Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

The inverse-variance trap: A simple ratio fix for combining skewed data

by Anton Grafström, Wilmer Prentius and Åsa Ranlund

Release date: June 29, 2026

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Survey Methodology, June 2026
Vol. 52, No. 1, pp. 117-128
Statistics Canada, Catalogue No. 12-001-X

The inverse-variance trap: A simple ratio fix for combining skewed data

Anton Grafström, Wilmer Prentius and Åsa Ranlund[1]

Abstract

Combining estimates from independent surveys via inverse-variance weights can lead to negative bias when unknown variances are estimated and the target variable is non-negative and positively skewed. In such cases, strong positive correlations typically arise between the estimators and their corresponding variance estimators, causing standard linear combinations with inverse-variance weights to exhibit negative bias. We introduce a strikingly simple method to reduce bias: replace the standard weight with the ratio of the estimator to the variance estimator. Under a linear model linking the two, we show that the new ratio-weighted estimator is approximately unbiased, whereas the conventional inverse-variance combination exhibits downward bias. Through simulations, we demonstrate that the new method brings both the bias and the mean squared error closer to the optimum for a wide range of different target variables. As our method uses only standardly reported summary statistics, it can be immediately adopted to reduce this widespread bias and improve the reliability of scientific findings in various fields.

Key Words: Combining survey information; Design-based sampling; Environmental surveys; Linear combination estimator.

1. Introduction

Combining information from multiple independent sources is a fundamental practice in the empirical sciences. To achieve the highest possible precision, researchers routinely synthesize the results of different experiments or surveys using a weighted average. The gold-standard approach, taught in textbooks and embedded in meta-analysis software, is inverse-variance weighting, where each estimate is weighted by the reciprocal of its variance (Cochran, 1954). As the true variances of the estimators are unknown for most (if not all) practical applications, variance estimates are often used instead of true variances. This method is statistically optimal under standard assumptions, providing the most precise unbiased linear combination of the available information. Its use is ubiquitous, forming the bedrock of evidence synthesis in a variety of fields.

However, the optimality of inverse-variance weighting hinges on assumptions that are frequently violated in practice. A critical, yet sometimes overlooked, challenge arises when the target variables are non-negative and exhibit substantial positive skewness. This data structure is not an edge case; it is the norm for a vast range of quantities. In epidemiology, disease counts in a region are often zero or low, with occasional hotspots. In economics, wealth is heavily skewed and is usually concentrated among a small proportion of individuals. In ecology, researchers are interested in the abundance of rare species or habitat area, where a large part of the sampling frame produces a zero response (Fisher, Corbet and Williams, 1943;

1. Anton Grafström, Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden. E-mail: anton.grafstrom@slu.se; Wilmer Prentius, Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden. E-mail: wilmer.prentius@slu.se; Åsa Ranlund, Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden. E-mail: asa.ranlund@slu.se.

Preston, 1948). In all these cases, the data are characterized by a preponderance of small values and a tail of large ones.

This characteristic, as noted by Gregoire and Schabenberger (1999) and further discussed by Grafström, Ekström, Jonsson, Esseen and Ståhl (2019), leads to a strong positive correlation between the estimator of a total and its corresponding variance estimator. One consequence of this dependence is that smaller estimates tend to be paired with smaller estimated variances, resulting in narrower confidence intervals. This asymmetry increases the likelihood that confidence intervals fall below the true value more often than above (Gregoire and Schabenberger, 1999).

When estimates from multiple independent surveys are combined using a linear combination with inverse-variance weights, this correlation can introduce a systematic bias. The survey with the smaller estimated total tends to have a smaller estimated variance and therefore receives a disproportionately large weight, resulting in a negatively biased combined estimator. This issue is particularly problematic when the variable of interest is rare or patchily distributed, precisely the situations in which combining estimates is most desirable to improve precision. There is therefore a clear need for an improved estimator that accounts for the correlation structure and mitigates this bias.

In this paper, we assume that each survey provides an unbiased estimator of the total, implying full coverage of the target population in each case. When multiple frames are required for full coverage, other methods must be considered; see Elliott, Raghunathan and Schenker (2018) for a review of such approaches.

As a motivating example, consider Sweden's National Forest Inventory (NFI), a sample-based system that provides national estimates for more than 100 target variables (Fridman, Holm, Nilsson, Nilsson, Ringvall and Ståhl, 2014). The NFI combines a permanent sample, optimized for monitoring temporal trends, with a temporary sample, designed to estimate the current state. To improve efficiency, the estimates of these two components must be combined. Traditionally, this has been done using weights inversely proportional to the estimated variances, as sample sizes are not well defined. However, for some variables, the NFI alone does not provide sufficient information, necessitating the integration of data from other independent surveys such as the National Inventory of Landscapes in Sweden (NILS). This integration is particularly critical in reporting the Swedish Annex I habitats under Article 17 of the EU Habitats Directive, where it is essential to utilize all available sources of information.

To address the bias and inefficiency introduced by conventional inverse-variance-based weighting in such contexts, we propose a new weighting strategy that explicitly accounts for the positive correlation between estimators and their variance estimators. The method is designed to reduce both the bias and the mean squared error of the combined estimator. Alternative weighting schemes, such as deriving pooled variance weights, have been proposed to mitigate these biases. Grafström et al. (2019) derived general pooled variance estimators for any design and provided the theory for combining independent samples (pooling data) from general designs. However, such methods are much more restrictive, as they require full knowledge of the designs and samples and also the same type of sampling unit in all surveys. In meta-

analysis, the bias introduced by skewed data has been dealt with in different ways, such as the logarithmic transformation of variables (Higgins, White and Anzures-Cabrera, 2008). Transformation introduces bias itself, requires access to full data, and is not applicable to general sampling designs.

This work contributes to the literature on the combination of information from independent surveys, improving the classical inverse-variance estimator (Cochran, 1954) for settings involving variables with positive skew. By grounding our method in the theory of unbiased estimation, we offer a robust and interpretable alternative to combine the results of multiple surveys. The remainder of this paper is structured as follows. Section 2 introduces the proposed weighting method; Section 3 illustrates its application with an example; and Section 4 concludes with final remarks.

2. Method

The objective is to estimate a population total, $Y = \sum_{i \in U} y_i$, of a target variable that takes the value y_i for unit i in the population $U = \{1,2,3,\ldots,N\}$ of size N. We assume that the target variable is non-negative and has a positively skewed distribution, i.e. that there are more small values than large values. Let $\hat{Y}_j$, for $j = 1,\ldots,K$ be K independent design-unbiased estimators of Y. Their variances are denoted as $V_j = V(\hat{Y}_j)$ along with the corresponding variance estimators $\hat{V}_j = \hat{V}(\hat{Y}_j)$, with $E(\hat{Y}_j) = V_j$, for $j = 1,\ldots,K$. It is well-known that the linear combination estimator

$$\hat{Y}_\alpha = \sum_{j=1}^{K} \alpha_j \hat{Y}_j, \quad \text{with} \quad \sum_{j=1}^{K} \alpha_j = 1, \tag{2.1}$$

is unbiased and

$$\alpha_j = \frac{1/V_j}{\sum_{i=1}^{K} 1/V_i},$$

is optimal in the sense that it has minimum variance among the possible linear combinations, see Shahar (2017) for a formal proof. In the combined estimator (2.1), each individual estimator is weighted by the inverse of its variance, and the weights are normalized to sum to unity. The naïve combination, where (2.1) is computed with variance estimates, $\hat{V}_1,\ldots,\hat{V}_K$, instead of the true variances, is denoted $\hat{Y}_{\hat{\alpha}}$.

If $E(\hat{Y}_j) = Y$ for $j = 1,2,\ldots,K$, and the weights $\hat{\alpha}_j \propto \hat{V}_j^{-1}$ sum to unity, then it follows that $E(\hat{Y}_{\hat{\alpha}}) = Y + \sum_{j=1}^{K} \mathrm{Cov}(\hat{\alpha}_j, \hat{Y}_j)$. Thus, a non-zero correlation between the weight $\hat{\alpha}_j$ and the estimator $\hat{Y}_j$ introduces bias. In our situation, where the target variable is non-negative and has a positively skewed distribution, it follows from the results by Grafström et al. (2019) that $\hat{Y}_j$ is roughly proportional to $\hat{V}_j$, making $\mathrm{Cov}(\hat{\alpha}_j, \hat{Y}_j)$ negative. The estimator $\hat{Y}_{\hat{\alpha}}$ then tends to be negatively biased.

Due to the approximate proportionality between $\hat{Y}_j$ and $\hat{V}_j$, we can theoretically construct a correction to $\hat{V}_j$. If Y were known, we would get an improved estimator of the variance of $\hat{Y}_j$ by the ratio estimator

$$\hat{V}_{jR} = \frac{Y}{\hat{Y}_j}\hat{V}_j,$$

as e.g. a low outcome of $\hat{Y}_j$ would imply a proportionally equal low outcome of $\hat{V}_j$ when $\hat{Y}_j$ and $\hat{V}_j$ are proportional. Thus, these ratio estimators of variances V_j effectively corrects for too low or too high outcomes of $\hat{V}_j$ by multiplication with the ratio $Y/\hat{Y}_j$. These theoretical variance estimators are then inserted in place of the true variances when we compute the weights. This results in the new estimated weights

$$\hat{\alpha}_{jR} = \frac{1/\hat{V}_{jR}}{\sum_{i=1}^{K} 1/\hat{V}_{iR}} = \frac{\hat{Y}_j/\hat{V}_j}{\sum_{i=1}^{K} \hat{Y}_i/\hat{V}_i}. \tag{2.2}$$

The unknown parameter Y cancels out in the derivation of our proposed weights. By substituting α_j with $\hat{\alpha}_{jR}$ in equation (2.1), we obtain our linear combination estimator, denoted $\hat{Y}_{\hat{\alpha}_R}$. In this estimator, each individual estimator is weighted by the ratio of the estimator to its corresponding variance estimator. The optimality of this weighting scheme is achieved with perfect proportionality between the estimators and their respective variance estimators. If the proportionality is approximate, then we expect the proposed weighting to be a good approximation to the optimal solution. Due to the assumption of a high correlation between $\hat{V}_j$ and $\hat{Y}_j$ we can use the model

$$\hat{V}_j = \gamma_j \hat{Y}_j + \varepsilon_j, \tag{2.3}$$

where $\gamma_j > 0$ are constants and $E(\varepsilon_j) = 0$ with ε_j independent of $\hat{Y}_j$ for $j = 1,\ldots,K$.

Theorem 1 (Bias reduction)**.** *Let $\hat{Y}_1,\ldots,\hat{Y}_K$ be unbiased estimators of a population total Y from K independent surveys, with the corresponding unbiased variance estimators $\hat{V}_1,\ldots,\hat{V}_K$. Assume the target variable is non-negative with a positively skewed distribution such that model (2.3) holds. Then the approximate (first-order) bias of $\hat{Y}_{\hat{\alpha}}$ is*

$$Bias(\hat{Y}_{\hat{\alpha}}) \approx -\frac{\sum_{j=1}^{K} \frac{\gamma_j}{V_j}\left(\sum_{i\neq j} 1/V_i\right)}{\left(\sum_{j=1}^{K} 1/V_j\right)^2},$$

whereas the approximate (first-order) bias of $\hat{Y}_{\hat{\alpha}_R}$ is zero.

A proof of Theorem 1 is provided in the appendix. Theorem 1 indicates that the ratio-weighted estimator $\hat{Y}_{\hat{\alpha}_R}$ has a lower absolute bias than the naïve combination $\hat{Y}_{\hat{\alpha}}$ when the estimators are roughly proportional to their variance estimators. The core message of Theorem 1 is that, when each estimator $\hat{Y}_j$ of a total is positively correlated with its variance estimator $\hat{V}_j$, the naïve inverse-variance weights tend to assign "too much" weight to the survey that happens to produce a smaller estimate (and thus a smaller estimated variance). In practice, positively skewed and non-negative variables (e.g., rare-species counts or patchy habitat areas) often produce $\rho_j = \mathrm{Corr}(\hat{Y}_j, \hat{V}_j)$ above 0.8, making $\hat{Y}_{\hat{\alpha}_R}$ preferable to $\hat{Y}_{\hat{\alpha}}$.

The use of this improved estimator requires only knowledge of the estimates and their associated estimated variances.

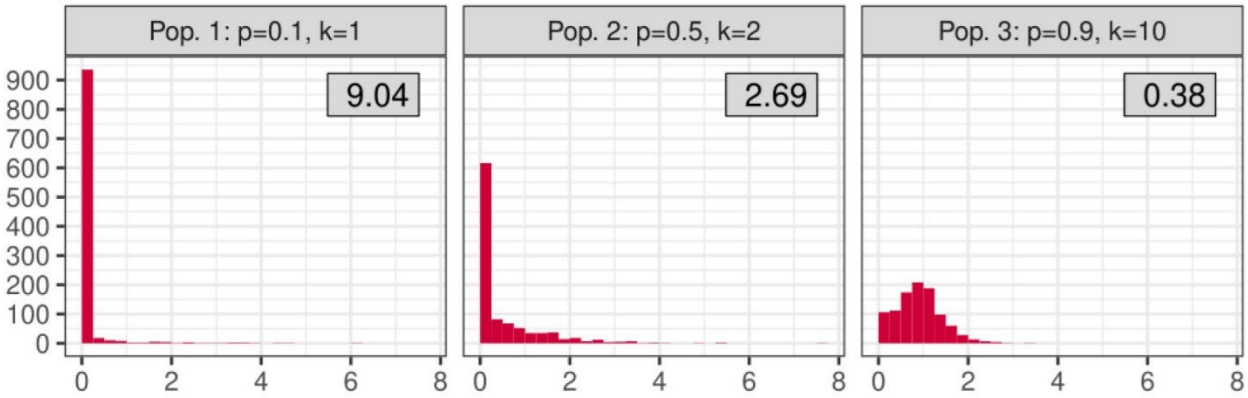
3. Simulation example

Here we provide an example for the case of $K = 3$ different estimators to illustrate the performance of the new weighting suggested in Section 2. Three populations, shown in Figure 3.1, were generated according to

$$Z_i = 1(i \le Np)\frac{X_i}{k}, X_i \sim \chi^2(k),$$

where p denotes the proportion of non-zero values out of the total population size of $N =$ 1,000 and $k =$ 1, 2 and 10.

Figure 3.1 Histograms of the three different populations used in the simulation study. p denotes the number of non-zero units in the population, and k denotes the degrees-of-freedom of the (mean-scaled) chi-squared distribution $\chi^2(k)$ used for the non-zero units



Note: Within each panel, the number in the upper right corner denotes the skewness for the random variable $Z^* = YX / k$, where Y denotes an independent Bernoulli random variable with parameter p.

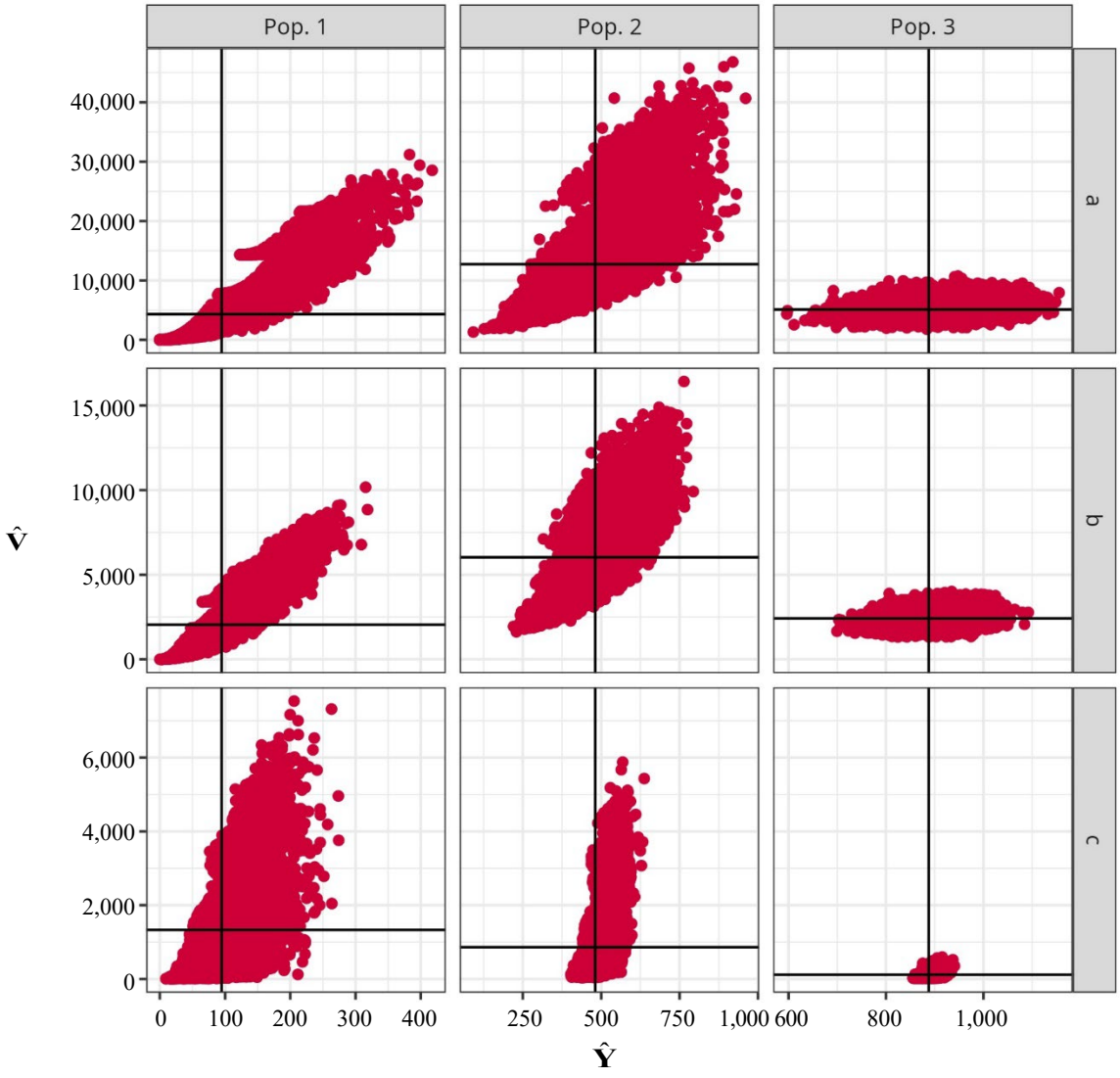
Three different sampling designs a, b, c were considered: Sampling designs a and b both used simple random sampling without replacement (SRS), with sample sizes $n_a = 50$, $n_b = 100$. For sampling design c, the population was first ordered by the target variable. The ordered population was then divided into ten equal-sized, sequential groups (deciles). From each of these deciles, a simple random sample of 10 units was drawn. The true variances of the Horvitz-Thompson (HT) estimators (Horvitz and Thompson, 1952) of the totals are shown in Table 3.1, and the relationship between the HT-estimates and the corresponding variance estimates is shown in Figure 3.2.

Table 3.1
The variances for the HT-estimators of the total for the three sampling designs a, b, c for each of the populations in Figure 3.1

Pop.	a	b	c
1	4,321.05	2,046.81	1,331.19
	(69.39%)	(47.76%)	(38.52%)
2	12,730.76	6,030.36	859.50
	(23.46%)	(16.15%)	(6.10%)
3	5,101.60	2,416.55	116.56
	(8.05%)	(5.54%)	(1.22%)

Notes: - Inside parentheses, the relative standard deviations.
- Horvitz-Thompson (HT).

Figure 3.2 Relation between estimates (x-axis) and corresponding variance estimates (y-axis) for the three populations (columns) and sampling designs (rows), for 20,000 samples each



Note: Black lines shows population totals and true variances for the estimators.

Four different combinations were considered: the optimal combination, using the true variances (V) of the estimators; using the sample size of the design (n); the naïve combination, using variance estimates $(\hat{V})$; and the proposed approach, using ratio variance estimates $(\hat{V}_R)$. The results of drawing 20,000 samples per sampling design and population, and combining these, are shown in Tables 3.2, 3.3 and 3.4, for relative RMSE, standard deviation and bias, respectively. Note that for Population 1, about 0.4% of the linear combinations could not be constructed, due to the sample a yielding a zero variance estimate. As such, these were left out of the results entirely. The likely effect of dropping the zeros is that the absolute bias is slightly underestimated for Pop. 1. How to best deal with zero estimates when using linear combinations based on data is currently an unsolved problem not considered in this paper.

Table 3.2
Relative Root Mean Squared Error (RMSE) for combining estimates using four different approaches: optimal (V), sample size (n), naïve $(\hat{V})$ and the proposed ratio $(\hat{V}_R)$. Results per combination in columns for sampling designs (a) SRS $n = 50$, (b) SRS $n = 100$ and (c) stratified $n = 10 \times 10$

Pop.	Type	ab	ac	bc	abc
1	V	39.13%	33.81%	30.18%	27.59%
	n	39.14%	34.66%	30.85%	28.19%
	$\hat{V}$	48.88%	44.14%	36.02%	43.12%
	$\hat{V}_R$	44.44%	39.21%	33.66%	34.90%
2	V	13.32%	5.92%	5.76%	5.57%
	n	13.32%	8.81%	8.73%	8.36%
	$\hat{V}$	14.10%	5.91%	5.77%	5.65%
	$\hat{V}_R$	13.47%	5.85%	5.72%	5.52%
3	V	4.56%	1.21%	1.19%	1.18%
	n	4.56%	2.79%	2.85%	2.78%
	$\hat{V}$	4.60%	1.20%	1.19%	1.18%
	$\hat{V}_R$	4.58%	1.21%	1.20%	1.18%

Note: Simple random sampling without replacement (SRS).

Table 3.3
Relative Standard Deviation for combining estimates using four different approaches: optimal (V), sample size (n), naïve $(\hat{V})$ and the proposed ratio $(\hat{V}_R)$. Results per combination in columns for sampling designs (a) SRS $n = 50$, (b) SRS $n = 100$ and (c) stratified $n = 10 \times 10$

Pop.	Type	ab	ac	bc	abc
1	V	39.13%	33.81%	30.18%	27.59%
	n	39.14%	34.66%	30.85%	28.19%
	$\hat{V}$	43.87%	39.99%	33.88%	34.99%
	$\hat{V}_R$	42.83%	37.93%	33.14%	32.47%
2	V	13.32%	5.92%	5.76%	5.57%
	n	13.32%	8.81%	8.73%	8.36%
	$\hat{V}$	13.83%	5.86%	5.71%	5.49%
	$\hat{V}_R$	13.44%	5.84%	5.70%	5.48%
3	V	4.56%	1.21%	1.19%	1.18%
	n	4.56%	2.79%	2.85%	2.78%
	$\hat{V}$	4.59%	1.20%	1.19%	1.18%
	$\hat{V}_R$	4.58%	1.21%	1.20%	1.18%

Note: Simple random sampling without replacement (SRS).

Table 3.4
Relative Bias for combining estimates using four different approaches: optimal (V), sample size (n), naïve ($\hat{V}$) and the proposed ratio ($\hat{V}_R$). Results per combination in columns for sampling designs (a) SRS $n = 50$, (b) SRS $n = 100$ and (c) stratified $n = 10 \times 10$

Pop.	Type	ab	ac	bc	abc
1	V	0.16%	0.05%	-0.04%	0.04%
	n	0.16%	0.10%	-0.03%	0.06%
	$\hat{V}$	-21.57%	-18.70%	-12.23%	-25.19%
	$\hat{V}_R$	-11.87%	-9.92%	-5.88%	-12.78%
2	V	0.02%	0.02%	0.01%	0.01%
	n	0.02%	0.03%	0.01%	0.02%
	$\hat{V}$	-2.75%	-0.79%	-0.79%	-1.36%
	$\hat{V}_R$	-0.99%	-0.38%	-0.48%	-0.73%
3	V	0.01%	0.01%	0.00%	0.00%
	n	0.01%	0.02%	0.00%	0.01%
	$\hat{V}$	-0.15%	-0.02%	-0.03%	-0.04%
	$\hat{V}_R$	0.06%	0.00%	-0.01%	-0.02%

Note: Simple random sampling without replacement (SRS).

As theorized, it is possible to reduce bias by using the ratio estimator. Even for populations which almost do not experience any skewness at all, the ratio-weighted estimator is able to reduce the bias or can at least be shown to not have a detrimental effect on the bias, SE and MSE compared with the naïve approach. However, as the linearity assumption behind the ratio estimator is not exactly fulfilled, as hinted by Figure 3.1, the proposed method is not able to completely remove the bias.

In population 1, the sample size weighting is better than the proposed ratio weighting, which is better than the naïve weighting. For populations 2 and 3 the proposed ratio weighting performs better than sample size weighting. In many applications, sample size weighting is not feasible as sample sizes are not available or cannot be determined. The results in Table 3.3 also highlights the problem with using the sample size as the weight for linear combinations, aside from the sample size not always being available in certain type of surveys, mainly environmental surveys where the observational units are indirectly sampled through sample plots. Sample size weighting is detrimentally affected when the efficiencies of the surveys vary considerably. When combining sampling designs with different efficiencies (design effects), here exemplified by the not-so-efficient SRS b $(n = 100)$ and the highly efficient ordered stratified design c $(n = 100)$, the true variance of the estimator is no longer proportional to the sample size, and can therefore be quite far away from the optimal solution. This leads to the proposed ratio weighting approach having better performance in populations 2 and 3, where it has little bias.

4. Final remarks

We introduced a novel approach for determining the weights in a linear combination of multiple estimators when dealing with non-negative and positively skewed target variables. We based the methodology on the fact that the estimators, under such conditions, exhibit a near-proportional relationship with their

variance estimators. Then we proved that, under such conditions, the proposed ratio-weighted estimator has a smaller approximate absolute bias than the standard inverse-variance estimator. Our simulation example clearly supports our theoretical findings. Moreover, the ratio-weighted estimator seems highly robust to deviations from near-proportionality. In all cases we evaluated, it performed better than the naïve combination. Thus, for a wide range of different target variables, ratio weights can be used to significantly reduce the bias of a linear combination estimator.

A major advantage is that our proposed approach does not need additional information and applies to any non-negative target variable characterized by a positively skewed distribution.

In some situations, especially when combining estimates from similar designs, deciding the weights based on the sample sizes is a viable option that can produce an optimal or close to optimal weighting. However, with very different designs, such weights can be far from optimal. Not only may the designs be very different, but also the frames that we sample from and the target variable distribution. The underlying population may be defined by a geographical area, which can be treated as a continuous frame of points. In this case, the target variable can be defined as a point response or a local average over some fixed area plot. Different plot sizes or shapes produce different responses, but yield the same population totals. An area can also be partitioned into different discrete (finite) populations that can be sampled using finite population methods, producing different frames, sampling units, and target variable distributions. In, for example, environmental monitoring, it is very common to find designs that use very different approaches to construct the frame from which we sample. However, these different designs are still used to estimate the same quantities. In such situations, sample sizes are not even close to comparable and should not be used to derive weights. For these reasons, the new weighting might be particularly advantageous for environmental surveys, where challenges such as varying designs, frames, plot sizes, and shapes complicate the identification of optimal weights. Given the prevalence of non-negative and positively skewed target variables in various surveys and meta-analysis, we anticipate that our weighting strategy will find widespread application, ultimately contributing to enhanced and more accurate combined estimates derived from multiple surveys.

Acknowledgements

The authors thank the associate editor and two anonymous referees for their helpful comments and suggestions.

Appendix

A. Proof of Theorem 1

Approximate bias of $\hat{Y}_{\hat{\alpha}}$

The bias of the combined estimator is given by $\text{Bias}(\hat{Y}_{\hat{\alpha}}) = E(\hat{Y}_{\hat{\alpha}}) - Y$. Since each estimator is unbiased for Y and the weights sum to unity, we have

$$\begin{aligned}\text{Bias}(\hat{Y}_{\hat{\alpha}}) &= E\left(\sum_{j=1}^{K}\hat{\alpha}_j\hat{Y}_j\right) - Y\sum_{j=1}^{K}E(\hat{\alpha}_j)\\ &= \sum_{j=1}^{K}E(\hat{\alpha}_j\hat{Y}_j) - \sum_{j=1}^{K}E(\hat{\alpha}_j)\,E(\hat{Y}_j) = \sum_{j=1}^{K}\text{Cov}(\hat{\alpha}_j,\hat{Y}_j).\end{aligned}$$

To derive the approximate bias, we make a first-order Taylor approximation of $\hat{\alpha}_j$ around the baseline $(V_1,\ldots,V_K)$. The approximation is

$$\hat{\alpha}_j \approx \alpha_j + \sum_{i=1}^{K}\left.\frac{\partial\hat{\alpha}_j}{\partial\hat{V}_i}\right|_{\text{baseline}}(\hat{V}_i - V_i).$$

The partial derivative of $\alpha_j = (1/V_j)\Big/\left(\sum_{i=1}^{K}1/V_i\right)$ with respect to V_i is

$$\frac{\partial\alpha_j}{\partial V_i} = \begin{cases} -\dfrac{(1/V_j^2)\left(\sum_{l\neq j}1/V_l\right)}{\left(\sum_{l=1}^{K}1/V_l\right)^2} & \text{if } i = j \\ \dfrac{(1/V_j)\,(1/V_i^2)}{\left(\sum_{l=1}^{K}1/V_l\right)^2} & \text{if } i \neq j \end{cases}$$

The covariance term becomes

$$\text{Cov}(\hat{\alpha}_j,\hat{Y}_j) \approx \text{Cov}\left(\sum_{i=1}^{K}\frac{\partial\alpha_j}{\partial V_i}(\hat{V}_i - V_i),\hat{Y}_j\right) = \sum_{i=1}^{K}\frac{\partial\alpha_j}{\partial V_i}\text{Cov}(\hat{V}_i,\hat{Y}_i).$$

With K independent surveys, the cross-covariances $\text{Cov}(\hat{V}_i,\hat{Y}_j)$ for $i \neq j$ are zero. Thus,

$$\text{Cov}(\hat{\alpha}_j,\hat{Y}_j) \approx \frac{\partial\alpha_j}{\partial V_j}\text{Cov}(\hat{V}_j,\hat{Y}_j) = -\frac{(1/V_j^2)\left(\sum_{i\neq j}1/V_i\right)}{\left(\sum_{i=1}^{K}1/V_i\right)^2}\text{Cov}(\hat{V}_j,\hat{Y}_j).$$

This expression shows that the terms contribute negatively to the bias when $\text{Cov}(\hat{V}_j,\hat{Y}_j) > 0$. Using the model (2.3), which gives $\text{Cov}(\hat{V}_j,\hat{Y}_j) = \gamma_j V(\hat{Y}_j) = \gamma_j V_j$, the total approximate bias is

$$\text{Bias}(\hat{Y}_{\hat{\alpha}}) \approx \sum_{j=1}^{K}\left(-\frac{(1/V_j^2)\left(\sum_{i\neq j}1/V_i\right)}{\left(\sum_{i=1}^{K}1/V_i\right)^2}\gamma_j V_j\right) = -\frac{\sum_{j=1}^{K}\frac{\gamma_j}{V_j}\left(\sum_{i\neq j}1/V_i\right)}{\left(\sum_{i=1}^{K}1/V_i\right)^2}.$$

Approximate bias of $\hat{Y}_{\hat{\alpha}_R}$

Define $W_j = \hat{Y}_j/\hat{V}_j$ for $j = 1,\ldots,K$. The weights are $\hat{\alpha}_{jR} = W_j\Big/\sum_{i=1}^{K}W_i$. The bias is $\text{Bias}(\hat{Y}_{\hat{\alpha}_R}) = \sum_{j=1}^{K}\text{Cov}(\hat{\alpha}_{jR},\hat{Y}_j)$.

We compute the first-order Taylor approximation of $\hat{\alpha}_{jR}$ around the baseline $(\hat{Y}_i = Y, \hat{V}_i = V_i = \gamma_i Y)$ for all i. The partial derivatives of W_i evaluated at the baseline are

$$\left.\frac{\partial W_i}{\partial \hat{Y}_i}\right|_{\text{baseline}} = \frac{1}{V_i}, \left.\frac{\partial W_i}{\partial \hat{V}_i}\right|_{\text{baseline}} = -\frac{Y}{V_i^2} = -\frac{1}{\gamma_i V_i}.$$

By the chain rule, we can express the Taylor expansion of $\hat{\alpha}_{k,R}$ as

$$\hat{\alpha}_{jR} \approx \alpha_{jR} + \sum_{i=1}^{K} \left(\left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{Y}_i}\right|_{\text{baseline}} \Delta Y_i + \left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{V}_i}\right|_{\text{baseline}} \Delta V_i \right),$$

where $\Delta Y_i = \hat{Y}_i - Y$ and $\Delta V_i = \hat{V}_i - V_i$. The partial derivatives are

$$\left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{Y}_i}\right|_{\text{baseline}} = \left.\frac{\partial \hat{\alpha}_{jR}}{\partial W_i}\right|_{\text{baseline}} \frac{1}{V_i}, \left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{V}_i}\right|_{\text{baseline}} = \left.\frac{\partial \hat{\alpha}_{jR}}{\partial W_i}\right|_{\text{baseline}} \frac{-1}{\gamma_i V_i}.$$

By the model (2.3) assumption, we have $\Delta V_i = \gamma_i \Delta Y_i + \varepsilon_i$. Substituting this into the expansion, the terms involving ΔY_i cancel out, and some algebra yields:

$$\left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{Y}_i}\right|_{\text{baseline}} \Delta Y_i + \left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{V}_i}\right|_{\text{baseline}} \Delta V_i = \left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{V}_i}\right|_{\text{baseline}} \varepsilon_i.$$

The expression for $\hat{\alpha}_{jR}$ simplifies to

$$\hat{\alpha}_{jR} \approx \alpha_{jR} + \sum_{i=1}^{K} \left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{V}_i}\right|_{\text{baseline}} \varepsilon_i.$$

As ε_i for $i = 1, \ldots, K$ are (assumed) independent of $\hat{Y}_j$, the approximate covariance term is

$$\text{Cov}(\hat{\alpha}_{jR}, \hat{Y}_j) \approx \text{Cov}\left(\sum_{i=1}^{K} \left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{V}_i}\right|_{\text{baseline}} \varepsilon_i, \hat{Y}_j \right) = \sum_{i=1}^{K} \left.\frac{\partial \hat{\alpha}_{jR}}{\partial \hat{V}_i}\right|_{\text{baseline}} \text{Cov}(\varepsilon_i, \hat{Y}_j) = 0.$$

Thus, the approximate (first-order) bias is zero.

$$\text{Bias}(\hat{Y}_{\hat{\alpha}_R}) = \sum_{j=1}^{K} \text{Cov}(\hat{\alpha}_{jR}, \hat{Y}_j) \approx 0.$$

References

Cochran, W.G. (1954). The combination of estimates from different experiments. *Biometrics*, 10(1), 101-129.

Elliott, M.R., Raghunathan, T.E. and Schenker, N. (2018). Combining estimates from multiple surveys. Wiley StatsRef: Statistics Reference Online, 1-10.

Fisher, R.A., Corbet, A.S. and Williams, C.B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *The Journal of Animal Ecology*, 42-58.

Fridman, J., Holm, S., Nilsson, M., Nilsson, P., Ringvall, A.H. and Ståhl, G. (2014). Adapting National Forest Inventories to changing requirements – The case of the Swedish National Forest Inventory at the turn of the 20th century. *Silva Fennica*, 48(3).

Grafström, A., Ekström, M., Jonsson, B.G., Esseen, P.-A. and Ståhl, G. (2019). On combining independent probability samples. *Survey Methodology*, 45(2), 349-364. Available at: https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2019002/article/00003-eng.pdf.

Gregoire, T.G. and Schabenberger, O. (1999). Sampling skewed biological populations: Behavior of confidence intervals for the population total. *Ecology*, 80(3), 1056-1065.

Higgins, J.P., White, I.R. and Anzures-Cabrera, J. (2008). Meta-analysis of skewed data: Combining results reported on log-transformed or raw scales. *Statistics in Medicine*, 27(29), 6072-6092.

Horvitz, D.G. and Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260), 663-685.

Preston, F.W. (1948). The commonness, and rarity, of species. *Ecology*, 29(3), 254-283.

Shahar, D.J. (2017). Minimizing the variance of a weighted average. *Open Journal of Statistics*, 7(2), 216-224.