aligned}
\tag{41}
$$

Conditionally to the sample s (and thus to n_0 and n_1), the nonresponse can be viewed as a second sampling phase. Let $0 \leq \pi_m < 1$ the propensity to nonrespond, all causes of nonresponse confounded, and let $0 \leq \pi_z \leq 1$ the propensity, among the nonrespondents, to nonrespond because their hunting bag was zero. Let z be the size of n_z, the set of nonrespondents who nonrespond because their harvest is null, with $n_z \leq n_m$ and $n_z \leq n_0$. We have:

$$
E(n_m) = n_s \pi_m
\tag{42}
$$

$$
E(n_z) = n_s \pi_z \pi_m
\tag{43}
$$

The nonresponse bias can be written as:

$$
\mathrm{NRBias}\,(\bar{y}_r) = E(n_z/n_s\,\bar{y}_r)
\tag{44}
$$

With n_z independent from $\bar{y}_r$, the nonresponse bias can also be written as:

$$
\mathrm{NRBias}\,(\bar{y}_r) = \pi_z \pi_m E(\bar{y}_r)
\tag{45}
$$

If $\pi_z = 0$ then $n_z = 0$ ($\forall n_m$) and $\mathrm{NRBias}\,(\bar{y}_r) = 0$ (the nonresponse is ignorable). Under the constraint $n_z \leq n_0$, if $\pi_z = 1$ then $n_z = n_m$ and $\mathrm{NRBias}\,(\bar{y}_r)$ is maximal for a given π_m.

4.2 Simulating the nonresponse mechanism

We now describe the way we implement the nonresponse mechanism specific to hunting bag surveys. For uni-phase SRSWOR the sampled population is of course U. For the multiphase sampling strategy, the algorithm we propose is successively applied to m_j for $j = 0, 1, \ldots, \ell$ (with $m_0 = U$) for generating $s_1, s_2, \ldots, s_L$. For the sake of notation simplicity, in what follows we describe the algorithm when sampling U.

4.2.1 Randomizing a set-size. In a Monte Carlo simulation of sampling, all set-sizes are necessarily integers. However, their expectations are not necessarily integers but must be approximately respected during the simulation. To randomly generate a set-size n such as $E(n) = N\pi = \alpha$, with $N \in \mathbb{N}^*$, $0 \leq \pi \leq 1$, and $n \in \mathbb{N}$, we used the two-point distribution:

$$
P_T(n|N, \pi) = \begin{cases} \omega & \text{if } n = \lfloor \alpha \rfloor + 1 \\ 1 - \omega & \text{if } n = \lfloor \alpha \rfloor \\ 0 & \text{otherwise} \end{cases}
\tag{46}
$$

with $\omega = \alpha - \lfloor \alpha \rfloor$, or equivalently:

$$
P_T(n|N, \pi) = \begin{cases} \omega^x(1 - \omega)^{1-x} & \text{if } n = \lfloor \alpha \rfloor + x, \; x \in \{0, 1\} \\ 0 & \text{otherwise} \end{cases}
\tag{47}
$$

of mean $E_T(n) = N\pi$ and variance $V_T(n) = \omega(1 - \omega)$.

4.2.2 Simulation algorithm. In the context of this article, the scheme we used to simulate the nonresponse mechanism consists of randomly defining a set of respondents $R \subset U$ of size N_R and a set of nonrespondents $M \subset U$ of size N_M, with $U = R \cup M$ and $R \cap M = \emptyset$. Within M

we define the subset $Z \subset U_0$ of size N_Z of hunters nonresponding because their hunting bag was zero. If $N_Z > 0$, then the null hunting bags are overrepresented within M, and there exists an upward nonresponse bias. The algorithm is the following:

```
1. randomly generate N_M such as E(N_M) = Nπ_m
2. N_R ← N - N_M
3. randomly generate N_Z such as E(N_Z) = Nπ_zπ_m
4. a sample Z is drawn from U_0 by SRSWOR(N_0, N_z)
5. C ← U - Z, N_C ← N - N_Z
6. a sample R is drawn from C by SRSWOR(N_C, N_R)
7. M ← U - R, (M ⊃ Z)
8. a sample s is drawn from U by SRSWOR(N, n_s)
9. m ← s ∩ M
10. z ← s ∩ Z
11. r ← s ∩ R
```

With N_Z independent from $\bar{y}_R$, the nonresponse bias can also be written as:

$$\text{NRBias}(\bar{y}_r) = \pi_z \pi_m \text{E}(\bar{y}_R) \tag{48}$$

Under this algorithm, the distributional properties of n_m and n_z are given in S1 Appendix. At step 3 of the algorithm, it is required that two constraints are satisfied, namely $N_Z \leq N_M$ and $N_Z \leq N_0$. These constraints are also examined in S1 Appendix. For the reader convenience, Fig 1 illustrate the algorithm.
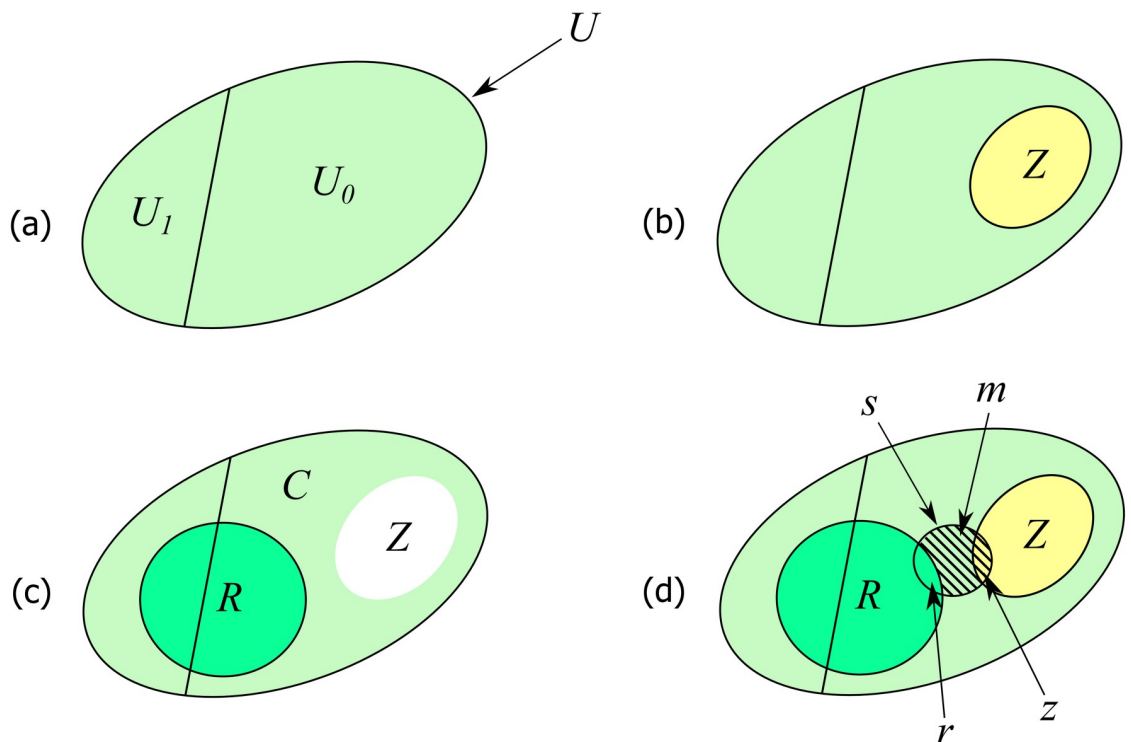
Fig 1. Scheme of the algorithm steps for implementing the nonresponse mechanism. (a) partition of U into strata $U_0 \ni k$ such as $y_k = 0$ and $U_1 \ni k$ such as $y_k > 0$; (b) step 4: a random subset Z is defined within U_0; (c) steps 5-6: the "hole" in U corresponds to Z, resulting in an undercoverage of the stratum U_0 when selecting the random set R within C: no elements which belong to Z can be included in R; (d) steps 7-11: random selection of the sample s within U: the random sets r, m (hatched area within s) and z result from the intersection of s with R, M, and Z, respectively (with z included within m since Z is included within M).

https://doi.org/10.1371/journal.pone.0213670.g001

Monte Carlo simulation study

According to the nonresponse mechanism we proposed in section 4.1 and the algorithm described in section 4.2.2, it is possible to vary the values of π_m and π_z at each phase. Besides, we do not want only theoretical results, but orders of magnitude rooted in reality. Consequently, since the multiphase sampling strategy is complex, and given the possibility to vary the nonresponse at each phase and the requirement of concrete results, we rely on Monte Carlo simulations to documentate the bias of the estimators. For ensuring the quality of our Monte Carlo simulations, we used several random number streams with huge period and very good properties by using function `MRG32k3a` proposed by L'Ecuyer [71] (Fig 1).

5.1 Superpopulation model

To simulate a set of individual hunting bags we need a superpopulation model ξ which should be a discrete distribution allowing to use any proportion of null values. To specify ξ for simulation purpose, a convenient choice is a two-parameter distribution such as the hurdle-at-zero Poisson model:

$$
\Pr\left(Y = y\right) = \begin{cases} p & \text{if } y = 0 \\ \phi \dfrac{e^{-\lambda}\lambda^y}{y!} & \text{for } y = 1, 2, \ldots \end{cases}
\tag{49}
$$

with $0 \leq p \leq 1$, $\phi = (1 - p)/(1 - e^{-\lambda})$ and $\phi \leq (1 - e^{-\lambda})^{-1}$ [70] (p. 352). This distribution is over- or underdispersed by respect to the Poisson distribution depending on the value of $\phi \neq 1$. If $\phi = 1$ then we have $p = e^{-\lambda}$ and we obtain the Poisson distribution as a particular instance. Mean and variance are [70] (p. 352):

$$
\mu = \phi\lambda
\tag{50}
$$

$$
\mu_2 = \phi\lambda(1 + \lambda) - \phi^2\lambda^2
\tag{51}
$$

Knowing μ, we can obtain the value of parameter λ as a solution of the transcendental equation $\mu = \lambda(1 - p)/(1 - e^{-\lambda})$, that is:

$$
\lambda = W_0(De^D) - D
\tag{52}
$$

with $D = \mu/(p - 1)$ and $W_0(x) \geq -1$ the upper branch of the Lambert function.

5.2 Nonresponse bias

When estimating a parameter ω using an estimator $\hat{\omega}$, we define the bias index $r = \mathrm{E}(\hat{\omega})/\omega$. The bias $\mathrm{E}(\hat{\omega}) - \omega$ is positive for $r > 1$, null for $r = 1$, and negative for $r < 1$. Here we plot the bias index $r = \mathrm{E}(\bar{y}_r)/\bar{y}$.

We simulate one finite population U of size $N = 10\,000$ by sampling the superpopulation model ξ with parameters $p = 0.955$ and $\lambda = 7$, that is for a superpopulation mean $\mu \simeq 0.315$. For the population simulated we have $N_0 = 9534$. We replicate 100 000 times the algorithm simulating the nonresponse mechanism with a sampling fraction $v = n/N = 0.5$, for $\pi_m = 0$ (0.05)0.9 and $\pi_z = 0(0.05)0.9$.

In accordance with the nonresponse bias expression (45), Fig 2 shows that π_z and π_m play a symmetrical role in the magnitude of the nonresponse bias. Recall that the sampling fraction v plays no role here. For instance, with $v = 0.05$ we would get exactly the same figure (we should just increase the number of simulations to get such smooth contour levels as depicted here).
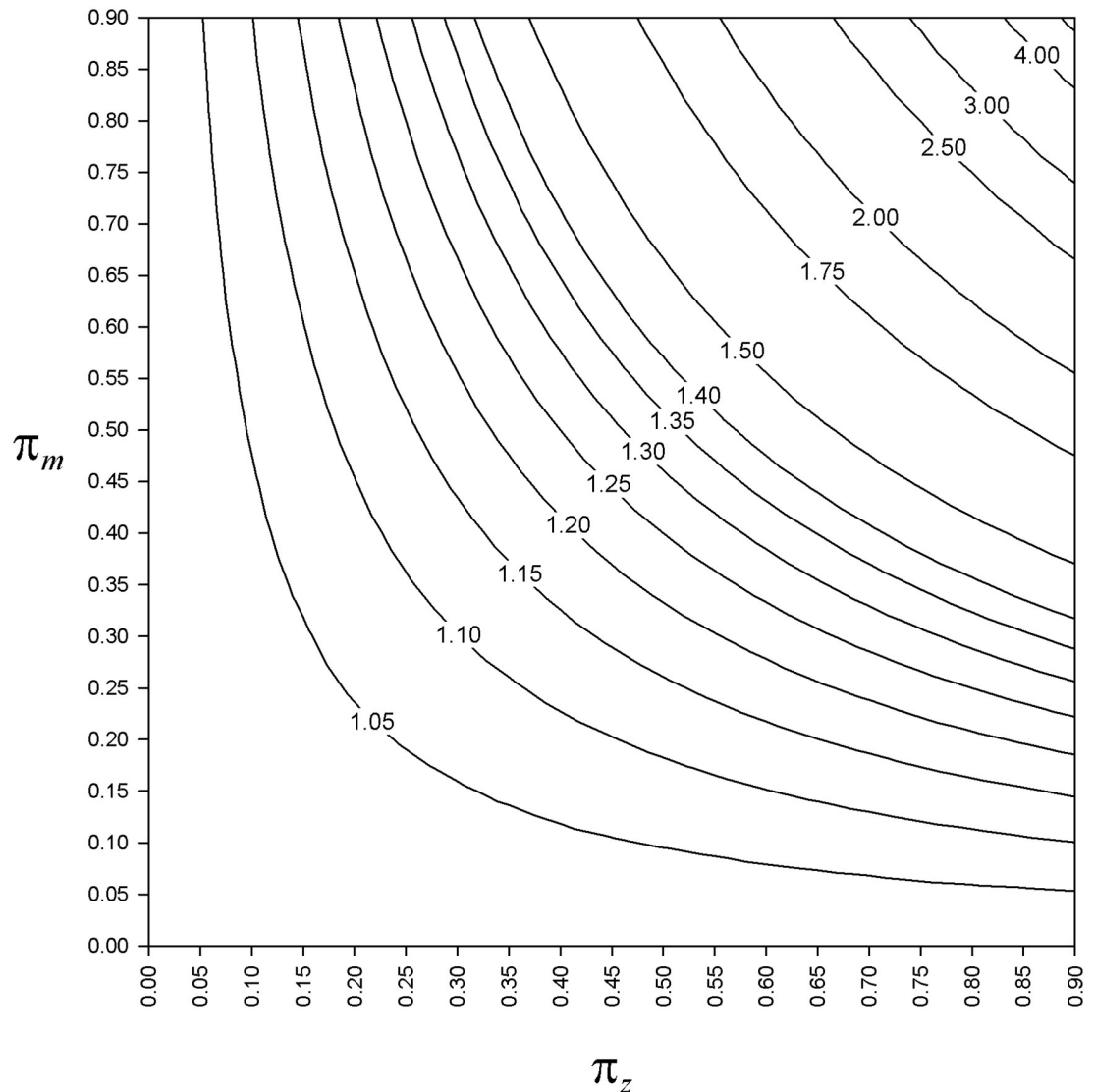
Fig 2. For the simulated population, bias index of the sample mean as a function of π_m and π_z. Contour levels for the bias index $r = \mathrm{E}(\bar{y}_r)/\bar{y}$ based on 100 000 simulations of the nonresponse mechanism, for $\pi_m = 0(0.05)0.9$ and $\pi_z = 0(0.05)0.9$. The contour level for $r = 1$ is confounded with the axes ($\pi_z = 0$ and $\pi_m = 0$). Details in the text.

For the finite population simulated, for $\pi_m = 0.85$ and $\pi_z = 0.30$, the classical estimator $\bar{y}_r$ leads to an overestimation of about 34%. Again for $\pi_m = 0.85$ but with $\pi_z = 0.20$, the overestimation is about 20%.

5.3 Nonresponse bias attenuation under two-phase design

Let $\pi_m(1)$ and $\pi_m(2)$ the values of π_m at the first and second phase of the two-phase sampling design (section 2), respectively. Using the estimator $\bar{y}_{HH}$ (9), whatever the value of $\pi_m(1) < 1$, for $\pi_m(2) = 0$ the nonresponse bias is eliminated. When $\pi_m(2) > 0$, then the nonresponse bias is only attenuated. For the same finite population as in section 5.2, under a two-phase sampling design with sampling fractions $\nu = \nu_m = 0.5$, we replicate the algorithm simulating the nonresponse mechanism for $\pi_z = 0.2$ (π_z remains constant across the phases), $\pi_m(1) = 0.85$, and
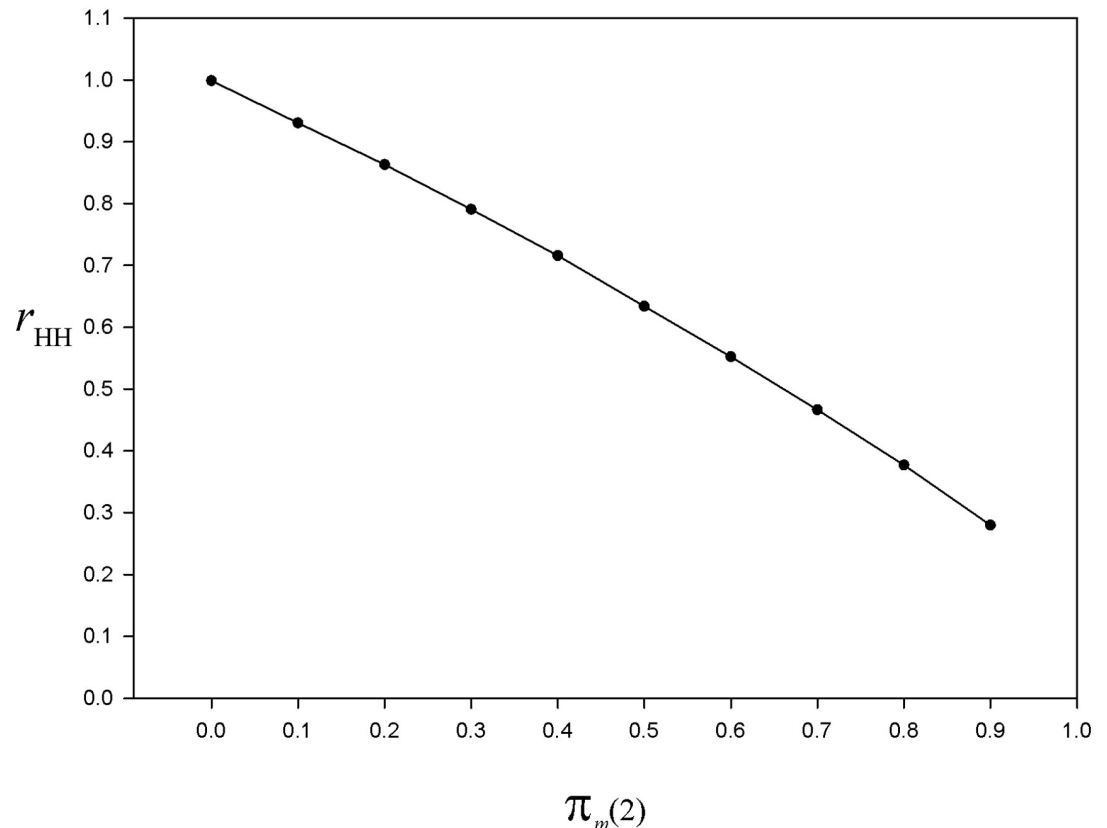
Fig 3. For the simulated population, bias index of the Hansen & Hurwitz estimator as a function of the nonresponse rate at the last phase. Curve of the bias index $r_{HH} = E(\bar{y}_{HH})/\bar{y}$ based on 1000 simulations of the nonresponse mechanism, for $\pi_z = 0.2$, $\pi_m(1) = 0.85$, and $\pi_m(2) = 0(0.1)0.9$. Details in the text.

$\pi_m(2) = 0(0.1)0.9$. First we run 1000 simulations and plot the bias index $r_{HH} = E(\bar{y}_{HH})/\bar{y}$. As expected, for $\pi_m(2) = 0$ we have $r_{HH} = 1$, that is, the nonresponse bias is eliminated (Fig 3). For $\pi_m(2) > 0$ we obtain $r < 1$, which means that using the estimator $\bar{y}_{HH}$ leads this time to an underestimation. This underestimation exceeds 70% for $\pi_m(2) = 0.9$ (Fig 3).

Again, for the two-phase sampling design, the nonresponse bias is not affected by the sampling fractions used. If we set $v_m = v$, with $v = 0.1(0.1)0.5$, we obtain the same results, except that there are Monte Carlo fluctuations (see Table 1).

Besides the behavior of the estimator $\bar{y}_{HH}$, we are also interested in that of the sampling variance estimator, that is $\hat{V}_p(\bar{y}_{HH})$ (12). Thus, in a second time, we run 1 000 000 simulations and plot the bias index $r_V = E(\hat{V}_p(\bar{y}_{HH}))/V_{MC}(\bar{y}_{HH})$ where $V_{MC}(\bar{y}_{HH})$ is a Monte Carlo approximate of $V_p(\bar{y}_{HH})$. As expected, we have $r_V = 1$ for $\pi_m(2) = 0$ (the sampling variance estimator is unbiased when nonresponse rate at the second phase is null) (Fig 4). When $\pi_m(2) > 0$, then we obtain $r_V < 1$, that is, using the estimator $\hat{V}_p(\bar{y}_{HH})$ leads to an underestimation of the sampling variance.

## 5.4 Nonresponse bias attenuation under multiphase design

In this section, we first examine the effect of the number of phases L (section 3) on the nonresponse bias attenuation. Let $\pi_m(i)$ denote the values of π_m at the i-th phase ($1 \leq i \leq L$). For the same finite population as in section 5.2, under a L-phase sampling design with $L = 2(1)6$, and

Table 1. For the simulated population, bias index of the Hansen & Hurwitz estimator as a function of the nonresponse rate at the last phase, with increasing sampling fractions.

| $\pi_m(2)$	$v = 0.1$	$v = 0.2$	$v = 0.3$	$v = 0.4$	$v = 0.5$
0.0	1.006	1.001	1.010	0.996	0.999
0.1	0.944	0.937	0.936	0.931	0.930
0.2	0.862	0.861	0.863	0.863	0.863
0.3	0.794	0.780	0.790	0.798	0.791
0.4	0.692	0.707	0.720	0.716	0.716
0.5	0.633	0.626	0.632	0.636	0.634
0.6	0.545	0.556	0.555	0.553	0.552
0.7	0.464	0.469	0.470	0.469	0.466
0.8	0.378	0.375	0.378	0.375	0.377
0.9	0.288	0.279	0.285	0.279	0.280

Bias index $r_{HH} = E(\bar{y}_{HH})/\bar{y}$ based on 1000 simulations of the nonresponse mechanism, for $\pi_z = 0.2$, $\pi_m(1) = 0.85$, and $\pi_m(2) = 0(0.1)0.9$. The sampling fraction varies as $v = 0.1(0.1)0.5$, with $v_m = v$. Details in the text.
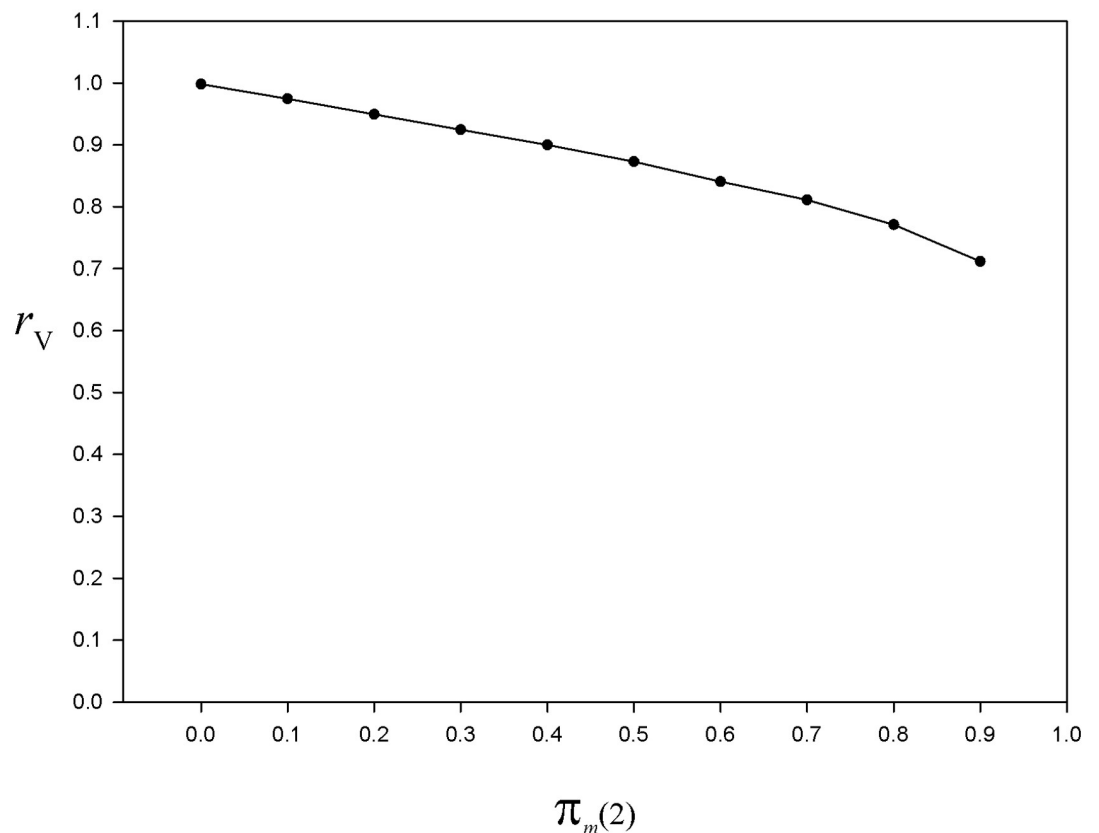
Fig 4. For the simulated population, bias index of the sampling variance estimator as a function of the nonresponse rate at the last phase. Curve of the bias index $r_V = E(\hat{V}_p(\bar{y}_{HH}))/V_{MC}(\bar{y}_{HH})$ based on 1 000 000 simulations of the nonresponse mechanism, for $\pi_z = 0.2$, $\pi_m(1) = 0.85$, and $\pi_m(2) = 0(0.1)0.9$. Details in the text.
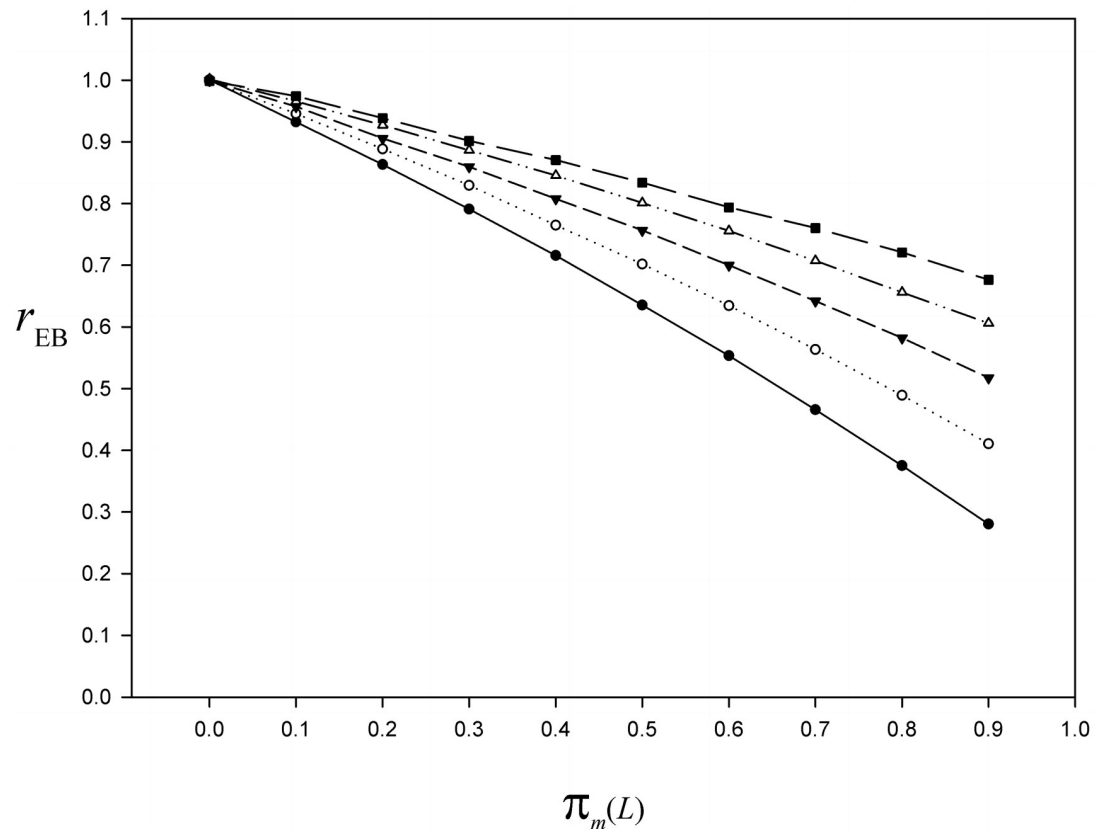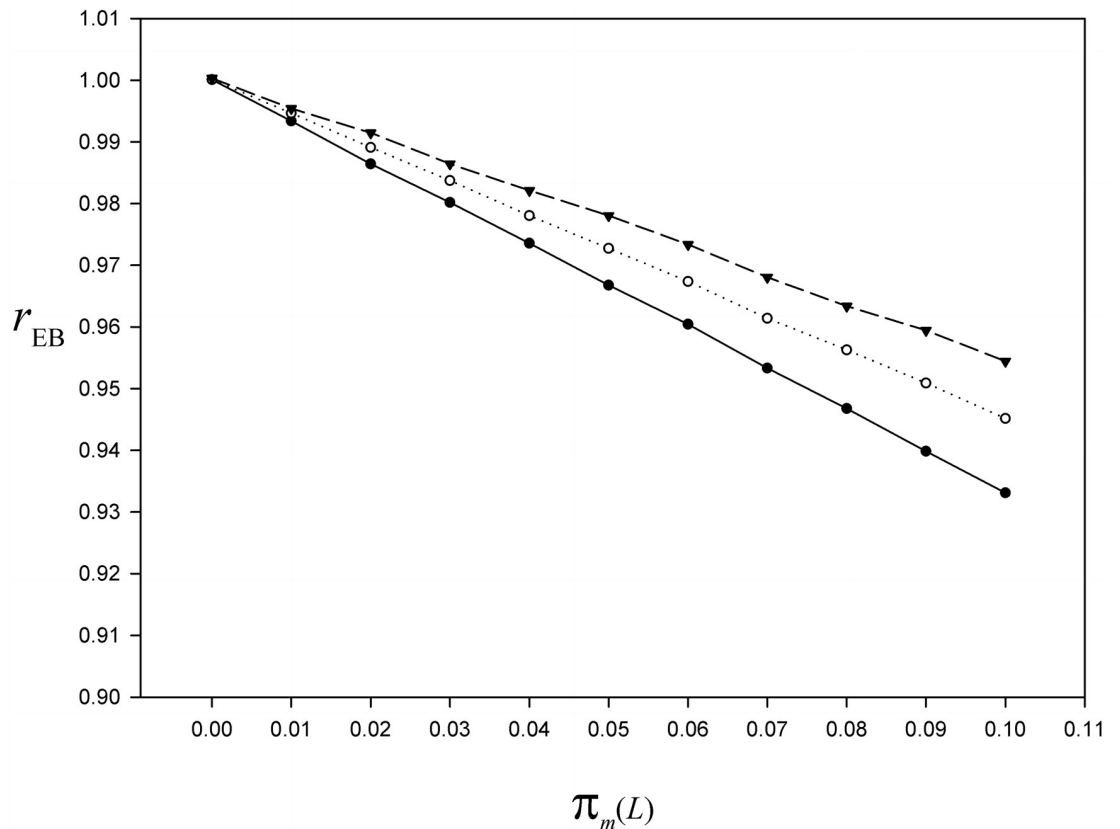
Fig 5. For the simulated population, bias index of the El-Badry estimator as a function of the nonresponse rate at the last phase. Curve of the bias index $r_{EB} = \mathrm{E}(\bar{y}_{EB})/\bar{y}$ based on 10 000 simulations of the nonresponse mechanism, for $L = 2(1)6$, $\pi_z = 0.2$, $\pi_m(i) = 0.85$ for $1 \le i < L$, and $\pi_m(L) = 0(0.1)0.9$ (*black dots L = 2*; *white circles L = 3*; *black triangles down L = 4*; *white triangles up L = 5*; *black squares L = 6*). Details in the text.

sampling fractions $\nu_i = 0.5$ ($1 \le i \le L$), we replicate the algorithm simulating the nonresponse mechanism for $\pi_z = 0.2$ (π_z remains constant across the phases), $\pi_m(i) = 0.85$ for $1 \le i < L$, and $\pi_m(L) = 0(0.1)0.9$. We run 10 000 simulations and plot the bias index $r_{EB} = \mathrm{E}(\bar{y}_{EB})/\bar{y}$. As expected, whatever the number of phases L, for $\pi_m(L) = 0$ we have $r_{EB} = 1$, that is, the nonresponse bias is eliminated (Fig 5). As previously, for $\pi_m(L) > 0$ we obtain $r_{EB} < 1$, which means that the estimator $\bar{y}_{EB}$ is biased downwards. The underestimation decreases as the number of phases L increases (Fig 5).

We now examine the underestimation for moderate values of $\pi_m(L)$, for instance up to 0.1. See Fig 6 for the plot of r_{EB} after 100 000 simulations, for $L = 2, 3, 4$.

Second, taking for example the case $L = 3$, we illustrate the variation of the underestimation according to the value of π_z. We replicate the algorithm simulating the nonresponse mechanism as previously except that, although π_z continues to remain constant across the phases, now it varies as $\pi_z = 0(0.1)0.9$. For each value of π_z, we run 10 000 simulations and plot the bias index $r_{EB} = \mathrm{E}(\bar{y}_{EB})/\bar{y}$. The underestimation is maximum for $\pi_z = 0$ and decreases as π_z increases since it is offset by the potential upward nonresponse bias (which increases in magnitude with π_z) (Fig 7). Thus, in the hypothetical case where π_z would be very high, the nonresponse bias would be strongly attenuated even with a low response rate at the last phase.
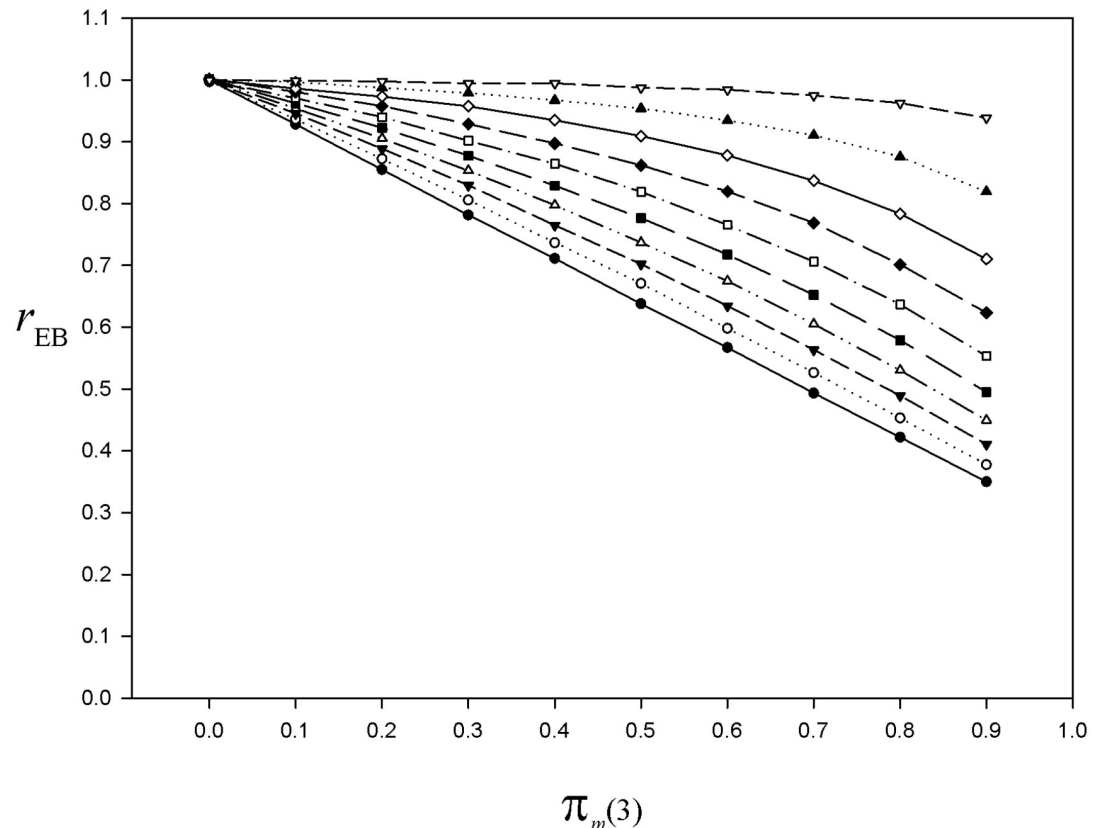
Fig 6. For the simulated population, bias index of the El-Badry estimator as a function of the nonresponse rate at the last phase (detail for moderate nonresponse rates). Curve of the bias index $r_{EB} = E(\bar{y}_{EB})/\bar{y}$ based on 100 000 simulations of the nonresponse mechanism, for $L = 2, 3, 4$, $\pi_z = 0.2$, $\pi_m(i) = 0.85$ for $1 \leq i < L$, and $\pi_m(L) = 0(0.01)0.1$ (*black dots L = 2; white circles L = 3; black triangles down L = 4*). Details in the text.

Discussion

In surveys, choices concerning the collecting mode (mail, web, phone), length, content, organization, wording and color of the questionnaire, type of outgoing postage, type of return postage, content of cover letter, the way the survey is publicized to surveyed people, the stakeholders in charge of the survey, are all factors susceptible of having an impact on final response rate and, hence, on potential nonresponse bias. Regarding questionnaire design, the simpler it is the most responses can be expected. There is however a limit to simplification and even a questionnaire which appears simple to the staff in charge of a hunting bag survey may be misunderstood by some of the surveyed people, and therefore may lead to nonresponse. In parallel, we live in a society experiencing an increasing demand for information. Consequently, more and more people are asked to participate in surveys, and they may increasingly see this as a burden. Therefore, people may become less and less inclined to cooperate [32]. This phenomena also holds for hunters of course. Since nonresponse is a psychosociological phenomena, it is very difficult to forecast which option or combination of options will have a significant positive impact on response rate. In practice, minimizing the anticipated nonresponse when designing the questionnaire requires a series of trial and error tests. For the very important issue of questionnaire design and administration, the reader is referred to [14, 72–75].

Fig 7. For the simulated population, bias index of the El-Badry estimator as a function of the nonresponse rate at the last phase, with varying values of π_z. Curve of the bias index $r_{EB} = E(\bar{y}_{EB})/\bar{y}$ based on 10 000 simulations of the nonresponse mechanism, for $L = 3$, $\pi_m(i) = 0.85$ for $1 \leq i < L$, $\pi_m(L) = 0(0.1)0.9$, and $\pi_z = 0(0.1)0.9$ (*black dots* $\pi_z = 0$; *white circles* $\pi_z = 0.1$; *black triangles down* $\pi_z = 0.2$, *white triangles up* $\pi_z = 0.3$, *black squares* $\pi_z = 0.4$, *white squares* $\pi_z = 0.5$, *black diamonds* $\pi_z = 0.6$, *white diamonds* $\pi_z = 0.7$, *black triangles up* $\pi_z = 0.8$, *white triangles down* $\pi_z = 0.9$). Details in the text.

Skalski & Millspaugh [76] state that "Estimating game harvest is among the most important activities of wildlife management agencies". Even though web-based systems tend to develop, usually hunting bag surveys still rely in part or totality upon self-administered mailed questionnaires. Despite the prominence of the nonresponse bias issue in mail hunting bag surveys [10, 14, 30, 41, 77–79], few papers have been published about statistical remedies facing this problem in this specific context, and fewer still wildlife agencies have taken this problem into account seriously and adequately. Some reports (e.g. [8]) or proceedings (e.g. [37]) on this topic do exist, but seem to have fallen into oblivion or are difficult to access. Accordingly, this paper is an opportunity to bring back to light the issue and to provide a practical, statistically sound solution, namely subsampling among nonrespondents in the framework of multiphase sampling for stratification.

In this paper we recalled the strategy proposed by Hansen & Hurwitz [48] and its generalization to any number of phases carried out by El-Badry [64]. At least in North America, the unbiased estimator introduced by Hansen & Hurwitz [48] is known (see [77], Footnote 3, [14, 80], [81], p. 239) and actually used (see [37] and [82]). Unfortunately, an unbiased sampling variance estimator was not necessarily used. For instance, after correcting the misformulated sampling variance printed in [44] (see our remark in section 2.3), Taylor et al. [82] simply used it by substituting sample estimates to population parameter values, which by no means leads to

an unbiased estimator. Oddly enough, no sampling variance estimator is given in a number of books dealing with the Hansen-Hurwitz's method (see for instance [53, 54, 65, 68, 83, 84]).

The generalization by El-Badry [64] was cited by Filion [14] but, to our knowledge, his strategy has not been used in the context of hunting bag surveys until the last French nation-wide hunting bag survey [85, 86]. MacDonald & Dillman [10] referred to El-Badry [64] but only for introduction generalities about nonresponse. Again, unfortunately for the practitioner, a sampling variance estimator was not given by El-Badry [64], nor in any of the rare books, theses or articles which address this design beyond the mere mention of its existence (see [87, 88], [62], pp. 104-105, [83], pp. 511-512, [65], p. 122, [67], p. 61 and p. 66, [68], pp. 406-409). We have filled this gap by providing an unbiased sampling variance estimator for any number of phases (section 3.4). We also provided the detailed expression of the sampling variance estimators for two- and three-phase sampling (see S1 Appendix), since such numbers of phases are the most likely to be used in practice, based on economic and logistical considerations, but it is safer to implement in a programming language the general expression we gave. We hypothesize that the lack of sampling variance estimator may have contributed to El-Badry's sampling strategy not becoming a regular element of wildlife agencies' toolbox.

For unbiased estimation of the total (or mean) and sampling variance, the El-Badry's sampling strategy requires a 100% response rate at the last phase of the multiphase sampling design, that is, when the hunters of a subsample drawn from the last mailing wave nonrespondents are interviewed (usually by phone). However, in practice, whatever the number of mailing waves, at the last phase the response rate cannot be 100%. Accordingly, the nonresponse bias cannot be totally eliminated by the multiphase sampling design. Nevertheless, a certain amount of bias attenuation should result from using the total estimator under the El-Badry's sampling strategy, depending on the nonresponse rate at the last phase L and potential magnitude of the nonresponse bias. To document this topic of paramount practical importance, we relied on Monte Carlo simulations. We found that a negative bias is induced by the nonresponse occurring at the last phase, both in estimating the mean (or total) (Figs 3 and 5) and the sampling variance (Fig 4). Moreover, the Monte Carlo study showed that the nonresponse bias attenuation (that is, when $\pi_m(L)$ is not 0) increases jointly with the number of phases (Fig 5). Actually, increasing the sampling effort with the aim of attenuating the nonresponse bias is only possible by increasing the number of phases. The fact that increasing the sampling size at any phase has no effect on the nonresponse bias should be recalled here since some authors saw this as a way to reduce the nonresponse bias (e.g. [89], p. 30).

Our Monte Carlo simulations also illustrate the fact that, in case of a very large potential nonresponse bias (caused by the conjunction of high values both of π_m and π_z), the El-Badry's sampling strategy leads to an important bias attenuation, even though the response rate at the last phase is not very high (Fig 7). As we assume in practice a moderate value for π_z, from the Monte Carlo case study we advocate that the response rate at the last phase should not be lower than 90%. Whatever the number of phases (i.e. especially with $L = 2$), a moderate underestimation of the nonresponse bias (say a maximum of 5%) is only achieved with a response rate of at least 93% in the last phase (Fig 6).

Although the cost/precision balance is an important topic, it would carry too far in this article to address these issues (see [48, 50, 62, 64, 69, 88], [52], pp. 371-372, and [53], pp. 977-979). However, two- or three-phase sampling are generally acceptable on economic and logistical grounds, and seems to provide a reasonable trade-off between additional mailing costs for multiple mailing waves, bias attenuation and increase of the sampling variance (since the sampling variance increases with the number of phases). In the case of a postseason survey conducted with an annual periodicity, in our experience, the success of the strategy we advocate in this article depends on several critical factors. First, the quality of the sampling frame is one of

the most relevant prerequisites, both in terms of coverage and correctness of contact information (postal address and phone numbers). Second, it is of the utmost importance that designing the questionnaire (either paper- or web-based) as well as receiving and processing the questionnaires be accomplished by the wildlife agency itself. We strongly advise against relying on unskilled organizations regarding hunting surveys such as market research organizations or opinion poll organizations: only printing and mailing could be outsourced. Third, the timing of mailing waves and of the last phase phone interview must be carefully planned and respected. Fourth, entry and control of the hunting bags reported (questionnaires completed) must be carried out on a continuous flow basis. Of course, is it possible that one or two surveys be necessary before entering in a perfectly mastered routine, but we think that the quality of the hunting bag estimating scheme worths these efforts.

The scope of the method reintroduced in this paper is broader than that of hunting bags surveys, and naturally covers other surveys that also have nonresponse issues [90]. By contrast, in terms of nonresponse bias, the recommendations must be domain-specific and cannot be automatically applied to other fields. In the present article, the Monte Carlo simulations were based on a nonresponse mechanism that makes sense in the field of hunting bag surveys, but which may have no phenomenological validity in another domain. The nonresponse mechanism we proposed (section 4.1) is simple but realistic enough to be useful. In this mechanism, we retain two parameters related to the nonresponse bias: the propensity to not respond (π_m) and, among the nonrespondents, the propensity to not respond because of a null harvest (π_z). In this situation, the nonresponse bias is equivalent to an undercoverage bias of the stratum U_0 (null hunting bags) by the set of respondents (see Fig 1).

It is conceivable that another nonignorable cause of nonresponse would be the fact for a hunter to have a very high hunting bag (which he/she would not be ready to disclose). Nevertheless, we advocate that this cause can be neglected compared to the issue of null bags, for at least two reasons. First, it seems unlikely that most of the very successful hunters do not respond because of their success. Indeed, it would be in contradiction with the fact that hunters tend overstating their bags for prestige or pride reasons, even though the survey was announced as anonymous (it is likely that some of the respondents do not believe this anonymity claim, because they received a nominative mail). Hunters do not seem to hesitate to report very high hunting bags, even when they exceed existing legal limits. Second, the overwhelming majority of hunters have a null harvest for a given game species, either because they were inactive or unsuccessful. For instance, in France, even for the most harvested wild bird species (i.e. without released birds), namely the common wood pigeon (*Columbia palumbus*), the proportion of hunters with null harvest was estimated at 78% according to the last survey (2013-2014 hunting season). Moreover, the proportion of hunters with null harvest for all allowed game species (about 90 species in France) was estimated at 30%. By contrast, very successful hunters are rarer.

Regarding our simulations, note that we found several other algorithms implementing the nonresponse mechanism described in section 4.1, all equivalent with reference to the nonresponse bias. We retained the algorithm that fit exactly the context of the Monte Carlo study under consideration, namely the framework of multiphase sampling for stratification. For other studies related to the same nonresponse mechanism, one of the other algorithms could be used, and will be documented at this occasion.

Although we cannot claim the generality of the finite population we used as an example for the Monte Carlo study, it is nevertheless rooted in reality. Indeed, we have set the population size to 10 000, which is the order of magnitude of the average number of (potentially) active hunters in a French department (a department is a mid-scale administrative entity used as a geographical stratum in the last nationwide French hunting bag survey, see [85, 86]). The two

parameters specifying the superpopulation model ξ are approximately the values for Eurasian teal (*Anas crecca*) hunting bags, estimated from the last nationwide French hunting bag survey (i.e. national-scale estimates). The value $\pi_m = 0.85$ corresponds approximately to the observed mean nonresponse rates (i.e. among geographical strata) for each of the two mailing waves in the last nationwide French hunting bag survey (average nonresponse rate of 86% for the first mailing wave, and of 88% for the second, [85]). Lastly, the value $\pi_z = 0.2$ corresponds to the order of magnitude of the proportion for nonrespondents in the second mailing wave, who declared by phone in the (last) third phase that they did not respond previously because of a low or null harvest (17%, see [91], encadré, p. 6). By using this example—which leads to an overestimation of about 20% with the usual estimator (see Fig 2)—we have been realistic in that overestimates of about 20% (or more) seem not to be exceptional [23, 41].

If we consider for instance a three-phase sampling design, and a response rate at the last phase greater than 90%, we can expect an underestimation of about 5% or less (Fig 6). In such a case, in terms of cost/benefit ratio, it is useless resorting to this sampling strategy when uni-phase SRSWOR leads to an overestimation equal or less than 5% (for the range of π_m and π_z values, see the contour level 1.05 on Fig 2). If we set $\pi_z = 0.20$, it might not be very useful to use this sampling strategy for instance when $\pi_m = 0.40$, since the uni-phase SRSWOR leads to an overestimation of about 9% (Fig 6). This might be the case for the Finnish hunting bag survey for which the response rate is currently about 60% (Leena Forsman, pers. comm.). At this stage of the reflexion, the genuine question is whether a given overestimation magnitude is inconsequential or not for the purpose at hand. As Chapman et al. [8] wrote, "The definition of any given error as 'inconsequential' is also quite relative, as under a different set of circumstances or in a different application such an error magnitude might not be inconsequential at all". In accordance with Fig 2, we agree with the recommendation for a 85% response rate to minimize the impact of nonresponse [79]. It must however be acknowledged that such a high response rate currently seems to be the exception rather than the rule. For instance, it is conversely the nonresponse rate that reached 85% at each mailing wave in the last French nationwide hunting bag survey. In this circumstance, the three-phase sampling design proved to be essential for attenuating a nonresponse bias that would otherwise have been far from negligible.

In the field of wildlife management, even in the most advanced countries (e.g. the U.S. or Canada, but see [1–3] for a discussion of this assertion), the nonresponse bias issue is not always addressed by the producers of hunting bag statistics [92]. In principle, one may suggest different reasons for this, which are non exclusive from each other: (i) poor awareness of the problem, (ii) financial or time constraints, (iii) hunting statistics as the result of a pure administrative request rather than a scientific question. A fourth reason could be that the variable of interest often is the trend in hunting bags rather than absolute hunting bag size (e.g. [93]), which may make more sense given the multiple biases potentially affecting absolute bag estimates [24]. As written by Wright [13]: "The important problem then is to estimate changes in the types and magnitude of the biases between years". Some authors claim that there is evidence that the nonresponse bias changes between years (e.g. [41]). At this stage, we have no general certainty, and careful case-by-case studies are needed. For trend assessment, in practice we only need that the nonresponse bias can be held relatively constant in time. It is especially of utmost importance that the nonresponse bias does not itself show a trend (up or down) over time, otherwise it will be impossible to interpret the presence/absence of a trend as representative of that of the actual hunting bags.

For trend assessment, under the nonresponse mechanism we proposed in the context of hunting bag surveys, if π_m may be considered as approximately constant, this condition must also hold for π_z since these two probabilities play a symmetrical role in producing nonresponse

bias. It is well known that the response rate shows a general decreasing trend over time, whatever the topic of the survey at hand (see for instance [47], Section 2.2). Such a situation also holds for hunting bag surveys. For instance, for the Illinois waterfowl hunter state survey, the overall response rate was 70-83% for the years 1982-1992 [94] and decreased to 44% for the 2015-2016 season [95]. In Finland, the response rate to the nationwide hunting survey was about 75% in 2012 and is currently about 60% (Leena Forsman, pers. comm.). However, it is likely that there is a threshold below which the response rate cannot decline further, depending on several factors such as the geographical scale of the survey, its periodicity, advertising for the survey and so on. Thus, at least in certain circumstances, we can think that the nonresponse rate may remain stable (the response rate cannot fall indefinitely to zero) and thus it seems possible to consider π_m as approximately constant (but only from a certain point in time that we do not know in advance). There remains the issue of π_z. Although changes in the hunting conditions may lead to a change in the proportion of null harvest among hunters taking a licence (i.e. potentially active hunters), this does not necessarily imply a change in π_z, for the propensity to report null harvest may be under the influence of several psychosociological processes. Anyway, the approximately constant nonresponse bias in time is a key assumption we think wiser not to make. Instead, it is safer using the adequate El-Badry's sampling strategy, ensuring a very moderate nonresponse bias in the estimation.

From a sampling point of view, hunting bag estimates can be established on the basis of a sample survey or *a census survey* (or simply, *a census*). A census may be viewed as a the limiting case of a sample survey, that is, when all the members of the frame are surveyed (a complete enumeration of all potentially active hunters). Whatever the type of survey, responding to a hunting bag questionnaire may be mandatory or on a voluntary basis only. When reporting hunting bags is mandatory, a fine may be provided for by the legislation in case of nonreporting (e.g. [96] for Denmark). A more effective incentive is obtained by conditioning the delivery of the license for a new hunting season on reporting the hunting bag for the previous hunting season. For instance, with such a measure, Denmark nowadays reaches a hunting bag return rate of almost 100% [21]. In practice, conditioning hunting on response is usually restricted to census. In case of a census, even when the reporting of hunting bags is mandatory, generally in the absence of a fine or prosecution such as mentioned above, the response rate is not 100%, according to hunters' compliance, which varies both at the individual and cultural level, at the nationwide or regional scale. A high response rate needs both strong adherence to the rule by the active hunter population and effective law enforcement by the authorities. So even though in some countries the response rate is nowadays close to 100% (e.g. Denmark, Norway), for a number of countries where hunting bag reporting is mandatory, the nonresponse bias issue is still relevant. Hence, in the case of a census with a moderate response rate, the El-Badry's sampling strategy might be applied (just consider the first phase sampling fraction $v_1 = 1$).

From a logistic point of view, using two or more mailing waves in hunting surveys was common in North America some time ago [7, 10, 14, 22, 26, 38, 43, 77, 79, 80, 94]. However, efforts were not always made to differentiate between the responses of the different waves [29], and the followup was only dedicated to gather more responses. For instance, Anderson et al. [94] report that in the case of the Illinois Waterfowl Hunter Survey for the years 1982-1992, the initial mailing and 2 followups to nonrespondents generated response rates of 70-83%. Under such conditions, the total hunting bag is often estimated by pooling the responses of the successive mailing waves, and possibly those gathered by telephone follow-up (e.g. [26]). When the responses from successive mailings are tabulated separately, some authors rely on the assumption that there exists a continuum of respondent types which

range from highly motivated to unmotivated individuals, that is to say, a linear increase in nonresponse bias with successive waves. Under this assumption, they fit a regression model aimed at correcting for nonresponse bias [10, 77], following an idea that goes back at least to Clausen & Ford [40]. We are not convinced by this approach. First, in accordance with Atwood [7], in mailing follow-up we think that "differences exhibited in data from successive waves of requests must be attributed to errors other than nonresponse errors". Besides, Sen [97] suggested that "the average kill per hunter may not change appreciably when successive reminders are used to reduce nonresponse". Second, the usefulness of a regression model fit on the basis of very few points can be questioned. Anyway, in this context, using a regression model deserves a thorough study, which could be the aim of another paper. Finally, at this stage of our knowledge, again, we advocate the El-Badry's sampling strategy especially because no assumptions are required. Moreover, a design-based approach allows to manage a great number of game species in the same survey, without having to assume that the nonresponse bias affects them all the same way—which is certainly not the case [37]—and without having to treat each species separately, in the sense that the same estimators apply to all in an automatic way.

If one considers as relevant the nonresponse mechanism we proposed in this paper, then there is a need to document the propensity of non respondents to not respond because of a null harvest (π_z), in addition to the propensity to not respond (π_m). Communication and educational actions towards hunters are needed for decreasing both parameters. This must be done through different channels, preferably at a local scale, by hunter's clubs and organizations. Anyway, we advocate that the El-Badry's sampling strategy is a good way to tackle the nonresponse bias issue, provided that the nonresponse rate at the last phase remains low. Needless to say, this sampling strategy has no effect on the other nonsampling biases such as the misclassification bias mentioned in the introduction. The negligible influence of misclassification error on the final estimates cannot be taken for granted, although some studies seem to be reassuring about identification errors (e.g. [98]). On the contrary, for some game species, it can be an important source of bias which deserves more studies (see [21] for a recent contribution). Some room for improvement hence remains to ensure the quality of hunting bag surveys but, in total accordance with Pendleton [78], we think that prerequisites are relying on a sampling frame of high quality (good coverage, accurate postal addresses and phone numbers) and attenuating the nonresponse bias using repeated sampling of nonrespondents.

Adaptive harvest management is gradually becoming the norm in wildlife agencies, following the very successful example of North Americans for waterbird hunting [99–101]. When hunting bag estimates are inputs in adaptive harvest management models, attenuating the nonresponse bias becomes of overwhelming importance, otherwise overestimates will be taken into account in the calculations, with the risk of misleading conclusions and unsuitable management recommendations. A recent analysis suggests adaptive harvest management is achievable even with minimum data availability, but regular and robust estimates of hunting bags are among those few absolute prerequisites [6]. Subsampling the nonrespondents in the framework of multiphase sampling for stratification is a usable solution, offering a good protection against high nonresponse bias.

Supporting information

S1 Appendix. Statistical and formal complements.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Philippe Aubry.

Methodology: Philippe Aubry.

Writing – original draft: Philippe Aubry, Matthieu Guillemain.

Writing – review & editing: Philippe Aubry, Matthieu Guillemain.

References

1. Artelle KA. When science-based management isn't. Science. 2014; 343(6177):1311. https://doi.org/10.1126/science.343.6177.1311-a PMID: 24653018

2. Artelle KA, Reynolds JD, Treves A, Walsh JC, Paquet PC, Darimont CT. Hallmarks of science missing from North American wildlife management. Science Advances. 2018; 4(3):eaao0167. https://doi.org/10.1126/sciadv.aao0167 PMID: 29532032

3. Artelle KA. Is wildlife conservation policy based in science? American Scientist. 2019; 107(1):38–45. https://doi.org/10.1511/2019.107.1.38

4. Anderson MG, Padding PI. The North American approach to waterfowl management: synergy of hunting and habitat conservation. International Journal of Environmental Studies. 2015; 72(5):810–829. https://doi.org/10.1080/00207233.2015.1019296

5. Holling CS. Adaptive environmental assessment and management. Chichester, UK: John Wiley & Sons; 1978.

6. Johnson FA, Alhainen M, Fox AD, Madsen J, Guillemain M. Making do with less: must sparse data preclude informed harvest strategies for European waterbirds? Ecological Applications. 2018; 28 (Suppl 2):427–441. https://doi.org/10.1002/eap.1659 PMID: 29205644

7. Atwood EL. Validity of mail survey data on bagged waterfowl. Journal of Wildlife Management. 1956; 20(1):1–16.

8. Chapman DG, Overton WS, Finkner AL. Methods of estimating dove kill. Raleigh, North Carolina, USA: Institute of Statistics, North Carolina State College; 1959.

9. Atwood EL. Abstract: A procedure for removing the effect of response bias errors from waterfowl hunter questionnaire responses. Biometrics. 1958; 14(1):132–133.

10. MacDonald D, Dillman EG. Techniques for estimating non-statistical bias in big game harvest surveys. Journal of Wildlife Management. 1968; 32(1):119–129. https://doi.org/10.2307/3798246

11. Sen AR. Response errors in Canadian waterfowl survey. Journal of Wildlife Management. 1973; 37 (4):485–491. https://doi.org/10.2307/3800313

12. Sen AR. Some nonsampling errors in the Canadian waterfowl mail survey. Journal of Wildlife Management. 1972; 36(3):951–954. https://doi.org/10.2307/3799455

13. Wright VL. Causes and effects of biases on waterfowl harvest estimates. Journal of Wildlife Management. 1978; 42(2):251–262. https://doi.org/10.2307/3800262

14. Filion FL. Human surveys in wildlife management. In: Schemnitz SD, editor. Wildlife Management Techniques Manual. Washington, DC, USA: The Wildlife Society; 1980. p. 441–453.

15. Chu A, Eisenhower D, Hay M, Morganstein D, Neter J, Waksberg J. Measuring the recall error in self-reported fishing and hunting activities. Journal of Official Statistics. 1992; 8(1):19–39.

16. Miller CA, Anderson WL. Digit preference in reported harvest among Illinois waterfowl hunters. Human Dimensions of Wildlife. 2002; 7(1):55–65. https://doi.org/10.1080/108712002753574783

17. Beaman J. Comment on "Digit preference in reported harvest among Illinois waterfowl hunters" by Craig A. Miller and William L. Anderson. Human Dimensions of Wildlife. 2002; 7(1):67–72. https://doi.org/10.1080/108712002753574792

18. Beaman J, Vaske JJ, Miller CA. Cognitive processes in hunters' recall of participation and harvest estimates. Journal of Wildlife Management. 2005; 69(3):967–975. https://doi.org/10.2193/0022-541X(2005)069%5B0967:CPIHRO%5D2.0.CO;2

19. Beaman J, Vaske JJ, Miller CA. Hunting activity record-cards and the accuracy of survey estimates. Human Dimensions of Wildlife. 2005; 10(4):285–292. https://doi.org/10.1080/10871200500292876

20. Crissey WF. Calculators and Ouija boards. In: Hawkins AS, Hanson RC, Nelson HK, Reeves HM, editors. Flyways. Pioneering waterfowl management in North America. Washington, DC, USA: U.S. Department of the Interior, Fish and Wildlife Service; 1984. p. 259–271.

21. Christensen TK, Balsby TS, Mikkelsen P, Lauritzen T. Vildtudbyttestatistik og vingeundersøgelsen for jagtsæsonerne 2015/16 og 2016/17. Aarhus, Danemark: Aarhus University; 2017.

22. Sen AR. Some recent developments in waterfowl sample survey techniques. Applied Statistics. 1971; 20(2):139–147. https://doi.org/10.2307/2346462

23. Hayne DW. Investigation of mail survey reporting by waterfowl hunters. Laurel, Maryland, USA: U.S. Fish and Wildlife Service, Bureau of Sport Fisheries and Wildlife, Patuxent Wildlife Research Center; 1964.

24. Martin EM, Carney SM. Population ecology of the mallard: IV. A review of duck hunting regulations, activity, and success, with special reference to the mallard. Resource publication 130 of the Fish and Wildlife Service. Washington, DC, USA: U.S. Department of the Interior, Fish and Wildlife Service; 1977.

25. Hedayat AS, Sinha BK. Design and inference in finite population sampling. New York, New York, USA: John Wiley & Sons; 1991.

26. Barker RJ, Geissler PH, Hoover BA. Sources of nonresponse to the federal waterfowl hunter questionnaire survey. Journal of Wildlife Management. 1992; 56(2):337–343. https://doi.org/10.2307/3808832

27. Sen AR. On the bias in estimation due to imperfect frame in the Canadian waterfowl surveys. Journal of Wildlife Management. 1970; 34(4):703–706. https://doi.org/10.2307/3799134

28. Sen AR. Developments in migratory game bird surveys. Journal of the American Statistical Association. 1976; 71(353):43–48. https://doi.org/10.1080/01621459.1976.10481474

29. Trost RE, Carney SM. Measuring the waterfowl harvest. In: Beattie KH, editor. Proceedings of the Sixth International Waterfowl Symposium. Long Grove, Illinois, USA: Ducks Unlimited; 1989. p. 134–147.

30. Couling LM, Sen AR, Martin EM. Reliability of kill and activity estimates in the U.S. waterfowl hunter survey. Washington, DC, USA: U.S. Department of the Interior, Fish and Wildlife Service; 1982.

31. Särndal CE, Lundström S. Estimation in surveys with nonresponse. Chichester, UK: John Wiley & Sons; 2005.

32. Bethlehem J. Cross-sectional research. In: Ader HJ, Mellenbergh GJ, editors. Research methodology in the social, behavioural and life sciences. Thousand Oaks, California, USA: Sage Publications; 1999. p. 110–142.

33. Enders CK. Applied missing data analysis. New York, New York, USA: The Guilford Press; 2010.

34. Arnab R. Survey sampling. Theory and applications. San Diego, California, USA: Academic Press; 2017.

35. Särndal CE, Swensson B, Wretman JH. Model assisted survey sampling. New York, New York, USA: Springer; 1992.

36. Tillé Y. Sampling algorithms. New York, New York, USA: Springer; 2006.

37. Overton WS. Post season mail survey techniques and procedures. In: Proceedings of the Annual Conference. vol. 7. Knoxville, Tennessee, USA: Southeastern Association of Game and Fish Commissioners; 1953. p. 71–81.

38. Martinson RK, Whitesell DE. Biases in a mail questionnaire survey of upland game hunters. Transactions of the North American Wildlife and Natural Resource Conference. 1964; 29:287–294.

39. Franzen R, Lazarsfeld PF. Mail questionnaire as a research problem. Journal of Psychology. 1945; 20(2):293–320. https://doi.org/10.1080/00223980.1945.9917260

40. Clausen JA, Ford RN. Controlling bias in mail questionnaires. Journal of the American Statistical Association. 1947; 42(240):497–511. https://doi.org/10.1080/01621459.1947.10501943

41. Barker RJ. Nonresponse bias in New Zealand waterfowl harvest surveys. Journal of Wildlife Management. 1991; 55(1):126–131. https://doi.org/10.2307/3809249

42. Strickland MD, Harju HJ, McCaffery KR, Miller HW, Smith LM, Stoll RJ. Harvest management. In: Bookhout TA, editor. Research and management techniques for wildlife and habitats. Fifth edition. Revised. Bethesda, Maryland, USA: The Wildlife Society; 1996. p. 445–473.

43. Eberhardt L, Murray RM. Estimating the kill of game animals by licensed hunters. In: Proceedings of the Social Statistics Section, American Statistical Association. Alexandria, Virginia, USA: American Statistical Association; 1960. p. 182–188.

44. Lessler JT, Kalsbeek WD. Nonsampling error in surveys. New York, New York, USA: John Wiley & Sons; 1992.

45. Rubin DB. Multiple imputation for nonresponse in surveys. New York, New York, USA: John Wiley & Sons; 1987.

46. Bethlehem J. Weighting nonresponse adjustments based on auxiliary information. In: Groves RM, Dillman DA, Eltinge JL, Little RJA, editors. Survey nonresponse. New York, New York, USA: John Wiley & Sons; 2002. p. 275–287.

47. Brick JM, Montaquila JM. Nonresponse and weighting. In: Pfeffermann D, Rao CR, editors. Handbook of Statistics 29A. Sample surveys: design, methods and applications. Oxford, UK: Elsevier; 2009. p. 163–185.

48. Hansen MH, Hurwitz WN. The problem of non-response in sample surveys. Journal of the American Statistical Association. 1946; 41(236):517–529. https://doi.org/10.1080/01621459.1946.10501894 PMID: 20279350

49. Hidiroglou M, Estevao V. Dealing with nonresponse using follow-up. In: Proceedings of the Joint Statistical Meetings—Section on Survey Research Methods, ASA. Alexandria, Virginia, USA: American Statistical Association; 2013. p. 1478–1489.

50. Dykes L, Singh S, Sedory SA, Louis V. Calibrated estimators of population mean for a mail survey design. Communications in Statistics—Theory and Methods. 2015; 44(16):3403–3427. https://doi.org/10.1080/03610926.2013.841932

51. Schmidt JI, Kellie KA, Chapin FS. Detecting, estimating, and correcting for biases in harvest data. Journal of Wildlife Management. 2015; 79(7):1152–1162. https://doi.org/10.1002/jwmg.928

52. Cochran WG. Sampling techniques. Third edition. New York, New York, USA: John Wiley & Sons; 1977.

53. Singh S. Advanced sampling theory with applications: how Michael'selected' Amy, Volume II. vol. 2. Dordrecht, The Netherlands: Kluwer Academic Publishers; 2003.

54. Gupta AK, Kabe DG. Theory of sample surveys. Singapore: World Scientific; 2011.

55. Jebe EH. Multiphase sampling. In: Kotz S, Balakrishnan N, Read CB, Vidakovic B, Johnson NL, editors. Encyclopedia of statistical sciences. Second edition. vol. 8. Hoboken, New Jersey, USA: John Wiley & Sons; 2006. p. 5051–5053.

56. Rao JNK. On double sampling for stratification and analytical surveys. Biometrika. 1973; 60(1):125–133. https://doi.org/10.1093/biomet/60.1.125

57. Johnston DC. Theory and application of selected multilevel sampling designs [Ph.D. thesis]. Colorado State University. Fort Collins, Colorado, USA; 1982.

58. Jeyaratnam S, Bowden DC, Graybill FA, Frayer WE. Estimation in multiphase designs for stratification. Forest Science. 1984; 30(2):484–491.

59. Rao JNK. Some nonresponse sampling theory when the frame contains an unknown amount of duplication. Journal of the American Statistical Association. 1968; 63(321):87–90. https://doi.org/10.1080/01621459.1968.11009223

60. Chaudhuri A, Stenger H. Survey sampling. Theory and methods. Second edition. Boca Raton, Florida, USA: Chapman & Hall/CRC; 2005.

61. Singh R, Singh Mangat N. Elements of survey sampling. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1996.

62. Rao PSRS. Randomization approach. In: Madow WG, Olkin I, Rubin DB, editors. Incomplete data in sample surveys. Volume 2: theory and bibliographies. New York, New York, USA: Academic Press; 1983. p. 97–105.

63. Lohr SL. Sampling: design and analysis. Second edition. Boston, Massachusetts, USA: Brooks/Cole; 2010.

64. El-Badry MA. A sampling procedure for mailed questionnaire. Journal of the American Statistical Association. 1956; 51(274):209–227. https://doi.org/10.1080/01621459.1956.10501321

65. Raj D, Chandhok P. Sample survey theory. New Delhi, India: Narosa Publishing House; 1998.

66. Ramakrishnan MK. Some results on the comparison of sampling with and without replacement. Sankhya Series A. 1969; 31(3):333–324.

67. Siripornpibul S. Survey designs and compensation methods for nonresponse problems [Ph.D. thesis]. University of Canterbury. Canterbury, UK; 2001.

68. Mukhopadhyay P. Theory and methods of survey sampling. Second edition. New Delhi, India: PHI Learning; 2009.

69. Srinath KP. Multiphase sampling in non-response problems. Journal of the American Statistical Association. 1971; 66(335):583–586. https://doi.org/10.1080/01621459.1971.10482310

70. Johnson NL, Kemp AW, Kotz S. Univariate discrete distributions. Third edition. Hoboken, New Jersey, USA: John Wiley & Sons; 2005.

71. L'Ecuyer P. Good parameters and implementations for combined multiple recursive random number generators. Operations Research. 1999; 47(1):159–164. https://doi.org/10.1287/opre.47.1.159

72. Filion FL. Importance of question wording and response burden in hunter surveys. Journal of Wildlife Management. 1981; 45(4):873–882. https://doi.org/10.2307/3808096

73. Dillman DA. The design and administration of mail surveys. Annual Review of Sociology. 1991; 17:225–249. https://doi.org/10.1146/annurev.so.17.080191.001301

74. Peterson RA. Constructing effective questionnaires. Thousand Oaks, California, USA: Sage Publications; 2000.

75. Saris WE, Gallhofer IN. Design, evaluation, and analysis of questionnaires for survey research. Hoboken, New Jersey, USA: John Wiley & Sons; 2007.

76. Skalski JR, Millspaugh JJ. The impact of hunter postseason questionnaire design on big game harvest estimation. Wildlife Society Bulletin. 2006; 34(2):329–337. https://doi.org/10.2193/0091-7648(2006)34%5B329:TIOHPQ%5D2.0.CO;2

77. Filion FL. Estimating bias due to nonresponse in mail surveys. Public Opinion Quarterly. 1975; 39 (4):482–492. https://doi.org/10.1086/268245

78. Pendleton GW. Nonresponse patterns in the federal waterfowl hunter questionnaire survey. Journal of Wildlife Management. 1992; 56(2):344–348. https://doi.org/10.2307/3808833

79. Sheriff SL, Schulz JH, Bales BD, Moore MT, Padding PI, Shipes DA. The current reliability of harvest information program surveys. In: Ver Steeg JM, Elden RC, Dolton DD, Padding PI, editors. Harvest information program: evaluation and recommendations. Washington, DC, USA: International Association of Fish and Wildlife Agencies, Migratory Shore and Upland Game Bird Working Group, Ad Hoc Committee on HIP; 2002. p. 51–68.

80. Ryel LA. The legal deer kill—How it's measured. In: Hine RL, Nehls S, editors. White tailed deer population management in the north central states. Proceedings of a Symposium held at the 41st Midwest Fish and Wildlife Conference, Urbana, Illinois, 10 December 1979. Bethesda, Maryland, USA: The Wildlife Society; 1980. p. 37–45.

81. Skalski JR, Ryding KE, Millspaugh JJ. Wildlife demography. Analysis of sex, age and count data. Burlington, Massachusetts, USA: Elsevier Academic Press; 2005.

82. Taylor CE, Otis DL, Hill HS, Ruth CR. Design and evaluation of mail surveys to estimate deer harvest parameters. Wildlife Society Bulletin. 2000; 28(3):717–723.

83. Sukhatme PV, Sukhatme BV, Sukhatme S, Asok C. Sampling theory of surveys with applications. Third edition. Ames, Iowa, USA: Iowa State University Press; 1984.

84. Chaudhuri A. Modern survey sampling. Boca Raton, Florida, USA: Chapman and Hall/CRC; 2014.

85. Aubry P, Anstett L, Ferrand Y, Reitz F, Klein F, Ruette S, et al. Enquête nationale sur les tableaux de chasse à tir. Saison 2013-2014. Résultats nationaux. Faune Sauvage. 2016; 310(supplément):1–8.

86. Guillemain M, Aubry P, Folliot B, Caizergues A. Duck hunting bag estimates for the 2013/14 season in France. Wildfowl. 2016; 66:126–141.

87. Foradori GT. Some non-response sampling theory for two-stage designs [Ph.D. thesis]. North Carolina State College. Raleigh, North Carolina, USA; 1961.

88. Hughes E, Rao JNK. Some problems of optimal allocation in sample surveys involving inequality constraints. Communications in Statistics—Theory and Methods. 1979; 8(15):1551–1574. https://doi.org/10.1080/03610927908827851

89. Landry P. Preliminary report on methods for collecting game bag statistics in European countries. In: Leeuwenberg F, Hepburn I, editors. Working group on game statistics. Proceedings of the second meeting, 6-7 octobre 1982, Doorwerth, Netherlands. Zoetermeer, The Netherlands: IUGB Working Group on Game Statistics, Wildlife Management Division; 1983. p. 25–46.

90. White PCL, Vaughan Jennings N, Renwick AR, Barker NHL. Questionnaires in ecology: a review of past use and recommendations for best practice. Journal of Applied Ecology. 2005; 42(3):421–430. https://doi.org/10.1111/j.1365-2664.2005.01032.x

91. Aubry P. Enquêtes sur les tableaux de chasse: pourquoi est-il essentiel d'y répondre, même quand on n'a rien prélevé. Faune Sauvage. 2017; 315:4–8.

92. Rupp SP, Ballard WB, Wallace MC. A nationwide evaluation of deer hunter harvest survey. Wildlife Society Bulletin. 2000; 28(3):570–578.

93. Massei G, Kindberg J, Licoppe A, Gačić D, Šprem N, Kamler J, et al. Wild boar populations up, numbers of hunters down? A review of trends and implications for Europe. Pest Management Science. 2015; 71(4):492–500. https://doi.org/10.1002/ps.3965 PMID: 25512181

94. Anderson WL, Thornburg DD, Whitton RM. Estimating Canada goose harvest in southern Illinois quota zones. Wildlife Society Bulletin. 1996; 24(2):233–237.

95. Williams BD, Schweizer LA, Campbell LK, Miller CA. Illinois waterfowl hunter report: harvest, youth hunts, and season preferences. Champaign, Illinois, USA: Illinois Natural History Survey; 2016.

96. Asferg T. Manglende indberetninger til vildtudbyttestatistikken i jagtsæsonen 2006/07. Aarhus, Danemark: Danmarks Miljøundersøgelser, Aarhus University; 2008.

97. Sen AR. Relative efficiency of sampling systems in the Canadian waterfowl harvest survey. Biometrics. 1970; 26(2):315–326. https://doi.org/10.2307/2529078

98. Wilson BC, Rohwer FC. In-hand duck identification by hunters at Mississippi flyway public hunting areas. Wildlife Society Bulletin. 1995; 23(3):472–480.

99. Nichols JD, Runge MC, Johnson FA, Williams BK. Adaptive harvest management of North American waterfowl populations: a brief history and future prospects. Journal of Ornithology. 2007; 148(Suppl 2):S343–S349. https://doi.org/10.1007/s10336-007-0256-8

100. Madsen J, Bunnefeld N, Nagy S, Griffin C, Defos du Rau P, Mondain-Monval JY, et al. Guidelines on sustainable harvest of migratory waterbirds. Revision 1. AEWA Conservation guidelines No. 5, AEWA Technical Series No. 62. Bonn, Germany: UNEP/AEWA; 2015.

101. Madsen J, Williams JH, Johnson FA, Tombre IM, Dereliev S, Huijken E. Implementation of the first adaptive management plan for a European migratory waterbird population: the case of the Svalbard pink-footed goose *Anser brachyrhynchus*. Ambio. 2017; 46(Suppl 2):S275–S289. https://doi.org/10.1007/s13280-016-0888-0