Visualization of generalized mean estimators using auxiliary information in survey sampling: additive case and stratification

Rodolphe Priam

Visualization of generalized mean estimators using auxiliary information in survey sampling: additive case and stratification

R. Priam*

January 24, 2021

Abstract

The mean estimators with ratio depend on multiple auxiliary variables and unknown parameters in a finite population setting. Recently a new generic approach for modeling multivariate mean estimators with matrices has been proposed in order to compute automatically their minimum mean squared error. This brings naturally a graphical analysis for comparing mean estimators via nonlinear curves of their approximated mean squared error or their bias. Herein generalized parametric ratio estimators with two auxiliary variables are proposed for a stratified sampling design. This leads to consider per stratum the main matrix in stake with moments of the auxiliary and target variables while keeping an underlying regression model for the optimization on a new aggregated model. A brief review of the generic method is also presented while extending to constrained parameters the approach. A perspective is the visualization of alternative models under this framework when empirical means are associated with ratio functions of auxiliary variables.

Keywords: ratio estimator, auxiliary variable, mean squared error, bias, stratification

1 Introduction

In survey theory, one generally studies an aggregated stastistics non available at the population level. For this purpose it may be constructed a random variable which has its expectation the nearest as possible to the population value. In order to assist the estimation, auxiliary variables are used in order to correct the avalaible empirical statistics by additional information, and reduce the variability for instance because the value of these variables is known at the level of the population. The correction usually implements either a ratio method with a multiplicative factor (Cochran, 1940), either a regression method with an additive factor (Rao, 1991). Hence, the expressions of existing estimators from the literature, see (Allen, Singh, and Smarandache, 2003; Diana and Perri, 2007; Vishwakarma and Kumar, 2015) for instance, are in an additional, multiplicative, quotient setting. When using more than one auxiliary variables, the mean squared error may be even reduced further (Olkin, 1958), this justifies to study such estimators. A new generalizing class (Priam, 2019) of ratio estimators has been introduced via a modeling of each function in stake separately with their expansions. The expression of a corresponding mean squared error (mse) has been proposed in a matricial form and its minimisation w.r.t. the parameters. Thus the comparison and the visualization of the mse in order to help at selecting the estimator for a whole population or an available survey sample becomes possible. Herein, this approach is extended with new generalized estimators with a stratified sampling design.

*rpriam@gmail.com

- For one auxiliary variable, let denote Y the variable of interest and X an auxiliary variable which is correlated with Y. The population mean $\bar{X}$ of X is known while the population mean $\bar{Y}$ of Y is unknown in the case of a sample. The observations y_i for Y and x_i for X are available for each sampled unit: the sample $(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$ is a random variable of size n on pairs of variable (X,Y). These pairs are drawn by simple random sampling herein without loss of generality from a population of size N. Let define the sample means $\bar{x}=\frac{1}{n}\sum_i^n x_i$ and $\bar{y}=\frac{1}{n}\sum_i^n y_i$. When $f(.;.)$ is a function with one variable and eventually a vector of parameters θ, it is defined the *parametric ratio estimator* as written as follows,

$$\bar{y}_{\mathrm{R}_f} \quad = \quad f_\theta(\bar{x};\bar{X})\bar{y}. \tag{1}$$

Exemples of this estimator are found in (Yasmeen, Amin, and Hanif, 2015) for particular function f_θ. Let denote:

$$\delta_x=\frac{\bar{x}-\bar{X}}{\bar{X}} \text{ and } \delta_y=\frac{\bar{y}-\bar{Y}}{\bar{Y}},$$

where $E_s[\delta_x]=0$ and $E_s[\delta_y]=0$. Let define $a=a(\theta)$, and $b=b(\theta)$ eventually constant, taking their values according to the chosen function $f_\theta(.;.)$. An explicit 2rd order serie approximation leads to:

$$f_\theta(\bar{x};\bar{X}) \quad = \quad 1+a(\theta)\delta_x+b(\theta)\delta_x^2+\cdots. \tag{2}$$

This is mostly related to (Diana, Giordan, and Perri, 2011) with explicit variables a and b too, but the study of derivatives are found in other communications from the literature. These researches also consider second order approximations but without a and b being fully variables as proposed herein. Note that serie approximations, such as in (Srivastava, 1971) (Srivastava and Jhajj, 1981) (Koyuncu and Kadilar, 2010) (Vishwakarma and Singh, 2012) (Singh and Khalid, 2018) (Singh and Yadav, 2018), are proposed in other publications but via alternative approaches. A list of the corresponding values for a and b for several functions (from the literature) is presented in (Priam, 2019).

- In a bidimensional setting ($p=2$), the notation is as follows. Two auxiliary variables X_j are available with population mean $\bar{X}_j$ and sample mean $\bar{x}_j$ for $j=1$ and $j=2$. Hence $x_i=(x_{i1},x_{i2})$ is bidimensional. Let define $\bar{x}_j=\frac{1}{n}\sum_i^n x_{ij}$ and $\bar{y}=\frac{1}{n}\sum_i^n y_i$ and similarly $\bar{X}_j$ and $\bar{Y}$ for the population means. Let denote the free parameters α_j aggregated in the vector $\alpha=(\alpha_1,\alpha_2)^T$, such that the constraint $\sum_{j=1}^2\alpha_j=1$ may be introduced. Let denote $\delta_{x_j}=(\bar{x}_j-\bar{X}_j)/\bar{X}_j$ such as $f_{\theta_j}(.;.)$ is obtained by replacing in $f_\theta(.;.)$ θ by a new vector θ_j eventually different for each j if not equal to θ, while also replacing δ_x by δ_{x_j}. Let also denote $C_{0j}=S_{0j}/\bar{X}_j\bar{Y}$, $C_{jk}=S_{jk}/\bar{X}_j\bar{X}_k$, $C_0^2=S_0^2/\bar{Y}^2$, $C_j^2=S_j^2/\bar{X}_j^2$, $Cov_s(\bar{y},\bar{x}_j)=\lambda_n S_{0j}$, $Cov_s(\bar{x}_j,\bar{x}_k)=\lambda_n S_{jk}$, $Var_s(\bar{y})=\lambda_n S_0^2$, $Var_s(\bar{x}_j)=\lambda_n S_j^2$, and $\lambda_n=\frac{1}{n}(1-f)$ where $f=n/N$. Let also denote the correlations $\rho_{0j}=S_{0j}/S_0S_j$ and $\rho_{jk}=S_{jk}/S_jS_k$.
Bivariate ratio estimators may be defined by the combination of functions $f_\theta(.;.)$. For instance, the additive parametric ratio estimator is defined via a weighted sum of ratio estimators, while the multiplicative parametric ratio estimator is via a product of ratio estimators, roughly speaking. When denoting the new parameters,

$$a_j=a_j(\theta_j),\text{ and } b_j=b_j(\theta_j),$$

per function $f_{\theta_j}(.;.)$, let (re)define estimators for two auxiliary variables as listed in Table 2. It is recognized the parametric combinations of additive and multiplicative ratio estimators, and other parametric bivariate combined ratio and regression estimators from the statistical literature.

- We consider the case when two auxiliary variables are available and the approximations of $f_{\theta_1}(.;.)$ and $f_{\theta_2}(.;.)$ respectively, such that in order to write the amse, a generalized estimator $\bar{y}_{\mathrm{R}_{est}}$ is linearized

$\bar{y}_{\mathrm{R}_{diff}}$	$\bar{y}+k_1\delta_{x_1}+k_2\delta_{x_2}$	
$\bar{y}_{\mathrm{R}_{rao}}$	$k_0\bar{y}+k_1\delta_{x_1}+k_2\delta_{x_2}$	
$\bar{y}_{\mathrm{R}_f^m}$	$\bar{y}f_{\theta_1}(\bar{x}_1;\bar{X}_1)f_{\theta_2}(\bar{x}_2;\bar{X}_2)$	
$\bar{y}_{\mathrm{R}_f^c}$	$(k_0\bar{y}+k_1\delta_{x_1})f_{\theta}(\bar{x}_2;\bar{X}_2)$	
$\bar{y}_{\mathrm{R}_f^{c2}}$	$(k_0\bar{y}+k_1\delta_{x_1}+k_2\delta_{x_2})f_{\theta_1}(\bar{x}_1;\bar{X}_1)f_{\theta_2}(\bar{x}_2;\bar{X}_2)$	
$\bar{y}_{\mathrm{R}_f^{c3}}$	$k_0\bar{y}f_{\theta_1}(\bar{x}_1;\bar{X}_1)f_{\theta_2}(\bar{x}_2;\bar{X}_2)+k_1\delta_{x_1}+k_2\delta_{x_2}$	(3)
$\bar{y}_{\mathrm{R}_f^{c4}}$	$k_0\bar{y}+k_1\delta_{x_1}f_{\theta_2}(\bar{x}_2;\bar{X}_2)+k_2\delta_{x_2}f_{\theta_1}(\bar{x}_1;\bar{X}_1)$	
$\bar{y}_{\mathrm{R}_f^a}$	$(k_1f_{\theta_1}(\bar{x}_1;\bar{X}_1)+k_2f_{\theta_2}(\bar{x}_2;\bar{X}_2))\,\bar{y}$	
$\bar{y}_{\mathrm{R}_f^{am}}$	$\left(k_0+k_1f_{\theta_j}(\bar{x}_1;\bar{X}_1)f_{\theta_2}(\bar{x}_2;\bar{X}_2)\right)\bar{y}$	
$\bar{y}_{\mathrm{R}_f^{am2}}$	$\left(k_1f_{\theta_1}(\bar{x}_1;\bar{X}_1)+k_2f_{\theta_1}^{-1}(\bar{x}_1;\bar{X}_1)\right)f_{\theta_2}(\bar{x}_2;\bar{X}_2)\bar{y}$	
$\bar{y}_{\mathrm{R}_f^{am3}}$	$\left(k_1f_{\theta_1}(\bar{x}_1;\bar{X}_1)f_{\theta_2}(\bar{x}_2;\bar{X}_2)+k_2f_{\theta_1}^{-1}(\bar{x}_1;\bar{X}_1)f_{\theta_2}^{-1}(\bar{x}_2;\bar{X}_2)\right)\bar{y}$	

Table 1: List of example of bivariates generalized ratio estimators.

with an expansion w.r.t. the random variables in stake as follows:

$$\bar{y}_{\mathrm{R}_{est}}-\bar{Y} \;\doteq\; T+U_0\delta_y+\sum_{j=1}^{2}U_j\delta_{x_j}+\sum_{j=1}^{2}V_{jj}\delta_{x_j}^2+\sum_{j=1}^{2}V_{0j}\delta_y\delta_{x_j}+V_{12}\delta_{x_1}\delta_{x_2}\,. \tag{4}$$

The expressions of the *hidden parameters* T, U_0, U_j, V_{jj}, V_{0j}, V_{jk} and their corresponding values depend on each estimator in stake.

Herein we are interested in comparing several estimators in a stratified sampling design, hence the previously proposed expression of the amse is extended. The plan is as follows. Section 2 is a brief review of the approach for computing and visualizing the amse with an extension to constraints on the parameters. Section 3 develops the additive case with the amse, the matrices involved and an alternative expression. Section 4 proposes an extension to a stratification with examples in order to deal with separate estimators via summations. Section 5 is for the experiments which lead to compare several selected estimators and find the best one, at the first order. Section 6 is the conclusion which discusses the contribution and several perpectives.

2 Review of a generalized matricial expression of the amse and its minimum

In this section, expressions for approximating the bias and the mean squared error are found via an expansion as explained below.

Approximation of the bias and the mse via a second order serie

The expectation of the expression for $\bar{y}_{\mathrm{R}_{est}}-\bar{Y}$ just before in (4) leads to the bias of the generalized estimator:

$$\mathrm{bias}\,[\bar{y}_{\mathrm{R}_{est}}] \;\doteq\; T+V_{11}\lambda_nC_1^2+V_{22}\lambda_nC_2^2+V_{01}\lambda_nC_{01}+V_{02}\lambda_nC_{02}+V_{12}\lambda_nC_{12}\,. \tag{5}$$

The expectation of the squared may lead to the approximated mean squared-error,

$$(\bar{y}_{\mathrm{R}_{est}}-\bar{Y})^2 \;\doteq\; \big[T+U_0\delta_y+\sum_{j=1}^{2}U_j\delta_{x_j}\big]^2+2T\big[\sum_{j=1}^{2}V_{jj}\delta_{x_j}^2+\sum_{j=1}^{2}V_{0j}\delta_y\delta_{x_j}+V_{12}\delta_{x_1}\delta_{x_2}\big]\,. \tag{6}$$

Such that:

$$\begin{aligned}\mathrm{mse}\,[\bar{y}_{\mathrm{R}_{est}}] &\doteq \mathrm{amse}\,[\bar{y}_{\mathrm{R}_{est}}]\\ &= T^2+\lambda_nU_0^2C_0^2+2\lambda_n(TV_{11}+0.5U_1^2)C_1^2\\ &\quad+2\lambda_n(TV_{22}+0.5U_2^2)C_2^2+2\lambda_n(TV_{01}+U_0U_1)C_{01}\\ &\quad+2\lambda_n(TV_{02}+U_0U_2)C_{02}+2\lambda_n(TV_{12}+U_1U_2)C_{12}\,.\end{aligned} \tag{7}$$

Approximation of the bias and the mse via a linearization

The random variable $\bar{y}_{\mathrm{R}_{est}}$ is approximatively a function of $\delta = (\delta_y, \delta_{x_1}, \delta_{x_2})^T$,

$$\bar{y}_{\mathrm{R}_{est}} \doteq Z(\delta)\,.$$

A first order approximation associated to a linearization (Wolter, 2007) leads to an expression for the mean squared error as follows,

$$\begin{aligned}
\mathrm{mse}_L\,[\bar{y}_{\mathrm{R}_{est}}] &\doteq \mathrm{amse}_L\,[\bar{y}_{\mathrm{R}_{est}}] \\
\mathrm{amse}_L\,[\bar{y}_{\mathrm{R}_{est}}] &= \tfrac{\partial Z}{\partial \delta^T} Var[\delta] \tfrac{\partial Z}{\partial \delta} + \mathrm{bias}^2\,[\bar{y}_{\mathrm{R}_{est}}] \\
&= \lambda_n \big(U_0^2 C_0^2 + 2\sum_j U_0 U_j C_{0j} + U_1 U_2 C_{12}\big) + \big(T + \sum_j V_{jj} \lambda_n C_j^2 + \sum_j V_{0j} \lambda_n C_{0j} + V_{12} \lambda_n C_{12}\big)^2 \\
&= \lambda_n \big(U_0^2 C_0^2 + 2\sum_j U_0 U_j C_{0j} + U_1 U_2 C_{12}\big) + T^2 + 2T\,\big(\sum_j V_{jj} \lambda_n C_j^2 + \sum_j V_{0j} \lambda_n C_{0j} + V_{12} \lambda_n C_{12}\big) \\
&\quad + \big(\sum_j V_{jj} \lambda_n C_j^2 + \sum_j V_{0j} \lambda_n C_{0j} + V_{12} \lambda_n C_{12}\big)^2 .
\end{aligned} \tag{8}$$

The positivity of the variance and the squared may induce the positivity of this amse. This new approximation of the mse may add several terms to the usual one in order to insure non negative values, hence it may be larger than the usual one:

$$\begin{aligned}
\mathrm{amse}_L\,[\bar{y}_{\mathrm{R}_{est}}] &= \mathrm{amse}\,[\bar{y}_{\mathrm{R}_{est}}] + \mathrm{bmse}_L\,[\bar{y}_{\mathrm{R}_{est}}] \\
\mathrm{bmse}_L\,[\bar{y}_{\mathrm{R}_{est}}] &= \lambda_n^2 \big(V_{11} C_1^2 + V_{22} C_2^2 + V_{01} C_{01} + V_{02} C_{02} + V_{12} C_{12}\big)^2 \\
&= \big(\mathrm{bias}\,[\bar{y}_{\mathrm{R}_{est}}] - T\big)^2 .
\end{aligned} \tag{9}$$

Note that an higher order for the taylor approximation for the variance and the bias is an appealing extension as a future perspective. Next after, the amse is rewritten with the help of a matrix for an automatic computation of its minimum with real data.

Matricial expression of the amse

For a matricial expression of the amse, it may be considered the vector Ψ and the matrix Q,

$$\Psi = \begin{pmatrix} T \\ U_0 \\ U_1 \\ U_2 \\ V_{11} \\ V_{22} \\ V_{01} \\ V_{02} \\ V_{12} \end{pmatrix}, Q = \begin{pmatrix}
1 & 0 & 0 & 0 & \lambda_n C_1^2 & \lambda_n C_2^2 & \lambda_n C_{01} & \lambda_n C_{02} & \lambda_n C_{12} \\
0 & \lambda_n C_0^2 & \lambda_n C_{01} & \lambda_n C_{02} & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda_n C_{01} & \lambda_n C_1^2 & \lambda_n C_{12} & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda_n C_{02} & \lambda_n C_{12} & \lambda_n C_2^2 & 0 & 0 & 0 & 0 & 0 \\
\lambda_n C_1^2 & 0 & 0 & 0 & \lambda_n^2 C_1^4 & \lambda_n^2 C_1^2 C_2^2 & \lambda_n^2 C_{01} C_1^2 & \lambda_n^2 C_{02} C_1^2 & \lambda_n^2 C_1^2 C_{12} \\
\lambda_n C_2^2 & 0 & 0 & 0 & \lambda_n^2 C_1^2 C_2^2 & \lambda_n^2 C_2^4 & \lambda_n^2 C_{01} C_2^2 & \lambda_n^2 C_{02} C_2^2 & \lambda_n^2 C_{12} C_2^2 \\
\lambda_n C_{01} & 0 & 0 & 0 & \lambda_n^2 C_{01} C_1^2 & \lambda_n^2 C_{01} C_2^2 & \lambda_n^2 C_{01}^2 & \lambda_n^2 C_{01} C_{02} & \lambda_n^2 C_{01} C_{12} \\
\lambda_n C_{02} & 0 & 0 & 0 & \lambda_n^2 C_{02} C_1^2 & \lambda_n^2 C_{02} C_2^2 & \lambda_n^2 C_{01} C_{02} & \lambda_n^2 C_{02}^2 & \lambda_n^2 C_{02} C_{12} \\
\lambda_n C_{12} & 0 & 0 & 0 & \lambda_n^2 C_1^2 C_{12} & \lambda_n^2 C_{12} C_2^2 & \lambda_n^2 C_{01} C_{12} & \lambda_n^2 C_{02} C_{12} & \lambda_n^2 C_{12}^2
\end{pmatrix}. \tag{10}$$

The amse is rewritten via a quadratic form with matrices in order to solve for the unknown parameters without requiring to identify any linear system. Let denote $\tilde{K} = (1, K)^T$ with for instance $K = (k_0, k_1, k_2)^T$ or $K = (k_0, k_1)$ or $K = (k_0)$ for the unknown parameters. Let also define via two blocks ξ_0 and ξ_K, the matrix $\Phi = [\xi_0 | \xi_K]$ for mapping the vector $\tilde{K}$ into the parameters Ψ such that $\Psi = \Phi \tilde{K}$. With these matrices defined just before, the approximated mean squared error amse_L is now (renamed and) written:

$$\begin{aligned}
\mathrm{amse}_M\,[\bar{y}_{\mathrm{R}_{est}}] &= \Psi^T Q \Psi \\
&= \tilde{K}^T \begin{bmatrix} \xi_0^T \\ \xi_K^T \end{bmatrix} Q [\xi_0 | \xi_K] \tilde{K} \\
&= \tilde{K}^T \begin{bmatrix} \xi_0^T Q \xi_0 & \xi_0^T Q \xi_K \\ \xi_K^T Q \xi_0 & \xi_K^T Q \xi_K \end{bmatrix} \tilde{K} \\
&= \xi_0^T Q \xi_0 + 2 \xi_0^T Q \xi_K K + K^T \xi_K^T Q \xi_K K\,.
\end{aligned} \tag{11}$$

To our knowledge, this is a new alternative approach to the usual solutions for writing the amse for ratio estimator. The matricial expression is enough flexible for several estimators without requiring multivariate derivatives.

Equivalent matricial expression of the amse

Let denote $\Upsilon = \Upsilon(\delta)$, hence we get:

$$\begin{aligned} Z(\delta) &= \Upsilon^T \Psi \\ &= \left(1, \delta_y, \delta_{x_1}, \delta_{x_2}, \delta_{x_1}^2, \delta_{x_2}^2, \delta_y\delta_{x_1}, \delta_y\delta_{x_2}, \delta_{x_2}\delta_{x_1} \right) \Psi . \end{aligned} \tag{12}$$

Then instead of writing the linearization as in (8), the variance is directly written here in a full form as follows,

$$\begin{aligned} \text{amse}_\Omega \left[\bar{y}_{\mathrm{R}_{est}}\right] &= Var[Z(\delta)] + E\left[Z(\delta)\right]^2 \\ &\doteq \Psi^T \left(\Omega + \alpha_\Upsilon \alpha_\Upsilon^T\right) \Psi \\ \\ \text{bias}_\alpha \left[\bar{y}_{\mathrm{R}_{est}}\right] &\doteq \Psi^T E[\Upsilon] \\ &= \Psi^T \alpha_\Upsilon . \end{aligned} \tag{13}$$

A justification for a sparse matrix in the analytical expression of Ω is as follows. Let consider that $\Upsilon(\delta)$ is a vector of the components $\Upsilon_k(\delta)$ for $1 \le k \le 9$ as univariate functions of δ. Thus let denote $D\Upsilon(0)$ the matrix of derivatives of $\Upsilon(\delta)$ at $\delta = 0_3 = (0,0,0)^T$. The linear approximation of Υ at $\delta = 0_3$ leads to $\Upsilon(\delta) \doteq \Upsilon(0_3) + D\Upsilon(0_3)\delta$, which induces that Ω may be written:

$$\begin{aligned} Var\left[\Upsilon(\delta)\right] &\doteq D\Upsilon(0_3) Var[\delta] D\Upsilon(0_3)^T \\ &\doteq \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_n C_0^2 & \lambda_n C_{01} & \lambda_n C_{02} \\ \lambda_n C_{01} & \lambda_n C_1^2 & \lambda_n C_{12} \\ \lambda_n C_{02} & \lambda_n C_{12} & \lambda_n C_2^2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ &\doteq \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_n C_0^2 & \lambda_n C_{01} & \lambda_n C_{02} & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_n C_{01} & \lambda_n C_1^2 & \lambda_n C_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_n C_{02} & \lambda_n C_{12} & \lambda_n C_2^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \tag{14}$$

While,

$$\alpha_\Upsilon \alpha_\Upsilon^T = \begin{pmatrix} 1 & 0 & 0 & 0 & \lambda_n C_1^2 & \lambda_n C_2^2 & \lambda_n C_{01} & \lambda_n C_{02} & \lambda_n C_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda_n C_1^2 & 0 & 0 & 0 & \lambda_n^2 C_1^4 & \lambda_n^2 C_1^2 C_2^2 & \lambda_n^2 C_{01} C_1^2 & \lambda_n^2 C_{02} C_1^2 & \lambda_n^2 C_1^2 C_{12} \\ \lambda_n C_2^2 & 0 & 0 & 0 & \lambda_n^2 C_1^2 C_2^2 & \lambda_n^2 C_2^4 & \lambda_n^2 C_{01} C_2^2 & \lambda_n^2 C_{02} C_2^2 & \lambda_n^2 C_{12} C_2^2 \\ \lambda_n C_{01} & 0 & 0 & 0 & \lambda_n^2 C_{01} C_1^2 & \lambda_n^2 C_{01} C_2^2 & \lambda_n^2 C_{01}^2 & \lambda_n^2 C_{01} C_{02} & \lambda_n^2 C_{01} C_{12} \\ \lambda_n C_{02} & 0 & 0 & 0 & \lambda_n^2 C_{02} C_1^2 & \lambda_n^2 C_{02} C_2^2 & \lambda_n^2 C_{01} C_{02} & \lambda_n^2 C_{02}^2 & \lambda_n^2 C_{02} C_{12} \\ \lambda_n C_{12} & 0 & 0 & 0 & \lambda_n^2 C_1^2 C_{12} & \lambda_n^2 C_{12} C_2^2 & \lambda_n^2 C_{01} C_{12} & \lambda_n^2 C_{02} C_{12} & \lambda_n^2 C_{12}^2 \end{pmatrix} . \tag{15}$$

Hence, exactly the same expression for the approximated mean squared error amse_L, the optimal unconstrained or constrained parameter vector K and its corresponding minimum are available with this alternative approach, by changing Q into its new expression, such that:

$$\text{amse}_\Omega \left[\bar{y}_{\mathrm{R}_{est}}\right] = \xi_0^T \left(\Omega + \alpha_\Upsilon \alpha_\Upsilon^T\right) \xi_0 + 2\xi_0^T \left(\Omega + \alpha_\Upsilon \alpha_\Upsilon^T\right) \xi_K K + K^T \xi_K^T \left(\Omega + \alpha_\Upsilon \alpha_\Upsilon^T\right) \xi_K K . \tag{16}$$

This makes possible to handle higher order terms directly from the obtained matrix instead of neglecting them automatically in the usual linearization method in one step: this can be seen as two consecutive linearizations. Additional variables are also handled directly with these matrices when the vectors Ψ and α_Υ^T are writen accordingly.

Optimal and constrained solutions

After derivation, the system to solve for identifying the optimal value of K is as follows:

$$\frac{\partial}{\partial K}\{\text{amse}_M\,[\bar{y}_{\text{R}_{est}}]\} \quad = \quad 0\,. \tag{17}$$

This leads to the regression problem, with optimal solution $K_{(opt)}$,

$$\xi_K^T Q \xi_0 + \xi_K^T Q \xi_K K_{(opt)} \quad = \quad 0\,. \tag{18}$$

This induces that the quantity which minimizes the amse is defined as follows:

$$\begin{aligned} K_{(opt)} &= argmin_K \;\text{amse}_M\,[\bar{y}_{\text{R}_{est}}] \\ &= -(\xi_K^T Q \xi_K)^{-1} \xi_K^T Q \xi_0\,. \end{aligned} \tag{19}$$

When the solution $K_{(opt)}$ is inserted in the expression of the approximated mse, the amse is minimum. Note that for the constraints such as the sum of the components of K is one, we need to add a Lagrangian which leads to update this unconstrained solution $K_{(opt)}$ as in the usual linear regression with constraints. The constraint is typically the sum to one for an additive estimator. More generally let suppose $\Gamma K = \gamma$. The constrained solution is written as for regression:

$$K^c_{(opt)} = K_{(opt)} + (\xi_K^T Q \xi_K)^{-1} \Gamma^T [\Gamma(\xi_K^T Q \xi_K)^{-1} \Gamma^T]^{-1} (\gamma - \Gamma K_{(opt)})\,. \tag{20}$$

For instance $\gamma = 1$ while Γ is a vector of 1 for the estimator $\bar{y}_{\text{R}^a_f}$ in order to insure $k_0 = 1$ in the case of only one parameter for K, otherwise it is required to add zero(s).

Minimum amse

When the unconstrained solution $K_{(opt)}$ is inserted in the expression of the approximated mse, the resulting amse is minimum:

$$\text{amse}_{M(min)}\,[\bar{y}_{\text{R}_{est}}] \quad = \quad \xi_0^T Q \xi_0 - \xi_0^T Q \xi_K (\xi_K^T Q \xi_K)^{-1} \xi_K^T Q \xi_0\,. \tag{21}$$

The corresponding bias denoted $\text{bias}_{(min)}\,[\bar{y}_{\text{R}_{est}}]$ is written afterwards when the quantities T, V_{01}, V_{02}, V_{11}, V_{22} and V_{12} from the optimal vector $\Psi_{(opt)} = \Phi(1, K^T_{(opt)})^T$, are replaced in (36), otherwise this is $\Psi^T_{(opt)} \alpha_\Upsilon$ as seen above in (13).

3 Several computations of the amse for the additive estimator

In the additive case, several approaches are possible for writing and minimizing an approximated mean squared error as explained in this section.

Additive parametric ratio estimator via direct approach

When computing the amse, an usual way is to directly develop the square. This is a particular case of the proposed generalized expression presented in the previous section. The mse for this estimator is written as a function of k_1 and k_2, a_1, a_2, b_1 and b_2 by taking the expectation of the expansion:

$$\begin{aligned} \bar{y}_{\text{R}^a_f} - \bar{Y} &\doteq (k_1 + k_2 - 1)\bar{Y} + (k_1 + k_2)\bar{Y}\delta_y + k_1 a_1 \bar{Y}\delta_{x_1} + k_2 a_2 \bar{Y}\delta_{x_2} \\ &\quad + k_1 b_1 \bar{Y}\delta^2_{x_1} + k_2 b_2 \bar{Y}\delta^2_{x_2} + k_1 a_1 \bar{Y}\delta_{x_1}\delta_y + k_2 a_2 \bar{Y}\delta_{x_2}\delta_y\,. \\ \text{mse}_{(k_1,k_2)}\,[\bar{y}_{\text{R}_a}] &\doteq A k_0^2 + B k_1^2 + 2C k_0 k_1 - 2D_0 k_0 - 2D_1 k_1 + E \\ &= [\,k_0\;k_1\,] \begin{bmatrix} A & C \\ C & B \end{bmatrix} \begin{bmatrix} k_0 \\ k_1 \end{bmatrix} - 2\,[\,D_0\;D_1\,] \begin{bmatrix} k_0 \\ k_1 \end{bmatrix} + E\,. \end{aligned} \tag{22}$$

Where, the additional terms from the linearization are,

$$\begin{aligned} A &= \bar{Y}^2 + \bar{Y}^2 \lambda_n (C_0^2 + (a_1^2 + 2b_1) C_1^2 + 4a_1 C_{01}) + \bar{Y}^2 \lambda_n^2 (a_1 C_{01} + b_1 C_1^2)^2 \\ B &= \bar{Y}^2 + \bar{Y}^2 \lambda_n (C_0^2 + (a_2^2 + 2b_2) C_2^2 + 4a_2 C_{02}) + \bar{Y}^2 \lambda_n^2 (a_2 C_{02} + b_2 C_2^2)^2 \\ C &= \bar{Y}^2 + \bar{Y}^2 \lambda_n (C_0^2 + b_1 C_1^2 + b_2 C_2^2 + 2a_1 C_{01} + 2a_2 C_{02} + a_1 a_2 C12) \\ &\quad + \bar{Y}^2 \lambda_n^2 (b_1 b_2 C_1^2 C_2^2 + a_1 b_2 C_{01} C_2^2 + a_2 b_1 C_{02} C_1^2 + a_1 a_2 C_{01} C_{02}) \\ D_0 &= \bar{Y}^2 + \bar{Y}^2 \lambda_n (b_1 C_1^2 + a_1 C_{01}) \\ D_1 &= \bar{Y}^2 + \bar{Y}^2 \lambda_n (b_2 C_2^2 + a_2 C_{02}) \\ E &= \bar{Y}^2\,. \end{aligned}$$

The mean squared error at the first order is also written as proposed in the formula (38) when it is denoted $T = (k_1+k_2-1)\bar{Y}$, $U_0 = (k_1+k_2)\bar{Y}$, $U_1 = k_1a_1\bar{Y}$, $U_2 = k_2a_2\bar{Y}$, $V_{11} = k_1b_1\bar{Y}$, $V_{22} = k_2b_2\bar{Y}$, $V_{01} = k_1a_1\bar{Y}$, $V_{02} = k_2a_2\bar{Y}$, and $V_{12} = 0$. There are two cases to solve for identifying the optimal values of k_1 and k_2. If $k_1+k_2 \neq 1$, in order to find the optimal values of k_1 and k_2 by minimizing the mse, the system coming from the derivatives is solved. The solution if its exists is given by:

$$\begin{aligned} k_{1;\text{opt}} &= \left\{BD_0 - CD_1\right\}\left\{AB - C^2\right\}^{-1} \\ k_{2;\text{opt}} &= \left\{AD_1 - CD_0\right\}\left\{AB - C^2\right\}^{-1}. \end{aligned}$$

Thus, the minimum value for the mse is obtained when these two optimal values $k_{0;\text{opt}}$ and $k_{1;\text{opt}}$ are inserted in the mse. By adapting the formula (2.10) in (Solanki and Singh, 2015) from another class of estimators, it is obtained the minimum of the mse as:

$$\text{mse}_{(\text{min})}\left[\bar{y}_{\text{R}_f^a}\right] \doteq E - \left\{BD_0^2 - 2CD_0D_1 + AD_1^2\right\}\left\{AB - C^2\right\}^{-1}. \tag{23}$$

The resulting minimal mse for $k_1+k_2 \neq 1$ is a quotient of two polynomials of the variables a_1, a_2, b_1 and b_2. Note that the generalized expression of the amse via matrices leads to the same solution, with $\Phi_{Ra} = [\xi_0|\xi_K]$ from (Priam, 2019) in the expression (21), such that:

$$\begin{bmatrix} A & C \\ C & B \end{bmatrix} = -\xi_k^T Q \xi_0 \text{ and } \begin{bmatrix} D_0 \\ D_1 \end{bmatrix} = \xi_k^T Q \xi_k. \tag{24}$$

An expression for the amse when $k_1+k_2 = 1$ is given in (Priam, 2019), a similar result comes from $K^c_{(opt)}$ defined at (20). In a next paragraph, an alternative is derived from the literature.

Expressions for the matrix Φ for additive and combined additive cases

In this paragraph matrices eventually not given previously and involved in a first order approximation of the amse are proposed for new estimators. Respectively for $\bar{y}_{\text{R}_f^{am}}$, $\bar{y}_{\text{R}_f^{am2}}$ and $\bar{y}_{\text{R}_f^{am3}}$, the hidden parameters as summarized in the Table 2. According to the expansions at the first order of these estimators, the corresponding matrices Φ are denoted Φ_{Ram}, Φ_{Ram2} and Φ_{Ram3} from the left to the right:

$$\begin{pmatrix} -\bar{Y} & \bar{Y} & \bar{Y} \\ 0 & \bar{Y} & \bar{Y} \\ 0 & 0 & a_1\bar{Y} \\ 0 & 0 & a_2\bar{Y} \\ 0 & 0 & b_1\bar{Y} \\ 0 & 0 & b_2\bar{Y} \\ 0 & 0 & a_1\bar{Y} \\ 0 & 0 & a_2\bar{Y} \\ 0 & 0 & a_1a_2\bar{Y} \end{pmatrix}, \begin{pmatrix} -\bar{Y} & \bar{Y} & \bar{Y} \\ 0 & \bar{Y} & \bar{Y} \\ 0 & a_1\bar{Y} & -a_1\bar{Y} \\ 0 & a_2\bar{Y} & a_2\bar{Y} \\ 0 & b_1\bar{Y} & (a_1^2-b_1)\bar{Y} \\ 0 & b_2\bar{Y} & b_2\bar{Y} \\ 0 & a_1\bar{Y} & -a_1\bar{Y} \\ 0 & a_2\bar{Y} & a_2\bar{Y} \\ 0 & a_1a_2\bar{Y} & -a_1a_2\bar{Y} \end{pmatrix}, \begin{pmatrix} -\bar{Y} & \bar{Y} & \bar{Y} \\ 0 & \bar{Y} & \bar{Y} \\ 0 & a_1\bar{Y} & -a_1\bar{Y} \\ 0 & a_2\bar{Y} & -a_2\bar{Y} \\ 0 & b_1\bar{Y} & (a_1^2-b_1)\bar{Y} \\ 0 & b_2\bar{Y} & (a_2^2-b_2)\bar{Y} \\ 0 & a_1\bar{Y} & -a_1\bar{Y} \\ 0 & a_2\bar{Y} & -a_2\bar{Y} \\ 0 & a_1a_2\bar{Y} & a_1a_2\bar{Y} \end{pmatrix}. \tag{25}$$

	$\bar{y}_{\text{R}_f^a}$	$\bar{y}_{\text{R}_f^{am}}$	$\bar{y}_{\text{R}_f^{am2}}$	$\bar{y}_{\text{R}_f^{am3}}$
T	$(k_1+k_2-1)\bar{Y}$	$(k_1+k_2-1)\bar{Y}$	$(k_1+k_2-1)\bar{Y}$	$(k_1+k_2-1)\bar{Y}$
U_0	$(k_1+k_2)\bar{Y}$	$(k_1+k_2)\bar{Y}$	$(k_1+k_2)\bar{Y}$	$(k_1+k_2)\bar{Y}$
U_1	$a_1k_1\bar{Y}$	$a_1k_2\bar{Y}$	$a_1(k_1-k_2)\bar{Y}$	$a_1(k_1-k_2)\bar{Y}$
U_2	$a_2k_2\bar{Y}$	$a_2k_2\bar{Y}$	$a_2(k_2+k_1)\bar{Y}$	$a_2(k_1-k_2)\bar{Y}$
V_{11}	$b_1k_1\bar{Y}$	$b_1k_2\bar{Y}$	$[b_1k_1+(a_1^2-b_1)k_2]\bar{Y}$	$[b_1k_1+(a_1^2-b_1)k_2]\bar{Y}$
V_{22}	$b_2k_2\bar{Y}$	$b_2k_2\bar{Y}$	$b_2(k_1+k_2)\bar{Y}$	$[b_2k_1+(a_2^2-b_2)k_2]\bar{Y}$
V_{01}	$a_1k_1\bar{Y}$	$a_1k_2\bar{Y}$	$a_1(k_1-k_2)\bar{Y}$	$a_1(k_1-k_2)\bar{Y}$
V_{02}	$a_2k_2\bar{Y}$	$a_2k_2\bar{Y}$	$a_2(k_1+k_2)\bar{Y}$	$a_2(k_1-k_2)\bar{Y}$
V_{12}	0	$a_1a_2k_2\bar{Y}$	$a_1a_2(k_1-k_2)\bar{Y}$	$a_1a_2(k_1+k_2)\bar{Y}$

Table 2: Parameterization for three combined additive estimators and the additive estimator.

Additive parametric ratio estimator via a matricial approach
An estimator is defined via a weighted sum of ratio estimators with new parameters $a_j = a_j(\theta_j)$ and $b_j = b_j(\theta_j)$ per function $f_{\theta_j}(.,.)$. The bias is directly obtained by linearity of the expectation operator such as the expectation of the sum is the sum of the expectations, while the mse approximation denoted amse_O is found with additional computation,

$$\begin{aligned}\text{bias}_O\left[\bar{y}_{R_f^a}\right] &\doteq \bar{Y}\sum_{j=1}^{2} k_j\left(a_j Cov_s(\delta_{x_j},\delta_y) + b_j Var_s(\delta_{x_j})\right) &&\doteq \bar{Y}K^T b_\theta \\ \text{var}_O\left[\bar{y}_{R_f^a}\right] &\doteq \bar{Y}^2\sum_{j_1,j_2=1}^{2} k_{j_1}k_{j_2} Cov_s(\delta_y + a_{j_1}\delta_{x_{j_1}}, \delta_y + a_{j_2}\delta_{x_{j_2}}) &&\doteq \bar{Y}^2 K^T A_\theta K\,.\end{aligned}$$

Hence,

$$\text{amse}_O\left[\bar{y}_{R_f^a}\right] = \bar{Y}^2 K^T b_\theta b_\theta^T K + \bar{Y}^2 K^T A_\theta K\,.$$

The matrix A_θ and vector b_θ are with cells,

$$\begin{aligned}a_{j_1 j_2;\theta} &= \frac{\lambda_n S_0^2}{\bar{Y}^2} + \frac{a_{j_1}\lambda_n S_{01}}{\bar{X}_{j_1}\bar{Y}} + \frac{a_{j_2}\lambda_n S_{02}}{\bar{X}_{j_2}\bar{Y}} + \frac{a_{j_1}a_{j_2}\lambda_n S_{jk}}{\bar{X}_{j_1}\bar{X}_{j_2}}\,. \\ b_{j;\theta} &= a_j\frac{\lambda_n S_{01}}{\bar{X}_{j_1}\bar{Y}} + b_j\frac{\lambda_n S_j^2}{\bar{X}_j^2}\,.\end{aligned}$$

Note that in comparison to the general expression from (Olkin, 1958) several terms in $O(n^{-2})$ are neglected, which can be introduced as a perspective. When $k_1 + k_2 = 1$, the optimization for K in closed form leads to:

$$\hat{K} = (e^T(A_\theta + b_\theta b_\theta^T)^{-1}e)^{-1}(A_\theta + b_\theta b_\theta^T)^{-1}e\,,$$

according to (Olkin, 1958) (formula 3.1, page 157) where $e = (1,1)^T$ is p-dimensional. Here instead of mininizing the variance it is directly minimized the approximated mean squared error. The minimal amse with the optimal weights may be written as follows:

$$\text{amse}_{O(\min)}\left[\bar{y}_{R_f^a}\right] = \frac{1}{e^T(A_\theta + b_\theta b_\theta^T)^{-1}e}\bar{Y}^2\,.$$

This way to compute the mean squared error shares some similarities with the generic approach proposed for generalized estimators with a quadratic form involved but the underlying objective function may be mostly suitable only for related to additive estimators.

4 Extension: second order estimator

For more precise mse, we consider the case when two auxiliary variables are available and also when third order terms in the approximations of $f_{\theta_1}(.;.)$ and $f_{\theta_2}(.;.)$ respectively with c_1 and c_2 as the coefficients associated to $\delta_{x_1}^3$ and $\delta_{x_2}^3$, such that a generalized estimator $\bar{y}_{R_{est}}$ is as follows:

$$\begin{aligned}\bar{y}_{R_{est}} - \bar{Y} \doteq\; & T + U_0\delta_y + \sum_j U_j\delta_{x_j} + \sum_j V_{jj}\delta_{x_j}^2 + \sum_j V_{0j}\delta_y\delta_{x_j} + \sum_{j,k;k>j} V_{jk}\delta_{x_j}\delta_{x_k} \\ & + \sum_j W_{jjj}\delta_{x_j}^3 + \sum_{j,k} W_{jkk;k\neq j}\delta_{x_j}\delta_{x_k}^2 + \sum_j W_{0jj}\delta_y\delta_{x_j}^2 + \sum_{j,k;k>j} W_{0jk}\delta_y\delta_{x_j}\delta_{x_k}\,.\end{aligned} \tag{26}$$

The expressions of the corresponding values for T, U_0, U_j, V_j, V_{0j}, V_{jk}, W_{jjj}, W_{0jj}, W_{jkk} and W_{0jk}, depend on the given estimator. The expectation of the expression just before leads to the bias of the generalized estimator. For the mean squared error it is obtained that:

$$\begin{aligned}(\bar{y}_{R_{est}} - \bar{Y})^2 &\doteq \left[T + U_0\delta_y + \sum_j U_j\delta_{x_j} + \sum_j V_{jj}\delta_{x_j}^2 + \sum_j V_{0j}\delta_y\delta_{x_j} + \sum_{j,k;k>j} V_{jk}\delta_{x_j}\delta_{x_k}\right]^2 \\ &\quad + 2T\left[\sum_j W_{jjj}\delta_{x_j}^3 + \sum_{j,k;k\neq j} W_{jkk}\delta_{x_j}\delta_{x_k}^2 + \sum_j W_{0jj}\delta_y\delta_{x_j}^2 + \sum_{j,k;k>j} W_{0jk}\delta_y\delta_{x_j}\delta_{x_k}\right] \\ &\doteq \left[T + U_0\delta_y + \sum_j U_j\delta_{x_j}\right]^2 \\ &\quad + 2\left[T + U_0\delta_y + \sum_j U_j\delta_{x_j}\right]\left[\sum_j V_{jj}\delta_{x_j}^2 + \sum_j V_{0j}\delta_y\delta_{x_j} + \sum_{j,k;k>j} V_{jk}\delta_{x_j}\delta_{x_k}\right] \\ &\quad + 2T\left[\sum_j W_{jjj}\delta_{x_j}^3 + \sum_{j,k;k\neq j} W_{jkk}\delta_{x_j}\delta_{x_k}^2 + \sum_j W_{0jj}\delta_y\delta_{x_j}^2 + \sum_{j,k;k>j} W_{0jk}\delta_y\delta_{x_j}\delta_{x_k}\right].\end{aligned}$$

The expectation w.r.t. the sampling may lead to the mean squared error in the general case which is a new result to our knowledge. Traditionally, only the squared in the first line plus the quadractic term with T as a product in the second row enter the approximated mean squared error as only the means, variances and covariances are available and such approximation is generaly enough accurate with the third order terms neglectable. By extension to the first order approximation, the random variable $\bar{y}_{\mathrm{R}_{est}}$ is a function of $\delta = (\delta_y, \delta_{x_1}, \delta_{x_2})^T$, a vector of random variables as follows:

$$\bar{y}_{\mathrm{R}_{est}} \doteq Z(\delta)\,.$$

Matricial expression

Let denote $\Upsilon = \Upsilon(\delta)$ such as the bias at (36) is written $\Psi^T \alpha_\Upsilon$ with $\alpha_\Upsilon = E[\Upsilon]$. Hence the Υ, the α_Υ and the Ω quantities are as follows,

$$\begin{aligned} Z(\delta) &= \Upsilon^T \Psi \\ &= \begin{pmatrix} 1, \delta_y, \delta_{x_1}, \delta_{x_2}, \delta^2_{x_1}, \delta^2_{x_2}, \delta_y\delta_{x_1}, \delta_y\delta_{x_2}, \delta_{x_2}\delta_{x_1}, \\ \delta^3_{x_1}, \delta^3_{x_2}, \delta_{x_1}\delta^2_{x_2}, \delta_{x_2}\delta^2_{x_1}, \delta_y\delta^2_{x_1}, \delta_y\delta^2_{x_2}, \delta_y\delta_{x_1}\delta_{x_2} \end{pmatrix} \Psi\,. \end{aligned} \tag{27}$$

And following the proposed linearization, exactly the same expression for the amse, the optimal unconstrained or constrained parameter vector K and its corresponding minimum amse are available with this alternative approach, by changing Q into a new expression with higher order terms,

$$\mathrm{amse}_\Omega\left[\bar{y}_{\mathrm{R}_{est}}\right] = \xi_0^T\left(\Omega + \alpha_\Upsilon\alpha_\Upsilon^T\right)\xi_0 + 2\xi_0^T\left(\Omega + \alpha_\Upsilon\alpha_\Upsilon^T\right)\xi_K K + K^T\xi_K^T\left(\Omega + \alpha_\Upsilon\alpha_\Upsilon^T\right)\xi_K K\,. \tag{28}$$

Here the matrices in stake are as follows, when keeping fourth order terms for the remaining rest instead of only third order terms,

$$\Omega \doteq \begin{pmatrix}
0 & 0 & 0 & 0 & E_1^2 & E_2^2 & E_{01} & E_{02} & E_{12} & E_{111} & E_{222} & E_{122} & E_{112} & E_{011} & E_{022} & E_{012} \\
0 & E_0^2 & E_{01} & E_{02} & E_{011} & E_{022} & E_{001} & E_{002} & E_{012} & E_{0111} & E_{0222} & E_{0122} & E_{0122} & E_{0011} & E_{0022} & E_{0012} \\
0 & E_{01} & E_1^2 & E_{12} & E_{111} & E_{122} & E_{011} & E_{012} & E_{112} & E_{1111} & E_{1222} & E_{1122} & E_{1112} & E_{0111} & E_{0122} & E_{0112} \\
0 & E_{02} & E_{12} & E_2^2 & E_{112} & E_{222} & E_{012} & E_{022} & E_{122} & E_{1112} & E_{2222} & E_{1222} & E_{1122} & E_{0122} & E_{0222} & E_{0122} \\
E_1^2 & E_{011} & E_{111} & E_{112} & E_{1111} & E_{1122} & E_{0111} & E_{0112} & E_{1112} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_2^2 & E_{022} & E_{122} & E_{222} & E_{1122} & E_{2222} & E_{0122} & E_{0222} & E_{1222} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{01} & E_{001} & E_{011} & E_{012} & E_{0111} & E_{0122} & E_{0011} & E_{0012} & E_{0112} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{02} & E_{002} & E_{012} & E_{022} & E_{0112} & E_{0222} & E_{0012} & E_{0022} & E_{0122} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{12} & E_{012} & E_{112} & E_{122} & E_{1112} & E_{1222} & E_{0112} & E_{0122} & E_{1122} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{111} & E_{0111} & E_{1111} & E_{1112} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{222} & E_{0222} & E_{1222} & E_{2222} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{122} & E_{0122} & E_{1122} & E_{1222} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{112} & E_{0112} & E_{1112} & E_{1122} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{011} & E_{0011} & E_{0111} & E_{0122} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{022} & E_{0022} & E_{0122} & E_{0222} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{012} & E_{0012} & E_{0112} & E_{0122} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}, \tag{29}$$

and,

$$\alpha_\Upsilon\alpha_\Upsilon^T \doteq \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & E_{11}^2 & E_{11}E_{22} & E_{01}E_{11} & E_{02}E_{11} & E_{11}E_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & E_{11}E_{22} & E_{22}^2 & E_{01}E_{22} & E_{02}E_{22} & E_{12}E_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & E_{01}E_{11} & E_{01}E_{22} & E_{01}^2 & E_{01}E_{02} & E_{01}E_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & E_{02}E_{11} & E_{02}E_{22} & E_{01}E_{02} & E_{02}^2 & E_{02}E_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & E_{11}E_{12} & E_{12}E_{22} & E_{01}E_{12} & E_{02}E_{12} & E_{12}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}. \tag{30}$$

Here E_i, E_{ij}, E_{ijk} and E_{ijkl} are the expectation of the product of random variables δ_y, δ_{x_1}, and δ_{x_2} as in (Olkin, 1958), formula 2.2a to 2.2d at page 156, such that $E_i = \lambda_n C_i$, $E_{ij} = \lambda_n C_{ij}$ as defined before while $E_{ii} = \lambda_n C_i^2$ but the case of three or four variables ask for new expressions as found originally in (Sukhatme, 1970). Hence the new matrix replacing Q is mostly filling its empty cells with missing higher order statistis from the random variables in stake. Note that the expressions of an amse leading to a limiting bound as in (Allen et al., 2003) and (Diana and Perri, 2007) may be different, hence such bound is of interest as a perspective. Here as in (Koyuncu and Kadilar, 2009a) for instance the statistics C_{ijk} and $C_{ijk\ell}$ are defined similarly than for the first and second order ones by adding respectively one and two squares in the sums for the corresponding random variable δ_y, δ_{x_1} or δ_{x_2}. This adds the terms for completing the square as expected hence the final approximated mean squared error may be larger than at the first order. Checking that the expectations of higher order terms in the expansion of $\bar{y}_{\mathrm{R}_{est}}$ in order to validate that they are neglectable might be a requirement.

Example of expressions for the matrix Φ

Let denote Ψ completed with the components $(W_{111}, W_{222}, W_{122}, W_{112}, W_{011}, W_{022}, W_{012})$for the third order, thus it is written for its estimator the hidden parameters as the components of $\Psi = \Phi_{\mathrm{R}_{est}} \tilde{K}$ as explained previously. For $\bar{y}_{\mathrm{R}_f^a}$, $\bar{y}_{\mathrm{R}_f^{am}}$, $\bar{y}_{\mathrm{R}_f^{am2}}$ and $\bar{y}_{\mathrm{R}_f^{am3}}$, it is obtained from the expansion at the third order of these estimators, respectively the corresponding matrices $\Phi = \Phi_{\mathrm{Ra}}$, $\Phi = \Phi_{\mathrm{Ram}}$, $\Phi = \Phi_{\mathrm{Ram2}}$ and $\Phi = \Phi_{\mathrm{Ram3}}$ where:

$$\Phi_{\mathrm{Ra}} = \begin{pmatrix} -\bar{Y} & \bar{Y} & \bar{Y} \\ 0 & \bar{Y} & \bar{Y} \\ 0 & a_1\bar{Y} & 0 \\ 0 & 0 & a_2\bar{Y} \\ 0 & b_1\bar{Y} & 0 \\ 0 & 0 & b_2\bar{Y} \\ 0 & a_1\bar{Y} & 0 \\ 0 & 0 & a_2\bar{Y} \\ 0 & 0 & 0 \\ 0 & c_1\bar{Y} & 0 \\ 0 & 0 & c_2\bar{Y} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & b_1\bar{Y} & 0 \\ 0 & 0 & b_2\bar{Y} \\ 0 & 0 & 0 \end{pmatrix}, \Phi_{\mathrm{Ram}} = \begin{pmatrix} -\bar{Y} & \bar{Y} & \bar{Y} \\ 0 & \bar{Y} & \bar{Y} \\ 0 & a_1\bar{Y} & 0 \\ 0 & 0 & a_2\bar{Y} \\ 0 & b_1\bar{Y} & 0 \\ 0 & 0 & b_2\bar{Y} \\ 0 & a_1\bar{Y} & 0 \\ 0 & 0 & a_2\bar{Y} \\ 0 & 0 & a_1a_2\bar{Y} \\ 0 & 0 & c_1\bar{Y} \\ 0 & 0 & c_2\bar{Y} \\ 0 & 0 & a_1b_2\bar{Y} \\ 0 & 0 & a_2b_1\bar{Y} \\ 0 & 0 & b_1\bar{Y} \\ 0 & 0 & b_2\bar{Y} \\ 0 & 0 & a_1a_2\bar{Y} \end{pmatrix} \tag{31}$$

$$\Phi_{\mathrm{Ram2}} = \begin{pmatrix} -\bar{Y} & \bar{Y} & \bar{Y} \\ 0 & \bar{Y} & \bar{Y} \\ 0 & a_1\bar{Y} & -a_1\bar{Y} \\ 0 & a_2\bar{Y} & a_2\bar{Y} \\ 0 & b_1\bar{Y} & (a_1^2-b_1)\bar{Y} \\ 0 & b_2\bar{Y} & b_2\bar{Y} \\ 0 & a_1\bar{Y} & -a_1\bar{Y} \\ 0 & a_2\bar{Y} & a_2\bar{Y} \\ 0 & a_1a_2\bar{Y} & -a_1a_2\bar{Y} \\ 0 & c_1\bar{Y} & (2a_1b_1-c_1-a_1^3)\bar{Y} \\ 0 & c_2\bar{Y} & c_2\bar{Y} \\ 0 & a_1b_2\bar{Y} & -a_1b_2\bar{Y} \\ 0 & a_2b_1\bar{Y} & a_2(a_1^2-b_1)\bar{Y} \\ 0 & b_1\bar{Y} & (a_1^2-b_1)\bar{Y} \\ 0 & b_2\bar{Y} & b_1\bar{Y} \\ 0 & a_1a_2\bar{Y} & -a_1a_2\bar{Y} \end{pmatrix}, \Phi_{\mathrm{Ram3}} = \begin{pmatrix} -\bar{Y} & \bar{Y} & \bar{Y} \\ 0 & \bar{Y} & \bar{Y} \\ 0 & a_1\bar{Y} & -a_1\bar{Y} \\ 0 & a_2\bar{Y} & -a_2\bar{Y} \\ 0 & b_1\bar{Y} & (a_1^2-b_1)\bar{Y} \\ 0 & b_2\bar{Y} & (a_2^2-b_2)\bar{Y} \\ 0 & a_1\bar{Y} & -a_1\bar{Y} \\ 0 & a_2\bar{Y} & -a_2\bar{Y} \\ 0 & a_1a_2\bar{Y} & a_1a_2\bar{Y} \\ 0 & c_1\bar{Y} & (2a_1b_1-c_1-a_1^3)\bar{Y} \\ 0 & c_2\bar{Y} & (2a_2b_2-c_2-a_2^3)\bar{Y} \\ 0 & a_1b_2\bar{Y} & a_1(b_2-a_2^2)\bar{Y} \\ 0 & a_2b_1\bar{Y} & a_2(b_1-a_1^2)\bar{Y} \\ 0 & b_1\bar{Y} & -(b_1-a_1^2)\bar{Y} \\ 0 & b_2\bar{Y} & -(b_2-a_2^2)\bar{Y} \\ 0 & a_1a_2\bar{Y} & a_1a_2\bar{Y} \end{pmatrix}. \tag{32}$$

5 Difference estimator

The previously presented mse for ratio have focused on ratio models that are usually met in the literature for mean estimation. An extension is the estimation of the difference of two means as usually called *estimation of a difference* or *estimation of change* (Sud and Srivastava, 2000). The rational behind this estimator is to check if the mean has changed its value in a significative way between two times, with corresponding two means denoted respectively $\bar{Y}_1$ and $\bar{Y}_2$. In the context of ratio estimator, the new statistic of interest may be the difference, $D_{12} = \bar{Y}_2 - \bar{Y}_1$, with corresponding estimator as follows,

$$\bar{d}_{\mathrm{R}_f} = \hat{\bar{Y}}_2 - \hat{\bar{Y}}_1 \,.$$

One first example of such estimator is for the usual ratio one, when:

$$\hat{\bar{Y}}_\ell = f_{\theta_\ell}(\bar{x}_\ell;\bar{X}_\ell)\bar{y}_\ell \,.$$

More generally this estimator may be one from the one proposed in the literature. This induces that random variables are added to the equation while components are added to the vector Ψ. Thus, this illustrates that the proposed approach is able to handle more models than just the ratio estimator for the automatic an expression of the mse with an eventual computation of parameters. An expression for the bias and for the mean squared error are found via an approximation as explained below.

Finite population statistics In a bidimensional setting, the notation is as follows. Two auxiliary variables X_j are available with population mean $\bar{X}_j$ and sample mean $\bar{x}_j$ for $j=1$ and $j=2$. Hence $x_i = (x_{i1}, x_{i2})$ is bidimensional. Let define $\bar{x}_j = \frac{1}{n}\sum_i^n x_{ij}$ and $\bar{y} = \frac{1}{n}\sum_i^n y_i$ and similarly $\bar{X}_j$ and $\bar{Y}$ for the population means. Let denote the $p=2$ free parameters α_j aggregated in the vector $\alpha = (\alpha_1,\alpha_2)^T$, such that the constraint $\sum_{j=1}^2 \alpha_j = 1$ may be introduced. Let denote $\delta_{x_j} = (\bar{x}_j - \bar{X}_j)/\bar{X}_j$ such as $f_{\theta_j}(.;.)$ is obtained by replacing in $f_\theta(.;.)$ θ by a new vector θ_j eventually different for each j if not equal to θ, while also replacing δ_x by δ_{x_j}. Let also denote $C_{0j} = S_{yx_j}/\bar{X}_j\bar{Y}$, $C_{jk} = S_{x_jx_k}/\bar{X}_j\bar{X}_k$, $C_0^2 = S_y^2/\bar{Y}^2$, $C_j^2 = S_{x_j}^2/\bar{X}_j^2$, $Cov_s(\bar{y},\bar{x}_j) = \lambda_n S_{yx_j}$, $Cov_s(\bar{x}_j,\bar{x}_k) = \lambda_n S_{x_jx_k}$, $Var_s(\bar{y}) = \lambda_n S_y^2$, $Var_s(\bar{x}_j) = \lambda_n S_{x_j}^2$, and $\lambda_n = \frac{1}{n}(1-f)$ where $f = n/N$. Let also denote the correlations $\rho_{0j} = S_{yx_j}/S_y S_{yx_j}$ and $\rho_{12} = S_{x_1x_2}/S_{x_1}S_{x_2}$.
Bivariate ratio estimators may be defined by the combination of the functions $f_\theta(.;.)$. For instance, the additive parametric ratio estimator is defined via a weighted sum of ratio estimators in (Priam, 2019), while the multiplicative parametric ratio estimator is via a product of ratio estimators. When denoting the new coefficients,

$$a_j = a_j(\theta_j),\, b_j = b_j(\theta_j)\,,$$

per function $f_{\theta_j}(.;.)$, let (re)define estimators for two auxiliary variables as listed in Table 2. It is recognized the parametric combinations of additive and multiplicative ratio estimators, and other parametric bivariate combined ratio and regression estimators. Note that in table 2, α_1 and α_2 have been removed in $\bar{y}_{\mathrm{R}_f^{c3}}$ and $\bar{y}_{\mathrm{R}_f^m}$. Hence the power α_1 and α_1 may be added in any existing estimator $\bar{y}_{\mathrm{R}_{est}}$ while renaming into $\bar{y}^{\alpha}_{\mathrm{R}_{est}}$ for more generality.

Difference modeling

For the new statistic of interest, the mean difference, $D_{12} = \bar{Y}_2 - \bar{Y}_1$, a corresponding estimator may be defined as $\bar{d}_{\mathrm{R}_f}$. Following the approach for ratio estimator, the expansion at the second order leads to the difference:

$$\begin{aligned}
\bar{d}_{\mathrm{R}_f} - (\bar{Y}_2 - \bar{Y}_1) &\doteq \left\{[1 + a_2\delta_{x_2} + b_2\delta_{x_2}^2]\bar{y}_2 - \bar{Y}_2\right\} - \left\{[1 + a_1\delta_{x_1} + b_1\delta_{x_1}^2]\bar{y}_1 - \bar{Y}_1\right\} \\
&= \left\{[1 + a_2\delta_{x_2} + b_2\delta_{x_2}^2]\bar{Y}_2(1+\delta_{y_2}) - \bar{Y}_2\right\} - \left\{[1 + a_1\delta_{x_1} + b_1\delta_{x_1}^2]\bar{Y}_1(1+\delta_{y_1}) - \bar{Y}_1\right\} \qquad (33)\\
&= a_2\bar{Y}_2\delta_{x_2} + b_2\bar{Y}_2\delta_{x_2}^2 + \bar{Y}_2\delta_{y_2} + a_2\bar{Y}_2\delta_{x_2}\delta_{y_2} - a_1\bar{Y}_1\delta_{x_1} - b_1\bar{Y}_1\delta_{x_1}^2 - \bar{Y}_1\delta_{y_1} - a_1\bar{Y}_1\delta_{x_1}\delta_{y_1}\,.
\end{aligned}$$

After having developped the terms, this induces that when denoting the vectors Ψ and $\Upsilon = \Upsilon(\delta)$ as:

$$\begin{aligned}
\Psi_{[12]} &= \left(\, T,\, U_0^{(1)},\, U_0^{(2)},\, U_1^{(1)},\, U_2^{(2)},\, V_{11}^{(1)},\, V_{22}^{(2)},\, V_{01}^{(1)},\, V_{02}^{(2)} \,\right) \\
\Upsilon_{[12]} &= \left(\, 1,\, \delta_{y_1},\, \delta_{y_2},\, \delta_{x_1},\, \delta_{x_2},\, \delta_{x_1}^2,\, \delta_{x_2}^2,\, \delta_{y_1}\delta_{x_1},\, \delta_{y_2}\delta_{x_2} \,\right),
\end{aligned} \qquad (34)$$

the difference is rewriten when identifying for the expression of the components in Ψ,

$$\bar{d}_{\mathrm{R}_f} - (\bar{Y}_2 - \bar{Y}_1) \doteq \Psi_{[12]}^T \Upsilon_{[12]}\,. \qquad (35)$$

Approximation of the bias and the mse via at first order

The expectation of the expression for $\bar{y}_{\mathrm{R}_{est}} - \bar{Y}$ just before leads to the bias of the generalized estimator:

$$\mathrm{bias}\left[\bar{d}_{\mathrm{R}_f}\right] \doteq -b_1\bar{Y}_1\lambda_n C_1^2 + b_2\bar{Y}_2\lambda_n C_2^2 - a_1\bar{Y}_1\lambda_n C_{10_1} + a_2\bar{Y}_2\lambda_n C_{20_2}\,. \qquad (36)$$

The expectation of the squared may lead to the approximated mean squared-error,

$$\begin{array}{rcl}[\bar{d}_{\mathrm{R}_f}-(\bar{Y}_2-\bar{Y}_1)]^2 & \doteq & \big[a_2\bar{Y}_2\delta_{x_2}+b_2\bar{Y}_2\delta_{x_2}^2+\bar{Y}_2\delta_{y_2}+a_2\bar{Y}_2\delta_{x_2}\delta_{y_2}-a_1\bar{Y}_1\delta_{x_1}-b_1\bar{Y}_1\delta_{x_1}^2-\bar{Y}_1\delta_{y_1} \\ & & -a_1\bar{Y}_1\delta_{x_1}\delta_{y_1}\big]^2 .\end{array} \tag{37}$$

Such that:

$$\begin{array}{rcl}\mathrm{mse}\left[\bar{d}_{\mathrm{R}_f}\right] & \doteq & \mathrm{amse}\left[\bar{d}_{\mathrm{R}_f}\right] \\ & \doteq & E\big[+\bar{Y}_1^2\delta_{y_1}^2+\bar{Y}_2^2\delta_{y_2}^2+\bar{Y}_1^2a_1^2\delta_{x_1}^2+\bar{Y}_2^2a_2^2\delta_{x_2}^2-2\bar{Y}_1\bar{Y}_2a_1a_2\delta_{x_2}\delta_{x_1}-2\bar{Y}_1\bar{Y}_2\delta_{y_1}\delta_{y_2} \\ & & +(+2\bar{Y}_1^2a_1\delta_{x_1}-2\bar{Y}_1\bar{Y}_2a_2\delta_{x_2})\delta_{y_1}+(-2\bar{Y}_1\bar{Y}_2a_1\delta_{x_1}+2\bar{Y}_2^2a_2\delta_{x_2})\delta_{y_2}\big] \\ & \doteq & \bar{Y}_1^2C_{0_1}^2+\bar{Y}_2^2C_{0_2}^2+\bar{Y}_1^2a_1^2C_1^2+\bar{Y}_2^2a_2^2C_2^2-2\bar{Y}_1\bar{Y}_2a_1a_2C_{12}-2\bar{Y}_1\bar{Y}_2C_{0_10_2} \\ & & 2\bar{Y}_1^2a_1C_{10_1}-2\bar{Y}_1\bar{Y}_2a_2C_{20_1}-2\bar{Y}_1\bar{Y}_2a_1C_{10_2}+2\bar{Y}_2^2a_2C_{20_2} .\end{array} \tag{38}$$

For the mse in a matricial form, it is followed the approach introduced herein for one mean. The expansion for the estimator is written in a vectorial form as $\Psi_{[12]}^T\Upsilon_{[12]}$ such that:

$$\begin{array}{rcl}\mathrm{amse}_\Omega\left[\bar{d}_{\mathrm{R}_f}\right] & \doteq & \Psi_{[12]}^T\left(\Omega_{[12]}+\alpha_{\Upsilon_{[12]}}\alpha_{\Upsilon_{[12]}}^T\right)\Psi_{[12]} \\ \mathrm{bias}_\alpha\left[\bar{d}_{\mathrm{R}_f}\right] & \doteq & \Psi_{[12]}^T\alpha_{\Upsilon_{[12]}} .\end{array} \tag{39}$$

6 Stratified sampling design and experiments

When a statification is introduced in the sampling procedure, the mean estimators are updated for dealing with the H strata. In a bidimensional setting with stratification, the notation is updated. Two auxiliary variables X_{jh} are available with population mean $\bar{X}_{jh}$ and sample mean $\bar{x}_{jh}$ for $j=1$ or $j=2$ and h for the stratum. Hence $x_{ih}=(x_{i1h},x_{i2h})$ is bidimensional. Let define $\bar{x}_{jh}=\frac{1}{n}\sum_i^n x_{ijh}$ and $\bar{y}_h=\frac{1}{n}\sum_i^n y_{ih}$ and similarly $\bar{X}_{jh}$ and $\bar{Y}_h$ for the population means per stratum. Let denote the free parameters α_{jh} per stratum and per auxiliary variable. Let denote $\delta_{x_{jh}}=(\bar{x}_{jh}-\bar{X}_{jh})/\bar{X}_{jh}$. Let also denote $C_{0j(h)}=S_{0j(h)}/\bar{X}_{jh}\bar{Y}_h$, $C_{jk(h)}=S_{jk(h)}/\bar{X}_{jh}\bar{X}_{kh}$, $C_{0(h)}^2=S_{yh}^2/\bar{Y}_h^2$, $C_{j(h)}^2=S_{j(h)}^2/\bar{X}_{jh}^2$, $Cov_s(\bar{y}_h,\bar{x}_{jh})=\lambda_{n_h}S_{0j(h)}$, $Cov_s(\bar{x}_{jh},\bar{x}_{kh})=\lambda_{n_h}S_{jk(h)}$, $Var_s(\bar{y}_h)=\lambda_{n_h}S_{0(h)}^2$, $Var_s(\bar{x}_{jh})=\lambda_{n_h}S_{j(h)}^2$, and $\lambda_{n_h}=\frac{1}{n_h}(1-f_h)$ where $f_h=n_h/N_h$. Let also denote the correlations $\rho_{0j(h)}=S_{0j(h)}/S_{0(h)}S_{0j(h)}$ and $\rho_{12(h)}=S_{12(h)}/S_{1(h)}S_{2(h)}$. For the approximations, $\delta_{x_{jh}}=(\bar{x}_{jh}-\bar{X}_{jh})/\bar{X}_{jh}$ and $\delta_{y_h}=(\bar{y}_h-\bar{Y}_h)/\bar{Y}_h$ are supposed less than one in absolute value. A stratification has the property to reduce the variability of the estimators, and even to optimize for the best structure of the strata in order to reduce further the variance.

Expression of the amse for the generalized mean estimator under stratification

Such estimators are presented in Table 3 by summation for the so-called separate estimators (Lone, Tailor, and Verma, 2017), while the combined estimators (Singh and Vishwakarma, 2008; Solanki and Singh, 2015; Lone, Tailor, and Singh, 2016; Kumar and Vishwakarma, 2019) are not considered herein.

The new expansion is written as previously in (4) when adding a weighted sum for the strata. The random variables are independent between strata hence the amse is obtained by summing each amse from each stratum:

$$\begin{array}{rcl}\mathrm{amse}_M\left[\bar{y}_{\mathrm{R}_{est}^{(str)}}\right] & = & \sum_{h=1}^H w_h^2\,\mathrm{amse}_M\left[\bar{y}_{\mathrm{R}_{est(h)}}\right] \\ & = & \sum_{h=1}^H w_h^2\,\Psi_{(h)}^T Q_{(h)}\Psi_{(h)} \\ & = & \sum_{h=1}^H w_h^2\,\left\{\xi_{0\,(h)}^T Q_{(h)}\xi_{0\,(h)}+2\xi_{0\,(h)}^T Q_{(h)}\xi_{\mathrm{K}\,(h)}K_{(h)}+K_{(h)}^T\xi_{\mathrm{K}\,(h)}^T Q_{(h)}\xi_{\mathrm{K}\,(h)}K_{(h)}\right\} .\end{array} \tag{41}$$

This induces that the solutions $K_{(opt)}^{(h)}$ for $K_{(h)}$ may be found separately by optimization of the criterion just before. The solutions are evently jointly if required with an aggregated expression. Let denote,

$$\xi_{0(str)}=\begin{pmatrix}\xi_{0\,(1)}\\ \xi_{0\,(2)}\\ \vdots\\ \\ \xi_{0\,(H)}\end{pmatrix},\ \xi_{\mathrm{K}(str)}=\begin{pmatrix}\xi_{\mathrm{K}\,(1)}\\ \xi_{\mathrm{K}\,(2)}\\ \vdots\\ \\ \xi_{\mathrm{K}\,(H)}\end{pmatrix}\ \text{and}\ Q_{(str)}=\begin{pmatrix}Q_{(1)} & & & (0)\\ & Q_{(2)} & & \\ & & \ddots & \\ (0) & & & Q_{(H)}\end{pmatrix}. \tag{42}$$

$$
\begin{array}{|l|l|}
\hline
\bar{y}_{\mathrm{R}^{(\mathrm{str})}_{diff}} & \sum_{h=1}^{H} w_h \left(\bar{y}_h + k_{1h}\delta_{x_{1h}} + k_{2h}\delta_{x_{2h}}\right) \\
\bar{y}_{\mathrm{R}^{(\mathrm{str})}_{\mathrm{Rao}}} & \sum_{h=1}^{H} w_h \left(k_{0h}\bar{y}_h + k_{1h}\delta_{x_{1h}} + k_{2h}\delta_{x_{2h}}\right) \\
\bar{y}_{\mathrm{R}^{m}_{f(\mathrm{str})}} & \sum_{h=1}^{H} w_h \, \bar{y}_h f_{\theta_1}(\bar{x}_{1h};\bar{X}_{1h}) f_{\theta_2}(\bar{x}_{2h};\bar{X}_{2h}) \\
\bar{y}_{\mathrm{R}^{c}_{f(\mathrm{str})}} & \sum_{h=1}^{H} w_h \left(k_{0h}\bar{y}_h + k_{1h}\delta_{x_{1h}}\right) f_{\theta}(\bar{x}_{2h};\bar{X}_{2h}) \\
\bar{y}_{\mathrm{R}^{c2}_{f(\mathrm{str})}} & \sum_{h=1}^{H} w_h \left(k_{0h}\bar{y}_h + k_{1h}\delta_{x_{1h}} + k_{2h}\delta_{x_{2h}}\right) f_{\theta_1}(\bar{x}_{1h};\bar{X}_{1h}) f_{\theta_2}(\bar{x}_{2h};\bar{X}_{2h}) \\
\bar{y}_{\mathrm{R}^{a}_{f(\mathrm{str})}} & \sum_{h=1}^{H} w_h \left(k_{1h} f_{\theta_1}(\bar{x}_{1h};\bar{X}_{1h}) + k_{2h} f_{\theta_2}(\bar{x}_{2h};\bar{X}_{2h})\right) \bar{y}_h \\
\bar{y}_{\mathrm{R}^{am}_{f(\mathrm{str})}} & \sum_{h=1}^{H} w_h \left(k_{0h} + k_{1h} f_{\theta_j}(\bar{x}_{1h};\bar{X}_{1h}) f_{\theta_2}(\bar{x}_{2h};\bar{X}_{2h})\right) \bar{y}_h \\
\bar{y}_{\mathrm{R}^{am2}_{f(\mathrm{str})}} & \sum_{h=1}^{H} w_h \left(k_{1h} f_{\theta_1}(\bar{x}_{1h};\bar{X}_{1h}) + k_{2h} f^{-1}_{\theta_1}(\bar{x}_{1h};\bar{X}_{1h})\right) f_{\theta_2}(\bar{x}_{2h};\bar{X}_{2h}) \bar{y}_h \\
\bar{y}_{\mathrm{R}^{am3}_{f(\mathrm{str})}} & \sum_{h=1}^{H} w_h \left(k_{1h} f_{\theta_1}(\bar{x}_{1h};\bar{X}_{1h}) f_{\theta_2}(\bar{x}_{2h};\bar{X}_{2h}) + k_{2h} f^{-1}_{\theta_{1h}}(\bar{x}_{1h};\bar{X}_{1h}) f^{-1}_{\theta_{2h}}(\bar{x}_{2h};\bar{X}_{2h})\right) \bar{y}_h \\
\hline
\end{array}
\qquad (40)
$$

Table 3: List of example of bivariates generalized ratio estimators under stratification.

When $K^{(1:H)}_{(opt)} = (K^{(1)T}_{(opt)}, K^{(2)T}_{(opt)}, \cdots, K^{(H)T}_{(opt)})^T$, this aggregated expression is more useful in the case of restricted parameters if required,

$$\xi^T_{\mathrm{K}(str)} Q_{(str)} \xi_{0(str)} + \xi^T_{\mathrm{K}(str)} Q_{(str)} \xi_{\mathrm{K}(str)} K^{(1:H)}_{(opt)} \quad = \quad 0 \,. \qquad (43)$$

This may be interesting for constraints on the bias. Here, it is written from the bias of each stratum,

$$\mathrm{bias}\left[\bar{y}_{\mathrm{R}^{(str)}_{est}}\right] \quad \doteq \quad \sum_{h=1}^{H} w_h \, \mathrm{bias}\left[\bar{y}_{\mathrm{R}_{est(h)}}\right] . \qquad (44)$$

As the overall bias may be written by a weighted sum (Tracy, Singh, and Singh, 1999) of the H bias within each stratum, this leads to add the constraint:

$$\textstyle\sum_h w_h T_h = 0 \qquad \text{or} \qquad T_h = 0 \ \text{ for each } h \,. \qquad (45)$$

As an example for two parameters k_0 and k_1, this leads to:

$$\Gamma = \begin{pmatrix} w_1 \\ w_1 \\ w_2 \\ w_2 \\ \vdots \\ w_H \\ w_H \end{pmatrix} \text{ and } \gamma = 1 \quad \text{ or } \quad \Gamma = \begin{pmatrix} 1 & 0 & & & 0 \\ 1 & 0 & & & 0 \\ 0 & 1 & & & 0 \\ 0 & 1 & & & 0 \\ \vdots & \vdots & \dots & \dots & \vdots \\ 0 & 0 & & & 1 \\ 0 & 0 & & & 1 \end{pmatrix} \text{ and } \gamma = \begin{pmatrix} 1 \\ 1 \\ \\ \vdots \\ 1 \end{pmatrix} . \qquad (46)$$

This expression may be appealing for removing the first term of the bias which does not depend explicitely of λ_h, as met in some multivariate estimators. For a direct optimisation, this requires the aggregated version in (43) of the objective function when it is added the linear restriction as in (20) but with the new linear system.

Examples of stratified estimators

An example of estimator for stratification and two auxiliary variables is as follows:

$$\bar{y}_{\mathrm{R}^{c(opt)}_{f(str)}} = \sum_{h=1}^{H} w_h \left\{ (k_{0h}\bar{y} + k_{1h}\delta_{x_{1h}}) \left[f_{\theta_2}(\bar{x}_{2h};\bar{X}_{2h})\right]^{\frac{a_{(\mathrm{h,opt})}}{a_2}} \right\} . \qquad (47)$$

This estimator updates the generalized estimator with an exponentiation by changing the value of a_2 to the optimal one $a_{(\mathrm{opt})}$. This estimator may depend on a_2 only in a nonlinear way from the second order term via the new expression of b_2 on the contrary to the former one without the exponantiation. The function involved here in this optimal estimator (at the first order) has the following expansion written:

$$
\begin{aligned}
f^{c(opt)}_{\theta_j}(\bar{x}_{jh};\bar{X}_{jh}) &= \left[f_{\theta_j}(\bar{x}_{jh};\bar{X}_{jh})\right]^{\frac{a_{(\mathrm{h,opt})}}{a_j}} \\
&= 1 + a_{(\mathrm{h,opt})}\delta_{x_{jh}} - \frac{a_{(\mathrm{h,opt})}\left(a_j^2 - a_{(\mathrm{h,opt})}a_j - 2b_j\right)}{2a_j}\delta^2_{x_{jh}} + \cdots .
\end{aligned}
\qquad (48)
$$

Selected estimators when associated to parametric functions are considered in the experiments next section with real data. Next sections, two extensions of the approach is presented, one for higher order estimation of the mse and one for the difference of two means instead of one mean.

7 Numerical experiments

In this subsection, the numerical results for several estimators are presented. In the previous section stratified estimators are discussed with their mean squared error. The approximation of the mean squared errors are kept at the first order, hence a comparison with higher orders is left as a perspective. The main purpose of this section is to check if this kind of estimators are relevant for decreasing the mean squared error as previously observed graphically without stratification. Note that the stratification allows for optimisation at the level of the size of the strata, which is not studied here neither. Next, the amse and the bias of several mean estimators are visualized for one real dataset from the literature in order to compare their values.

Real dataset

The dataset is from (Koyuncu and Kadilar, 2009b; Lone et al., 2017; Muneer, Khalil, and Shabbir, 2020). This dataset is considered in several other communications on ratio estimation with stratified sampling for separate and combined approaches. The version for the numerical results of this section is given in Table 4.

Stratum	1	2	3	4	5	6
N_h	127	117	103	170	205	201
n_h	31	21	29	38	22	39
$S_{0(h)}$	883.84	644.92	1033.40	810.58	403.65	711.72
$S_{1(h)}$	30486.70	15180.77	27549.69	18218.93	8497.77	23094.14
$S_{2(h)}$	555.58	365.46	612.95	458.03	260.85	397.05
$\bar{Y}_h$	703.74	413.00	573.17	424.66	267.03	393.84
$\bar{X}_{1(h)}$	20804.59	9211.79	14309.3	9478.85	5569.95	12997.59
$\bar{X}_{2(h)}$	498.28	318.33	431.36	311.32	227.20	313.71
$\rho_{01(h)}$	0.9360	0.9960	0.9940	0.9830	0.9890	0.9650
$\rho_{02(h)}$	0.9790	0.9760	0.9840	0.9830	0.9640	0.9830
$\rho_{12(h)}$	0.9396	0.9696	0.9770	0.9640	0.9670	0.9960

Table 4: Real dataset with six strata.

Examples of numerical computation of the amse and the bias

Some estimators with their mean squared error and bias, as given in (Lone et al., 2017) for this dataset, are listed in the Table 5 in order to compare the values with the proposed generalized approach. Note that it is retrieved exactly the same values when using the formula for each estimator from the usual approach. The differences come with the linearization as seen in the Table 5. Between the two versions of the amse, with or without the terms in λ_n^2, there is for these examples of estimators only a relative difference of less than 5% except for the more biased estimator, as observed in Table 5. More precisely, they are from the first row to the last one equal to 1.02%, 1.31%, 0.00%, 6.00%, 4.80%, 0.00%. This suggests that the usual approach to neglect these terms may be relevant for non biased estimators but there may be no reason to remove them in the proposed matricial approach actually.

Graphical comparisons

The numerical results lead to compare empirically diverse estimators presented in the previous sections. Here, they are computed via the proposed generic approach. With the aggregated estimator the graphical output from rstudio (R Core Team, 2017) is in Figure 1, with the PRE (gain over the standard mean estimator) and the bias of the estimators compared: $\bar{y}_{\mathrm{R}^{c}_{f(str)}}$, $\bar{y}_{\mathrm{R}^{c(opt)}_{f(str)}}$, $\bar{y}_{\mathrm{R}^{c4}_{f(str)}}$ and $\bar{y}_{\mathrm{R}^{(\mathrm{str})}_{diff}}$. In Figures 1 and 2, the x-axis is for $a_j \in [-1;1]$ in $\bar{y}_{\mathrm{R}^{c}_{f(str)}}$ and $\bar{y}_{\mathrm{R}^{c4}_{f(str)}}$ while $b_j = 0.5(a_j^2 + a_j)$. This same axis is for b_j in $[-10;10]$ for $\bar{y}_{\mathrm{R}^{c(opt)}_{f(str)}}$, hence with a factor ten, such that 0.5 means 5.

Per stratum, the graphical output from rstudio is in Figure 2 for the PRE and the bias in order to compare also the behavior of the estimators without stratification. According to the graphical output obtained, the estimator $\bar{y}_{\mathrm{R}^{c4}_{f(str)}}$

Estimator	a_1	b_1	a_2	b_2	Value of bias	Value of amse without terms λ_n^2	Value of amse with terms λ_n^2
$\sum_h w_h \bar{y}_h \frac{\bar{X}_{1h}}{\bar{x}_{1h}}$	-1	1	0	0	2.08	129.82	131.14
$\sum_h w_h \bar{y}_h \frac{\bar{x}_{2h}}{\bar{X}_{2h}}$	0	0	1	0	23.23	6925.35	7016.15
$\sum_h w_h \bar{y}_h \frac{\bar{X}_{1h}}{\bar{x}_{1h}} \frac{\bar{x}_{2h}}{\bar{X}_{2h}}$	-1	1	1	0	1.30	1343.30	1343.78
$\sum_h w_h \bar{y}_h \frac{\bar{x}_{1h}}{\bar{X}_{1h}} \frac{\bar{x}_{2h}}{\bar{X}_{2h}}$	1	0	1	0	78.77	17370.17	18412.08
$\sum_h w_h \bar{y}_h \frac{\bar{X}_{1h}}{\bar{x}_{1h}} \frac{\bar{X}_{2h}}{\bar{x}_{2h}}$	-1	1	-1	1	21.16	1580.84	1656.73
$\sum_h w_h \bar{y}_h \frac{\bar{x}_{1h}}{\bar{X}_{1h}} \frac{\bar{X}_{2h}}{\bar{x}_{2h}}$	1	0	-1	1	2.57	3677.04	3678.28

Table 5: Some estimators with $amse = \Psi^T Q \Psi$ and $bias = \Psi^T \alpha_\Upsilon$ from the proposed approach when the additional terms are kept (last column) or not, for the dataset considered. The estimator here is $\bar{y}_{\mathrm{R}^{am}_{f(\mathrm{str})}}$ with $k_{0h} = 0$ and $k_{1h} = 1$ for all h.

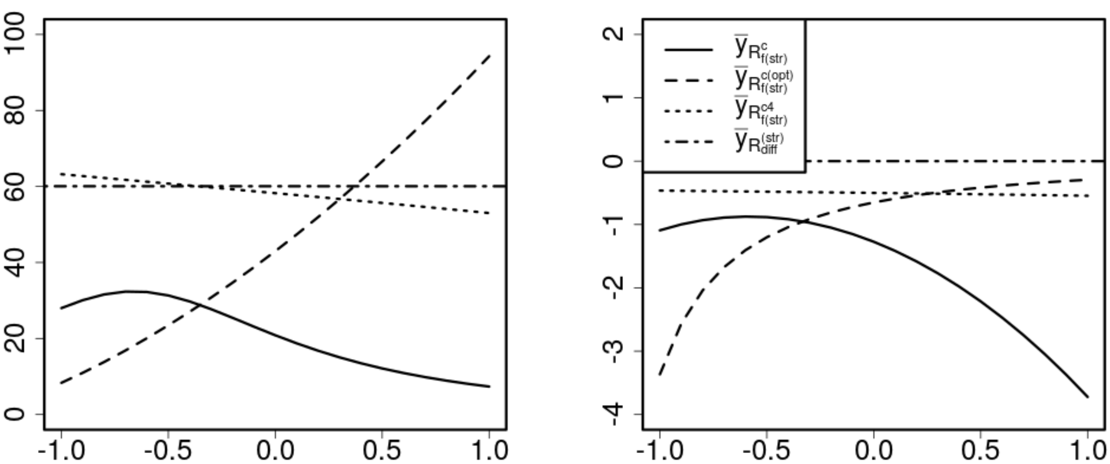

Figure 1: PRE (left) and BIAS (right) for the considered dataset from the proposed analytical expressions.

is not able to perform better than the difference estimator except in a stratum and nearly for the aggregated estimator. The estimator $\bar{y}_{\mathrm{R}^c_{f(str)}}$ (with optimal parameters in Table 6) is not able to reach its optimum value in the aggregated case because it is able only to perform well in each stratum separately with a different optimal value for a_j.

According to the computed value, the estimator $\bar{y}_{\mathrm{R}^{c(opt)}_{f(str)}}$ may be a better solution with only a parameter to set $b = b_j$ chosen equal in all the strata. Graphically, this estimator is able to perform better than the other estimators when b_j varies in $[-10; 10]$ and at the first order. Note that the large improvement observed here in the Figure 1 may be optimistic considering the high correlations and high variances of the variables in the six strata, but also other parametric functions are available from the literature. The bias remains small in all the obtained output for the four estimators at the first order.

8 Discussion and perspectives

Herein, we propose to review several existing ratio estimators from the literature via generic models when two auxiliary variables are available.

- A new expression for a generalized expression of an approximated mse is introduced for a stratified sampling

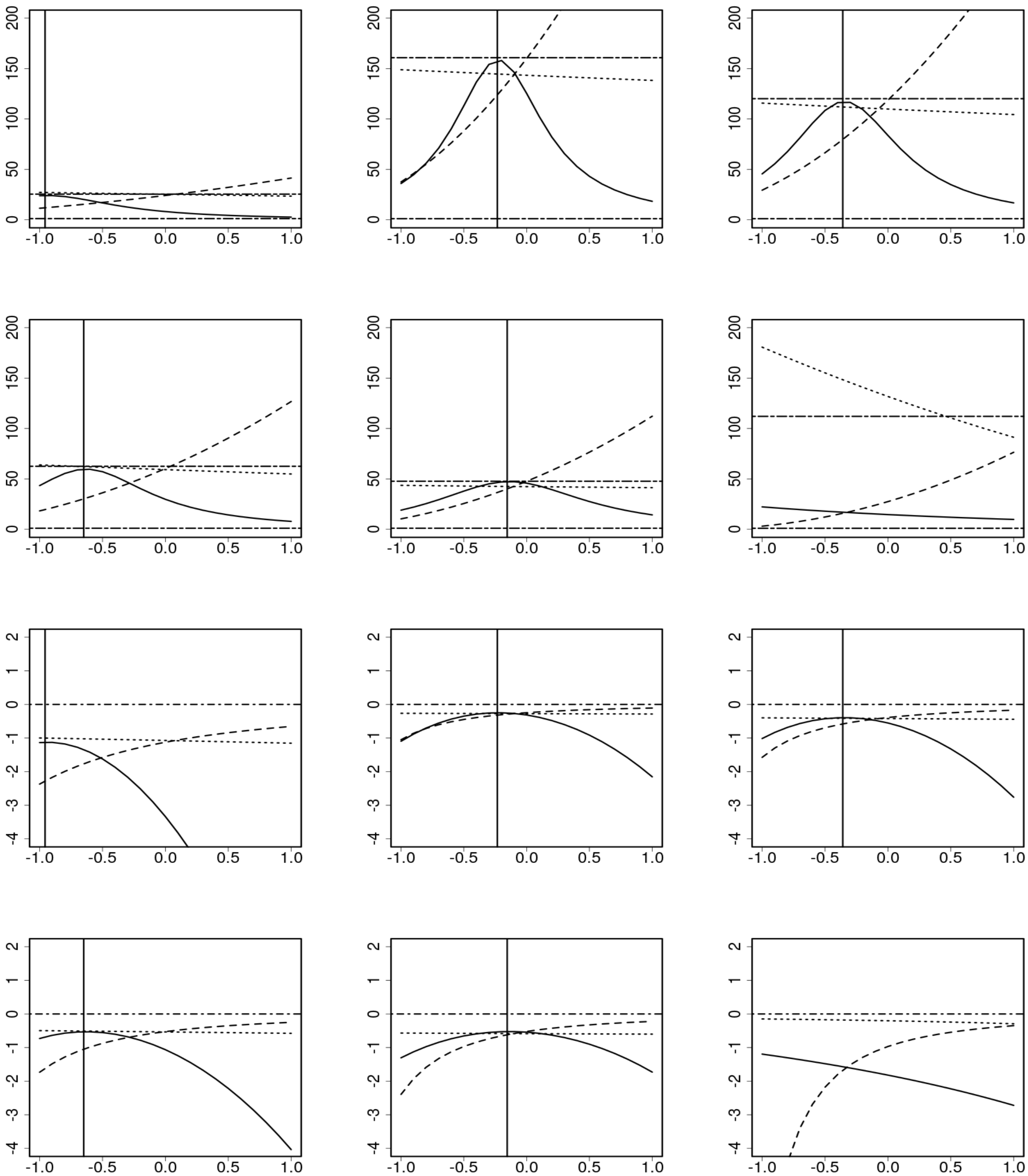

Figure 2: PRE (two top rows) and BIAS (two bottom rows) per stratum ($h = 1$ to $h = 3$ for first row and $h = 4$ to $h = 6$ for second row) for the considered dataset from the proposed analytical expressions. The vertical line is for the value of the optimal parameter $a_{(\mathrm{h,opt})}$.

designs, in order to compare generalized ratio estimators with stratification. A new justification for the expression of the generic amse from (Priam, 2019) is also proposed. The computation of the unknown parameters is automatic while requiring only the expansion of the estimators when it belongs to a family of a generalized estimator. Additive estimators, which have been not considered further before, have their corresponding matrices Φ given herein.

Stratum	1	2	3	4	5	6
w_h	0.1376	0.1268	0.1116	0.1842	0.2221	0.2178
$a_{(\mathrm{h,opt})}$	-0.9570	-0.2335	-0.3589	-0.6494	-0.1549	-3.9093
$b_{(\mathrm{h,opt})}$	-0.0030	0.0093	0.0121	0.0206	0.0285	0.0927
$C^2_{0(h)}$	1.5773	2.4384	3.2506	3.6434	2.2850	3.2657
$C^2_{1(h)}$	2.1473	2.7158	3.7068	3.6943	2.3276	3.1570
$C^2_{2(h)}$	1.2432	1.3180	2.0192	2.1646	1.3182	1.6019
$C_{01(h)}$	1.7226	2.5631	3.4504	3.6064	2.2808	3.0985
$C_{02(h)}$	1.3709	1.7497	2.5210	2.7605	1.6730	2.2483
$C_{12(h)}$	1.5352	1.8344	2.6729	2.7260	1.6938	2.2398

Table 6: Values of the optimal parameters per stratum for a combined additive estimator, plus moment statistics of the dataset.

- The proposed extensions allows for checking different hypothesis. The second order expression of the mse leads to validate the one order one because a divergent estimator is expected to remain divergent at higher order as the remaining term from the serie approximation may be not bounded. This means that when $\mathrm{amse}_{\Omega}^{(\ell)}\left[\bar{y}_{\mathrm{R}_{est}}\right]$ denotes the amse at the order ℓ^{th} it may be considered to compare:

$$\mathrm{amse}_{\Omega}^{(2)}\left[\bar{y}_{\mathrm{R}_{est}}\right] - \mathrm{amse}_{\Omega}^{(1)}\left[\bar{y}_{\mathrm{R}_{est}}\right] \text{ with } \mathrm{amse}_{\Omega}^{(1)}\left[\bar{y}_{\mathrm{R}_{est}}\right].$$

  By this way it becomes possible to check if the remaining term from the linearization at second order is small as wanted in theory and in practice. Not that this supposes a careful completion for the matrix Ω and more generally Q in order to keep the relevant terms. More precisely, the usual hypothesis from the literature may lead to check if the difference just before is of order $O(n^{-1})$. The reason why is that in the literature it is usually supposed that ratio estimators are able to reach such bound at the first order leaving the higher terms smaller. Similarly, the expression with ratio estimators for variance of change may be considered in order to check if adding a ratio to the means is able to decrease the mean squared error of the difference.

- When only a sample is available, the variance and bias of the estimators are estimated with mostly sample statistics and unequal probabilities. Alternative estimations for the mean squared error may be possible for instance via Horvitz-Thompsom variance estimator (Horvitz and Thompson, 1952) (Younis and Shabbir, 2019) or alternative related ways. These approaches lead to a quadratic function hence they are of main interest for the proposed generalized estimators and for the estimation of their parameters.

- The linearisation may not be always optimal for small sample, or a large coefficient of variation. The expansion may be true only if $|\delta_x| < 1$ but δ_x is a random variable hence this is not because for the one sample available it is obtained $|\delta_x| < 1$ that the inequality remains true. This is why the variance of $\bar{x}$ may need to be checked too. Actually, in the literature, some conditions are given for the usual ratio estimator, as follows. It may be required $n > 30$ and the coefficient of variations smaller than 0.1 otherwise the true mean squared error may be underestimated. This limits the use of the analytical formula obtained from the usual linearisation, and even may reduce the number of datasets suitable with such solution. When $|\delta_x| > 1$ the multiplicative correction may not behave well at all for negative values, more generally and the approximation is not performing well even for positive values as expected from a Taylor serie which is only tangencial to the true curve. This illustrates visually an eventual problem when estimating the mse with such approximation: some real samples may not respect the conditions of validity. Graphically, a second order approach does not improve much the first order one with the linearization. This also underlines that the linearization suffers from the not symetrical shape of the function around $\delta_x = 0$. This suggests to prefer an estimator of the mse with a more accurate fitting of the function f_θ, which is possible when δ_x is not small. This leads to check the standard deviation denoted σ_{δ_x}, as, $sd(\delta_x) = \sqrt{\lambda_n} C_x$. An enough small value of this quantity insures that the values of δ_x remain small and around 0 which is the requirement for the Taylor linearisation, but graphically the quadratic approximation for a ratio is quickly not accurate when δ_x is not near zero which is likely to happens in a random setting. This may be measured by the probability to be out of an interval around zero with for instance a gaussian approximation or a t-Student one as follows,

$$\delta_x/\sigma_{\delta_x} \sim \mathcal{N}(0,1) \text{ or } \delta_x/\sigma_{\delta_x} \sim \mathcal{T}(n).$$

This translates in term of probability for the random variable δ_x as follows with a constant η which is expected to be small as 10^{-2} for instance, $P(|\delta_x| < 1) = 1 - \eta$. For the range of values of δ_x it is also deduced if the Taylor serie is relevant by computing the corresponding maximal distance[1] between the true curve f_θ and the approximating curve $\tilde{f}_\theta$ which is the polynomial approximation. An alternative expansion to the Taylor serie may be considered. For the reasons presented just before, a quadratic function is estimated in order to mimic the function $f_\theta(\bar{x};\bar{X})$, in the positive part, for an interval around zero. This is written with $\tilde{\beta}_k = \tilde{\beta}_k(\theta)$ as follows,

$$\tilde{f}_\theta(\bar{x};\bar{X}) \doteq \tilde{\beta}_0 + \tilde{\beta}_1\delta_x + \tilde{\beta}_2\delta_x^2.$$

A new interval for the random variable δ_x is to be considered in comparison to the usual linearization. Two constant values exist such that, $P(\delta_x^{min} \leq \delta_x \leq \delta_x^{max}) = 1 - \eta$. As δ_x is left bounded by -1 because $\bar{x} > 0$, this induces $\delta_x^{min} > -1$. Even if the distributional hypothesis is able to lead to such bounds, it may be preferred a non parametric approach, such as a resampling procedure via boostrapping or en empirical likelihood method for more precise values.

The perspective remains a further study of the behaviour of the mses and the bias w.r.t. the function $f_\theta(.;.)$ with eventually higher orders (Koyuncu and Kadilar, 2010) and additional auxiliary variables in the approximations and other generalized estimators. A further understanding of the quality of the approximation in the general case is also appealing.

References

Allen, J., H. P. Singh, and F. Smarandache (2003). A family of estimators of population mean using multi-auxiliary information in presence of measurement errors. *International Journal of Social Economics 30*, 837–849.

Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science 30*(2), 262–275.

Diana, G., M. Giordan, and P. F. Perri (2011). An improved class of estimators for the population mean. *Statistical Methods & Applications 20*(2), 123–140.

Diana, G. and P. F. Perri (2007). Estimation of finite population mean using multi-auxiliary information. *Metron LXV*(1), 99–112.

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

Koyuncu, N. and C. Kadilar (2009a). Family of estimators of population mean using two auxiliary variables in stratified random sampling. *Communications in Statistics - Theory and Methods 38*(14), 2398–2417.

Koyuncu, N. and C. Kadilar (2009b). Ratio and product estimators in stratified random sampling. *Journal of Statistical Planning and Inference 139*(8), 2552–2558.

Koyuncu, N. and C. Kadilar (2010). On improvement in estimating population mean in stratified random sampling. *Journal of Applied Statistics 37*(6), 999–1013.

Kumar, M. and G. Vishwakarma (2019, 06). Generalized classes of regression-cum-ratio estimators of population mean in stratified random sampling. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*.

Lone, H. A., R. Tailor, and H. P. Singh (2016). Generalized ratio-cum-product type exponential estimator in stratified random sampling. *Communications in Statistics - Theory and Methods 45*(11), 3302–3309.

Lone, H. A., R. Tailor, and M. R. Verma (2017). Efficient separate class of estimators of population mean in stratified random sampling. *Communications in Statistics - Theory and Methods 46*(2), 554–573.

Muneer, S., A. Khalil, and J. Shabbir (2020). A parent-generalized family of chain ratio exponential estimators in stratified random sampling using supplementary variables. *Communications in Statistics - Simulation and Computation 0*(0), 1–22.

[1] with $\tilde{g}_\theta = f_\theta - \tilde{f}_\theta$, $\tilde{\varepsilon}_m \leq |\tilde{g}_\theta| \leq \tilde{\varepsilon}_M$ and $0 < \tilde{\varepsilon}_m, \tilde{\varepsilon}_M \ll 1$.

Olkin, I. (1958). Multivariate ratio estimation for finite populations. *Biometrika 45*(1/2), 154–165.

Priam, R. (2019). Visualization of generalized mean estimators using auxiliary information in survey sampling. *Communications in Statistics - Theory and Methods*, 1–22.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rao, T. J. (1991). On certain methods of improving ratio and regression estimators. *Communications in Statistics - Theory and Methods 20*(10), 3325–3340.

Singh, G. N. and M. Khalid (2018). An effective estimation strategy for population mean under random non-response situations in two-phase successive sampling. *Communications in Statistics - Theory and Methods 47*(19), 4641–4660.

Singh, H. P. and G. K. Vishwakarma (2008). A family of estimators of population mean using auxiliary information in stratified sampling. *Communications in Statistics - Theory and Methods 37*(7), 1038–1050.

Singh, H. P. and A. Yadav (2018). Improved generalized family of estimators of population mean using information on transformed auxiliary variables. *Pakistan Journal of Statistics and Operation Research 14*(4), 913–934.

Solanki, R. S. and H. P. Singh (2015). Efficient classes of estimators in stratified random sampling. *Statistical Papers 56*(1), 83–103.

Srivastava, S. K. (1971). A generalized estimator for the mean of a finite population using multi-auxiliary information. *Journal of the American Statistical Association 66*(334), 404–407.

Srivastava, S. K. and H. S. Jhajj (1981). A class of estimators of the population mean in survey sampling using auxiliary information. *Biometrika 68*(1), 341–343.

Sud, U. and A. Srivastava (2000, 01). Estimation of population mean in repeat surveys in the presence of measurement errors. *Journal of the Indian Society of Agricultural Statistics 53*.

Sukhatme, P. (1970). *Sampling Theory of Surveys with Applications*. Indian Society of Agricultural Statistics, New Delhi and the Iowa State University Press, Ames, Iowa.

Tracy, D., P. H. Singh, and R. Singh (1999, 09). Constructing an unbiased estimator of population mean in finite populations using auxiliary information. *Statistical Papers 40*, 363–368.

Vishwakarma, G. K. and M. Kumar (2015). An efficient class of estimators for the mean of a finite population in two-phase sampling using multi-auxiliary variates. *Communications in Mathematics and Statistics 3*(4), 477–489.

Vishwakarma, G. K. and H. P. Singh (2012). A general procedure for estimating the mean using double sampling for stratification and multi-auxiliary information. *Journal of Statistical Planning and Inference 142*(5), 1252 – 1261.

Wolter, K. (2007). *Introduction to Variance Estimation*. Springer Series Statistics for Social and Behavioral Sciences.

Yasmeen, U., M. N. Amin, and M. Hanif (2015). Generalized exponential estimators of finite population mean using transformed auxiliary variate. *International Journal of computational Mathematics 1*(2), 1–10.

Younis, F. and J. Shabbir (2019). Generalized ratio-type and ratio-exponential-type estimators for population mean under modified horvitz-thompson estimator in adaptive cluster sampling. *Journal of Statistical Computation and Simulation 0*(0), 1–11.