5-1-2017

A New Estimator based on Auxiliary Information through Quantitative Randomized Response Techniques

Nilgün Özgül
Hacettepe University, nilgunozgul@yahoo.com

Hülya Çıngı
Hacettepe University, hcingi@hacettepe.edu.tr

A New Estimator based on Auxiliary Information through Quantitative Randomized Response Techniques

Nilgün Özgül
Hacettepe University
Ankara, Turkey

Hülya Çıngı
Hacettepe University
Ankara, Turkey

An exponential-type estimator is developed for the population mean of the sensitive study variable based on various Randomized Response Techniques (RRT) using a non-sensitive auxiliary variable. The mean squared error (MSE) of the proposed estimator is derived for generalized RRT models. The proposed estimator is compared with competitors in a simulation study and an application. The proposed estimator is found to be more efficient using a non-sensitive auxiliary variable.

Keywords: Randomized response techniques, sensitive question, auxiliary variable, exponential estimator, efficiency

Introduction

In surveys on sensitive topics, estimation of the population mean with a direct questioning technique may cause respondents to refuse answering or to give untruthful answers on purpose. Respondents may encounter questions about drug use, illegal income, political views, abortion, homosexual activities, and AIDS in some social, medical, and epidemiological questionnaires. On these surveys, respondents do not feel comfortable and they may choose not to answer or may intentionally provide false answers. This can bring about significant bias in the estimation of population parameters.

Random response techniques (RRT) are used to reduce nonrespondent's rates and biased responses to sensitive questions. Warner (1965) introduced the randomization technique for the proportion of a population characterized by a sensitive variable, which was followed by studies where the response to a sensitive question results in a quantitative variable. Quantitative RRT are used to

---

Dr. Özgül one is a Professor in the Department of Statistics. Email them at: nilgunozgul@yahoo.com.

estimate the mean value of some behavior in a population. For example, the sensitive study variable may be the total number of abortions a woman has had or the average weekly alcohol consumption or annual earnings of people. These RRT are sub-classified as either additive or multiplicative techniques.

In additive RRT, respondents are asked to scramble their responses using a randomization device such as a deck of cards. Each of the cards in the deck has a number. The numbers in the deck follow a known probability distribution, such as Normal, Chi-square, Uniform, Poisson, Binomial, Weibull, etc. The respondent is asked to add the real response to the number listed on card picked, and then report only the sum to the interviewer. Multiplicative RRT are similar to additive RRT. Again, a deck of cards with known probability distribution is used, but now when the respondents scramble their responses, they are asked to report the product of the real response and the number listed on the selected card. The interviewer cannot see the card, but records the reported number. RRT can also be categorized by how the respondents are instructed to randomize. If all respondents are asked to randomize their response, the model is characterized as a full randomization RRT model. If some of the respondents are instructed to randomize their response, the model is characterized as a "partial RRT model" (Özgül, 2013).

Thornton and Gupta (2004) extended Warner's (1971) approach by using partial additive models for estimating the mean of sensitive quantitative variables in RRT. The multiplicative model was later investigated in depth by Eichhorn and Hayre (1983), who referred to it as the scrambled responses method. Similarly, Bar-Lev, Bobovitch, and Boukai (2004) proposed a method which uses a partial model that generalizes Eichhorn and Hayre's results and yields an estimate which, under mild conditions, has a uniformly smaller variance. Further developments focused on the use of auxiliary variables to improve the precision. Diana and Perri (2011), Sousa, Shabbir, Real, and Gupta (2010), and Gupta, Shabbir, Sousa, and Real (2012) suggested mean estimators using the auxiliary variable for estimating of the quantitative sensitive variable in RRT. Bahl and Tuteja (1991), Shabbir and Gupta (2011), Grover and Kaur (2014), and Özgül and Cingi (2014) studied exponential-type estimators to obtain more efficient estimates for various sampling methods. In the current study, an exponential-type estimator of the mean of a sensitive variable is proposed using a non-sensitive auxiliary variable for generalized partial quantitative RRT.

Various Estimators Based on Auxiliary Information through Quantitative RRT

Diana and Perri (2011) introduced a general mechanism to scramble responses and proposed a class of regression estimators for the mean of a sensitive variable using a non-sensitive auxiliary variable. To estimate μ_y, a sample of individuals is selected from the population and each respondent is asked to perform a Bernoulli trial with a probability of success P. If this is successful, the respondent then gives the true values of both Y and X. In the case of failure, the respondent gives their answers by using the values given in S and R, which are the various randomized designs for the variables Y and X, respectively. The interviewer does not know the outcome of the Bernoulli experiment. Then, the distribution of the responses is given in (1) as

$$(Z,U) \rightarrow \begin{cases} (Y,X) & \text{with probability } P \\ (S,R) & \text{with probability } (1-P) \end{cases} \tag{1}$$

where Y is the sensitive variable of interest with unknown mean μ_y and unknown variance S_y^2, X is the non-sensitive variable with known mean μ_x and known variance S_x^2, Z is the reported response for the sensitive variable Y, and U is the reported response for the first non-sensitive variable X. In S and R, the respondents answer the questions using the additive or multiplicative technique. For the additive technique, each respondent is requested to draw a value from the distribution of the scrambling variable, add it to the real response, and report back to the interviewer. For the multiplicative model, the respondent responds with the product of the drawn value and their true response. The scrambling variables are defined as W and T which have pre-assigned distributions such as Normal, Chi-square, Uniform, Poisson, Binomial, Weibull, etc. W is the scrambling variable with known true mean μ_w and known variance S_w^2 in S and T is the scrambling variable with known true mean μ_t and variance S_t^2 in R (Özgül, 2013).

Under the generic scheme given in (1), the following class of estimators based on a SRSWR sample $\{(z_1, u_1), (z_2, u_2),\ldots, (z_n, u_n)\}$ of n responses is

$$\hat{\mu}_{\text{DP}} = \frac{\bar{z} + b(\mu_u - \bar{u}) - c}{h}, \quad (h \neq 0) \tag{2}$$

366

where b is a suitably selected real constant and

$$\bar{z} = \frac{1}{n} \sum_{i=1}^{n} z_i, \quad \bar{u} = \frac{1}{n} \sum_{i=1}^{n} u_i$$

are the sample means of the reported responses for the sensitive variable and the non-sensitive auxiliary variable, respectively. Here, c and h depend exclusively on the scrambling design.

The variance of $\hat{\mu}_{DP}$ is

$$\text{Var}\left(\hat{\mu}_{DP}\right) = \frac{1}{nh^2}\left(S_z^2 - 2BS_{zu} + B^2 S_u^2\right) \tag{3}$$

where

$$S_z^2 = \frac{\sum_{i=1}^{N}\left(z_i - \mu_z\right)^2}{N-1}, \quad S_u^2 = \frac{\sum_{i=1}^{N}\left(u_i - \mu_z\right)^2}{N-1}$$

are the population variances of z and u, respectively,

$$S_{zu} = \frac{\sum_{i=1}^{N}\left(z_i - \mu_z\right)\left(u_i - \mu_u\right)}{N-1}$$

is the population covariance between z and u, $B = S_{zu}/S_u^2$ is the population regression coefficient between z and u, and

$$\mu_z = \frac{1}{N} \sum_{i=1}^{N} z_i, \quad \mu_u = \frac{1}{N} \sum_{i=1}^{N} u_i$$

are the population means of z and u, respectively. The minimum variance of $\hat{\mu}_{DP}$ is

$$\text{Var}\left(\hat{\mu}_{DP}\right)_{\min} = \frac{S_z^2}{nh^2}\left(1 - \rho_{zu}^2\right) \tag{4}$$

where $\rho_{zu} = S_{zu} / (S_z S_u)$ is the population correlation coefficient between z and u.

Sousa et al. (2010) proposed a ratio estimator for the mean of a sensitive variable using a non-sensitive auxiliary variable. The respondent is asked to provide true responses for X. The Sousa et al. estimator is

$$\hat{\mu}_{SR} = \bar{z} \left(\frac{\mu_x}{\bar{x}} \right) \tag{5}$$

where $\bar{z}$ is the sample mean of the reported responses for the sensitive variable $(Z = Y + W)$,

$$\mu_x = \frac{1}{N} \sum_{i=1}^{N} x_i$$

is the known population mean of non-sensitive auxiliary variable, and

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

is the sample mean of non-sensitive auxiliary variable. The Bias and MSE of $\hat{\mu}_{SR}$, under first order of the approximation, is

$$\text{Bias}(\hat{\mu}_{SR}) \cong \lambda \mu_z \left(C_x^2 - C_{zx} \right) \tag{6}$$

$$\text{MSE}(\hat{\mu}_{SR}) = \lambda \mu_z^2 \left[C_z^2 + C_x^2 - 2\rho_{xz} C_x C_z \right] \tag{7}$$

where

$$\lambda = \frac{1}{n} - \frac{1}{N}$$

and $C_z = S_z / \mu_z$ and $C_x = S_x / \mu_x$ are the coefficients of variation of Z and X, respectively.

Gupta et al. (2012) proposed regression-cum-ratio estimator using a non-sensitive auxiliary variable.

$$\hat{\mu}_{\mathrm{GRR}} = \left[b_1 \bar{z} + b_2 \left(\mu_x - \bar{x} \right) \right] \left(\frac{\mu_x}{\bar{x}} \right) \tag{8}$$

where $\bar{z}$, µx, and $\bar{x}$ are defined as above for (5), and b_1 and b_2 are constants. The Bias and minimum MSE of $\hat{\mu}_{\mathrm{GRR}}$, under first order of the approximation, is

$$\mathrm{Bias}\left(\hat{\mu}_{GRR} \right) \cong \left(b_1 - 1 \right) \mu_z + b_1 \lambda \mu_z \left\{ C_x^2 - C_{zx} \right\} + b_2 \lambda \mu_x C_x^2 \tag{9}$$

$$\mathrm{MSE}\left(\hat{\mu}_{GRR} \right)_{\min} \cong \frac{\lambda \mu_z^2 C_z^2 \left(1 - \rho_{zx}^2 \right) \left(1 - \lambda C_x^2 \right)}{\lambda C_z^2 \left(1 - \rho_{zx}^2 \right) + \left(1 - \lambda C_x^2 \right)} \tag{10}$$

Suggested Exponential-Type Estimator Based on Auxiliary Information through Quantitative RRT

Applying the general formulation of Diana and Perri (2011) and following Grover and Kaur (2014), an exponential-type estimator for the mean of a sensitive variable is proposed using a non-sensitive auxiliary variable in RRT. Consider the following improved exponential estimator based on a SRSWOR sample $\{(z_1, u_1), (z_2, u_2),\ldots, (z_n, u_n)\}$ of n responses:

$$\hat{\mu}_{\mathrm{NH(exp)}} = \frac{\left[\left\{ b_1 \bar{z} + b_2 \left(\mu_u - \bar{u} \right) \right\} \exp\left(\frac{\alpha \left(\mu_u - \bar{u} \right)}{\alpha \left(\mu_u + \bar{u} \right) + \beta} \right) \right] - c}{h}, \quad \left(h \neq 0 \right) \tag{11}$$

where b_1 and b_2 are suitably selected real constant and α and β are already assumed to be either any known constants or functions of any known population parameters of the auxiliary variable, such as standard deviation (σ_x), coefficient of variation (C_x), coefficient of skewness $\{\beta_1(x)\}$, coefficient of kurtosis $\{\beta_2(x)\}$, coefficient of correlation (ρ_{yx}) (Cingi & Kadilar, 2009). Here, c and h depend exclusively on the scrambling design.

To obtain the MSE equation for the proposed estimator, we define following relative error terms

$$e_0 = \frac{\left(\bar{z} - \mu_z \right)}{\mu_z}, e_1 = \frac{\left(\bar{u} - \mu_u \right)}{\mu_u}$$

such that

$$\mathrm{E}(e_0)=\mathrm{E}(e_1)=0;\ \mathrm{E}(e_0^2)=\lambda C_z^2,\ \mathrm{E}(e_1^2)=\lambda C_u^2,\ \mathrm{E}(e_0 e_1)=\lambda \rho_{zu} C_z C_u$$

where

$$\rho_{zu}=\frac{S_{zu}}{S_z S_u},\ C_u^2=\frac{S_u^2}{\mu_u^2},\ C_u=\frac{S_u}{\mu_u},\ C_z=\frac{S_z}{\mu_z},$$

$$S_z^2=\frac{\sum_{i=1}^{N}(z_i-\mu_z)^2}{N-1},\ S_u^2=\frac{\sum_{i=1}^{N}(u_i-\mu_z)^2}{N-1}$$

Expressing (11) in terms of the e's:

$$\hat{\mu}_{\mathrm{NH(exp)}}=\frac{\left\{b_1 \bar{Z}(1+e_0)-b_2\mu_u e_1\right\}\exp\left\{-\gamma e_1(1+\gamma e_1)^{-1}\right\}-c}{h}$$

$$=\frac{\left\{b_1 \bar{Z}(1+e_0)-b_2\mu_u e_1\right\}\left\{1-\gamma e_1+\tfrac{3}{2}\gamma^2 e_1^2+\ldots\right\}-c}{h} \tag{12}$$

where

$$\gamma=\frac{\alpha\mu_u}{2(\alpha\mu_u+\beta)}$$

Assuming $|e_1|<1$, expanding the right hand side of (10), and retaining terms up to the second degree of the e's we have

$$\hat{\mu}_{\mathrm{NH(exp)}}-\mu_z \cong \mu_z\left[(b_1-1)+b_1 e_0-b_1\gamma(e_1+e_0 e_1)+\tfrac{3}{2}b_1\gamma^2 e_1^2\right]$$

$$+b_2\mu_u\left(-e_1-\gamma e_1^2\right) \tag{13}$$

Taking the expectation both sides of (13), the Bias Equation of $\hat{\mu}_{\mathrm{NH(exp)}}$ is obtained to the first degree of approximation as

$$\text{Bias}\left(\hat{\mu}_{\text{NH(exp)}}\right) \cong \frac{\mu_z\left\{\left(b_1-1\right)+\lambda b_1\gamma C_u\left(\frac{3}{2}\gamma C_u-\rho_{zu}C_z\right)\right\}+\lambda b_2\mu_u\gamma C_u^2-c}{h} \quad (14)$$

Squaring both sides of (13), retaining terms of the e's up to the second degree and taking the expectation, we get the MSE Equation of $\hat{\mu}_{\text{NH(exp)}}$ to the first degree of approximation as

$$\text{MSE}\left(\hat{\mu}_{\text{NH(exp)}}\right) = \frac{\mu_z^2}{h^2}\left[\left(b_1-1\right)^2 + \lambda b_1^2 A - \lambda\gamma b_1\left\{D+\left(\gamma C_u^2-\rho_{zu}C_zC_u\right)\right\}\right] \\ + \lambda b_2^2\mu_u^2 C_u^2 + 2\lambda b_2\mu_z\mu_u\left[b_1D-\gamma C_u^2\right] \quad (15)$$

where $A = C_z^2 + 4\gamma^2 C_u^2 - 4\gamma\rho_{zu}C_zC_u$, $D = 2\gamma C_u^2 - \rho_{zu}C_zC_u$, and C_u is the coefficient of variation of u.

To minimize $\text{MSE}\left(\hat{\mu}_{\text{NH(exp)}}\right)$, consider the following normal equations:

$$\frac{\partial MSE\left(\hat{\mu}_{\text{NH(exp)}}\right)}{\partial b_i} = 0, \quad i = 1,2$$

On solving these two normal equations simultaneously, the optimum values of b_1 and b_2 are, respectively,

$$b_{1(\text{opt})} = \frac{-C_u^2\left[2-\lambda\gamma D+\lambda\gamma\left(\gamma C_u^2-\rho_{zu}C_zC_u\right)\right]}{2\left[\lambda D^2-\left(1+\lambda A\right)C_u^2\right]} \\ b_{2(\text{opt})} = \frac{\mu_z\left[D\left\{2+\lambda\gamma D+\lambda\gamma\left(\gamma C_u^2-\rho_{zu}C_zC_u\right)\right\}-2\gamma C_u^2\left(1+\lambda A\right)\right]}{2\mu_u\left[\lambda D^2-\left(1+\lambda A\right)C_u^2\right]} \quad (16)$$

On substituting the optimum values of b_1 and b_2 from (15) into (14), the minimum MSE of the proposed estimator $\hat{\mu}_{\text{NH(exp)}}$, up to first order of approximation, is given by

$$\text{MinMSE}\left(\hat{\mu}_{\text{NH(exp)}}\right) = \frac{1}{h^2}\left[\begin{array}{c}\dfrac{\lambda\mu_z^2 C_z^2\left(1-\rho_{zu}^2\right)}{1+\lambda C_z^2\left(1-\rho_{zu}^2\right)}\\[2em] -\dfrac{\lambda^2\mu_z^2\gamma^2 C_u^2\left\{4C_z^2\left(1-\rho_{zu}^2\right)+\gamma^2 C_u^2\right\}}{4\left\{1+\lambda C_z^2\left(1-\rho_{zu}^2\right)\right\}}\end{array}\right] \tag{17}$$

The expressions of c, h, and the MSE and mean equations change depending on the specified models. Two additive models and two multiplicative models are specified. In the first model, M_1, the additive technique is applied for the sensitive variable while the direct technique is utilized for the non-sensitive auxiliary variable: $\{Z = PY + (1 - P)(Y + W),\ U = X\}$. In the second model, M_2, the multiplicative model is applied for the sensitive variable while the direct technique is utilized for the non-sensitive auxiliary variable: $\{Z = PY + (1 - P)(YW),\ U = X\}$. In the third model, M_3, the additive model is applied for both the sensitive variable and the non-sensitive auxiliary variable: $\{Z = PY + (1 - P)(Y + W),\ U = PX + (1 - P)(X + T)\}$. In the fourth model, M_4, the multiplicative model is applied for both the sensitive variable and the non-sensitive auxiliary variable: $\{Z = PY + (1 - P)(YW),\ U = PX + (1 - P)(XT)\}$. In some surveys dealing with sensitive topics, the auxiliary variable that researchers determine to be non-sensitive may be sensitive for respondents. Therefore, in the third model M_3 and fourth model M_4, randomized devices are also used for the auxiliary variable. Mean, variance, and correlation equations, which will be used in MSE equation in (17), are presented in Appendix A according to these four models (Özgül, 2013).

Efficiency Comparisons

A comparison of the proposed estimator with the Diana and Perri (2011) estimator $\hat{\mu}_{\text{DP}}$, the Sousa et al. (2010) estimator $\hat{\mu}_{\text{SR}}$, and the Gupta et al. (2012) estimator $\hat{\mu}_{\text{GRR}}$ is now considered. To compare the efficiencies of the various existing estimators with the proposed estimator, we compare their MSE under the model 1 M_1, in which the respondent is asked to provide true responses for X. The MSEs of estimators under that model with SRSWOR are given below:

$$\text{Var}\left(\hat{\mu}_{\text{DP}}\right)_{\min} = \lambda S_z^2\left(1-\rho_{zx}^2\right) \tag{18}$$

$$\text{MSE}\left(\hat{\mu}_{\text{SR}}\right) = \lambda\mu_z^2\left(C_z^2 + C_x^2 - 2\rho_{xz}C_xC_z\right) \tag{19}$$

$$\text{MSE}\left(\hat{\mu}_{\text{GRR}}\right)_{\min} \cong \frac{\lambda\mu_z^2C_z^2\left(1-\rho_{zx}^2\right)\left(1-\lambda C_x^2\right)}{\lambda C_z^2\left(1-\rho_{zx}^2\right)+\left(1-\lambda C_x^2\right)} \tag{20}$$

$$
\begin{aligned}
\text{MinMSE}\left(\hat{\mu}_{\text{NH(exp)}}\right) &= \left[\begin{array}{l}\dfrac{\lambda\mu_z^2C_z^2\left(1-\rho_{zx}^2\right)}{1+\lambda C_z^2\left(1-\rho_{zx}^2\right)} \\[2mm] \quad -\dfrac{\lambda^2\mu_z^2\gamma^2C_x^2\left\{4C_z^2\left(1-\rho_{zx}^2\right)+\gamma^2C_x^2\right\}}{4\left\{1+\lambda C_z^2\left(1-\rho_{zx}^2\right)\right\}}\end{array}\right] \\[4mm]
&= \lambda\mu_z^2C_z^2\left(1-\rho_{zx}^2\right)-\frac{\lambda^2\mu_z^2C_z^4\left(1-\rho_{zx}^2\right)^2}{1+\lambda C_z^2\left(1-\rho_{zx}^2\right)}-\frac{\lambda^2\mu_z^2\gamma^2C_x^2\left\{4C_z^2\left(1-\rho_{zx}^2\right)+\gamma^2C_x^2\right\}}{4\left\{1+\lambda C_z^2\left(1-\rho_{zx}^2\right)\right\}} \\[4mm]
&= \lambda\mu_z^2C_z^2\left(1-\rho_{zx}^2\right)-\frac{\dfrac{Var\left(\hat{\mu}_{DP}\right)_{\min}^2}{\mu_z^2}}{1+\dfrac{Var\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_z^2}}-\frac{\lambda^2\mu_z^2\gamma^2C_x^2\left\{4C_z^2\left(1-\rho_{zx}^2\right)+\gamma^2C_x^2\right\}}{4\left\{1+\lambda C_z^2\left(1-\rho_{zx}^2\right)\right\}}
\end{aligned} \tag{21}
$$

From (18) and (21),

$$\text{Var}\left(\hat{\mu}_{\text{DP}}\right)_{\min} - \text{MinMSE}\left(\hat{\mu}_{\text{NH(exp)}}\right) = \frac{\dfrac{Var\left(\hat{\mu}_{DP}\right)^{2}_{\min}}{\mu_{z}^{2}}}{1 + \dfrac{Var\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_{z}^{2}}} + \frac{\lambda^{2}\mu_{z}^{2}\gamma^{2}C_{x}^{2}\left\{4C_{z}^{2}\left(1-\rho_{zx}^{2}\right)+\gamma^{2}C_{x}^{2}\right\}}{4\left\{1+\lambda C_{z}^{2}\left(1-\rho_{zx}^{2}\right)\right\}}$$

and so $\text{Var}\left(\hat{\mu}_{\text{DP}}\right)_{\min} - \text{MinMSE}\left(\hat{\mu}_{\text{NH(exp)}}\right) > 0$ always.

From (19) and (21),

$$\text{MSE}\left(\hat{\mu}_{\text{SR}}\right) - \text{MinMSE}\left(\hat{\mu}_{\text{NH(exp)}}\right) = \lambda\mu_{z}^{2}\left[C_{z}^{2}+C_{x}^{2}-2\rho_{xz}C_{x}C_{z}\right] - \lambda\mu_{z}^{2}C_{z}^{2}\left(1-\rho_{zx}^{2}\right)$$

$$+ \frac{\dfrac{Var\left(\hat{\mu}_{DP}\right)^{2}_{\min}}{\mu_{z}^{2}}}{1 + \dfrac{Var\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_{z}^{2}}} + \frac{\lambda^{2}\mu_{z}^{2}\gamma^{2}C_{x}^{2}\left\{4C_{z}^{2}\left(1-\rho_{zx}^{2}\right)+\gamma^{2}C_{x}^{2}\right\}}{4\left\{1+\lambda C_{z}^{2}\left(1-\rho_{zx}^{2}\right)\right\}}$$

$$= \lambda\mu_{z}^{2}\left(C_{x}-2\rho_{xz}C_{z}\right)^{2} + \frac{\dfrac{Var\left(\hat{\mu}_{DP}\right)^{2}_{\min}}{\mu_{z}^{2}}}{1 + \dfrac{Var\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_{z}^{2}}} + \frac{\lambda^{2}\mu_{z}^{2}\gamma^{2}C_{x}^{2}\left\{4C_{z}^{2}\left(1-\rho_{zx}^{2}\right)+\gamma^{2}C_{x}^{2}\right\}}{4\left\{1+\lambda C_{z}^{2}\left(1-\rho_{zx}^{2}\right)\right\}}$$

and so $\text{MSE}\left(\hat{\mu}_{\text{SR}}\right) - \text{MinMSE}\left(\hat{\mu}_{\text{NH(exp)}}\right) > 0$ always.

From (20) and (21),

$$\mathrm{MSE}\left(\hat{\mu}_{\mathrm{GRR}}\right)_{\min} - \mathrm{MinMSE}\left(\hat{\mu}_{\mathrm{NH(exp)}}\right)$$

$$= \frac{\lambda\mu_z^2 C_z^2\left(1-\rho_{zx}^2\right)\left(1-\lambda C_x^2\right)}{\lambda C_z^2\left(1-\rho_{zx}^2\right)+\left(1-\lambda C_x^2\right)} - \lambda\mu_z^2 C_z^2\left(1-\rho_{zx}^2\right) + \frac{\dfrac{\mathrm{Var}\left(\hat{\mu}_{DP}\right)_{\min}^2}{\mu_z^2}}{1+\dfrac{\mathrm{Var}\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_z^2}} + \frac{\lambda^2\mu_z^2\gamma^2 C_x^2\left\{4C_z^2\left(1-\rho_{zx}^2\right)+\gamma^2 C_x^2\right\}}{4\left\{1+\lambda C_z^2\left(1-\rho_{zx}^2\right)\right\}}$$

$$= \frac{\dfrac{4\,\mathrm{Var}\left(\hat{\mu}_{DP}\right)_{\min}^2}{\mu_z^2}\left(1+\mathrm{Var}\left(\hat{\mu}_{DP}\right)_{\min}\right)+\left(\lambda C_x^2-1-\dfrac{\mathrm{Var}\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_z^2}\right)\left(2\dfrac{\mathrm{Var}\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_z}+\lambda\gamma^2\mu_z C_x^2\right)^2}{4\left(1+\dfrac{\mathrm{Var}\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_z^2}\right)^2\left(\lambda C_x^2-1-\dfrac{\mathrm{Var}\left(\hat{\mu}_{DP}\right)_{\min}}{\mu_z^2}\right)}$$

and so $\mathrm{MSE}\left(\hat{\mu}_{\mathrm{GRR}}\right)_{\min} - \mathrm{MinMSE}\left(\hat{\mu}_{\mathrm{NH(exp)}}\right) > 0$ provided that $\left\{\lambda C_x^2-\lambda C_z^2\left(1-\rho_{zx}^2\right)\right\} > 1$.

Simulation Study

A simulation study is presented to show the performance of the proposed estimator in comparison to other estimators using the auxiliary variable for RRT models. The proposed estimator $\hat{\mu}_{\mathrm{NH(exp)}}$ is compared with the Diana and Perri (2011) estimator $\hat{\mu}_{\mathrm{DP}}$, the Sousa et al. (2010) estimator $\hat{\mu}_{\mathrm{SR}}$, and the Gupta et al. (2012) estimator $\hat{\mu}_{\mathrm{GRR}}$. Three finite populations of size 1000 are generated from a multivariate normal distribution. The three populations each have theoretical mean $\mu = [5, 5]$ of $[Y, X]$ and have different covariance matrices. The populations are generated based on correlation levels between the variables. The correlation levels are classified as low, medium and high. The covariance matrices and the correlations are presented below. The scrambling variable W is considered to be a normal random variable with mean equal to zero and standard deviation equal to 0.30. The scrambling variable T is considered to be a normal random variable with mean equal to zero and standard deviation equal to 0.20. We use the simulation studies of Gupta et al. to determine the parameters that are easier to compare.

The covariance matrices and the correlation coefficients for each population are given below:

Population I (Low Correlation):

$$\Sigma_1 = \begin{bmatrix} 9.0 & 5.4 \\ 5.4 & 4.0 \end{bmatrix}, \quad \rho_{yx} = 0.30$$

Population II (Medium Correlation):

$$\Sigma_2 = \begin{bmatrix} 9.0 & 3.6 \\ 3.6 & 4.0 \end{bmatrix}, \quad \rho_{yx} = 0.60$$

Population III (High Correlation):

$$\Sigma_3 = \begin{bmatrix} 9.0 & 5.4 \\ 5.4 & 4.0 \end{bmatrix}, \quad \rho_{yx} = 0.90$$

Table 1. Theoretical and empirical MSEs of estimators according to degree of the correlation between the sensitive and non-sensitive variable for model 1 (M₁)

Population	n	MSE	Estimators			
			$\hat{\mu}_{DP}$	$\hat{\mu}_{SR}$	$\hat{\mu}_{GRR}$	$\hat{\mu}_{NH(exp)}$
I	50	Theoretical	0.1684	0.1894	0.1672	**0.1601**
$\rho_{yx} = 0.30$		Empirical	0.1728	0.1916	0.1754	**0.1705**
	100	Theoretical	0.0838	0.0942	0.0828	**0.0782**
		Empirical	0.0840	0.0945	0.0840	**0.0808**
	200	Theoretical	0.0415	0.0466	0.0414	**0.0398**
		Empirical	0.0408	0.0465	0.0411	**0.0388**
	300	Theoretical	0.0274	0.0308	0.0273	**0.0268**
		Empirical	0.0272	0.0308	0.0272	**0.0262**
II	50	Theoretical	0.1197	0.1203	0.1191	**0.0972**
$\rho_{yx} = 0.60$		Empirical	0.1187	0.1191	0.1187	**0.0982**
	100	Theoretical	0.0595	0.0599	0.0594	**0.0494**
		Empirical	0.0608	0.0613	0.0610	**0.0498**
	200	Theoretical	0.0295	0.0296	0.0291	**0.0239**
		Empirical	0.0300	0.0302	0.0297	**0.0244**
	300	Theoretical	0.0194	0.0196	0.0194	**0.0162**
		Empirical	0.0193	0.0193	0.0193	**0.0158**
III	50	Theoretical	0.0358	0.0472	0.0358	**0.0058**
$\rho_{yx} = 0.90$		Empirical	0.0372	0.0480	0.0374	**0.0060**
	100	Theoretical	0.0178	0.0235	0.0178	**0.0098**
		Empirical	0.0186	0.0239	0.0186	**0.0100**
	200	Theoretical	0.0088	0.0116	0.0088	**0.0033**
		Empirical	0.0091	0.0120	0.0091	**0.0014**
	300	Theoretical	0.0058	0.0077	0.0058	**0.0010**
		Empirical	0.0060	0.0079	0.0060	**0.0010**

Table 2. Theoretical and empirical MSEs of estimators according to degree of the correlation between the sensitive and non-sensitive variable for model 2 (M₂)

Population	n	MSE	Estimators			
			$\hat{\mu}_{DP}$	$\hat{\mu}_{SR}$	$\hat{\mu}_{GRR}$	$\hat{\mu}_{NH(exp)}$
I	50	Theoretical	3.5609	3.5628	3.1889	**3.1872**
$\rho_{yx} = 0.30$		Empirical	2.6095	2.5585	2.2517	**2.2515**
	100	Theoretical	1.6867	1.6877	1.5985	**1.5980**
		Empirical	1.5985	1.6009	1.3017	**1.3015**
	200	Theoretical	0.7497	0.7501	0.7317	**0.7316**
		Empirical	1.0440	1.0409	0.8691	**0.8622**
	300	Theoretical	0.4373	0.4376	0.4311	**0.4310**
		Empirical	0.8338	0.8312	0.7260	**0.7259**

Table 2, continued.

Population	n	MSE	$\hat{\mu}_{DP}$	$\hat{\mu}_{SR}$	$\hat{\mu}_{GRR}$	$\hat{\mu}_{NH(exp)}$
					Estimators	
II	50	Theoretical	3.1069	3.1073	2.8349	**2.8334**
$\rho_{yx} = 0.60$		Empirical	2.1729	2.1638	1.9602	**1.9517**
	100	Theoretical	1.4717	1.4719	1.4078	**1.4073**
		Empirical	1.3937	1.3960	1.1353	**1.1316**
	200	Theoretical	0.6541	0.6542	0.6412	**0.6410**
		Empirical	0.9189	0.9192	0.7682	**0.7624**
	300	Theoretical	0.3816	0.3816	0.3772	**0.3770**
		Empirical	0.7939	0.7964	0.6994	**0.6898**
III	50	Theoretical	2.8921	2.9277	2.6571	**2.6557**
$\rho_{yx} = 0.90$		Empirical	2.0042	2.0091	1.8101	**1.7838**
	100	Theoretical	1.3699	1.3868	1.3150	**1.3145**
		Empirical	1.3401	1.3760	1.0960	**1.0958**
	200	Theoretical	0.6089	0.6164	0.5978	**0.5976**
		Empirical	0.9457	0.9552	0.8018	**0.7990**
	300	Theoretical	0.3552	0.3596	0.3514	**0.3512**
		Empirical	0.8639	0.8687	0.7715	**0.7684**

Table 3. Theoretical and empirical MSEs of estimators according to degree of the correlation between the sensitive and non-sensitive variable for model 3 (M3)

Population	n	MSE	$\hat{\mu}_{DP}$	$\hat{\mu}_{SR}$	$\hat{\mu}_{GRR}$	$\hat{\mu}_{NH(exp)}$
					Estimators	
I	50	Theoretical	0.1670	0.1883	0.1659	**0.1603**
$\rho_{yx} = 0.30$		Empirical	0.1722	0.1905	0.1740	**0.1698**
	100	Theoretical	0.0831	0.0937	0.0835	**0.0780**
		Empirical	0.0835	0.0938	0.0847	**0.0798**
	200	Theoretical	0.0411	0.0464	0.0411	**0.0394**
		Empirical	0.0405	0.0462	0.0408	**0.0382**
	300	Theoretical	0.0271	0.0306	0.0271	**0.0257**
		Empirical	0.0270	0.0307	0.0271	**0.0257**
II	50	Theoretical	0.1183	0.1191	0.1178	**0.0964**
$\rho_{yx} = 0.60$		Empirical	0.1188	0.1180	0.1173	**0.0978**
	100	Theoretical	0.0589	0.0593	0.0588	**0.0484**
		Empirical	0.0606	0.0605	0.0600	**0.0488**
	200	Theoretical	0.0291	0.0293	0.0294	**0.0238**
		Empirical	0.0297	0.0298	0.0300	**0.0242**
	300	Theoretical	0.0192	0.0194	0.0197	**0.0157**
		Empirical	0.0191	0.0191	0.0193	**0.0157**

Table 3, continued.

Population	n	MSE	$\hat{\mu}_{DP}$	$\hat{\mu}_{SR}$	$\hat{\mu}_{GRR}$	$\hat{\mu}_{NH(exp)}$
					Estimators	
III	50	Theoretical	0.0357	0.0459	0.0357	**0.0056**
$\rho_{yx} = 0.90$		Empirical	0.0375	0.0466	0.0370	**0.0061**
	100	Theoretical	0.0178	0.0229	0.0178	**0.0096**
		Empirical	0.0188	0.0233	0.0186	**0.0100**
	200	Theoretical	0.0088	0.0113	0.0088	**0.0031**
		Empirical	0.0090	0.0116	0.0090	**0.0017**
	300	Theoretical	0.0058	0.0075	0.0058	**0.0028**
		Empirical	0.0059	0.0077	0.0060	**0.0028**

Table 4. Theoretical and empirical MSEs of estimators according to degree of the correlation between the sensitive and non-sensitive variable for model 4 (M$_4$)

Population	n	MSE	$\hat{\mu}_{DP}$	$\hat{\mu}_{SR}$	$\hat{\mu}_{GRR}$	$\hat{\mu}_{NH(exp)}$
					Estimators	
I	50	Theoretical	2.0340	2.8564	1.8945	**1.8568**
$\rho_{yx} = 0.30$		Empirical	2.6028	3.3407	2.2134	**1.9528**
	100	Theoretical	0.9635	1.3530	0.9327	**0.9223**
		Empirical	1.4757	1.6144	1.2717	**1.1878**
	200	Theoretical	0.4282	0.6013	0.4222	**0.4199**
		Empirical	0.8659	0.8751	0.7673	**0.7651**
	300	Theoretical	0.2498	0.3508	0.2478	**0.2469**
		Empirical	0.6889	0.6764	0.6284	**0.6265**
II	50	Theoretical	1.6396	2.4315	1.5521	**1.5146**
$\rho_{yx} = 0.60$		Empirical	2.2175	2.9569	1.8395	**1.6508**
	100	Theoretical	0.7767	1.1518	0.7577	**0.7479**
		Empirical	1.3376	1.6164	1.1199	**1.0010**
	200	Theoretical	0.3452	0.5119	0.3415	**0.3394**
		Empirical	0.8464	0.9276	0.7488	**0.7016**
	300	Theoretical	0.2014	0.2990	0.2001	**0.1993**
		Empirical	0.7073	0.7303	0.6484	**0.6456**
III	50	Theoretical	1.3325	2.2892	1.2748	**1.2421**
$\rho_{yx} = 0.90$		Empirical	1.7606	2.1009	1.4869	**1.4323**
	100	Theoretical	0.6312	1.0843	0.6187	**0.6104**
		Empirical	1.1714	1.2880	1.0017	**0.9817**
	200	Theoretical	0.2806	0.4820	0.2781	**0.2763**
		Empirical	0.7664	0.7792	0.6878	**0.6821**
	300	Theoretical	0.1637	0.2811	0.1628	**0.1622**
		Empirical	0.6670	0.6459	0.6185	**0.6139**

The process was repeated 5000 times and for different sample sizes: $n = 50$, 100, 200, and 300. The value of the design parameter P changes from 0.10 to 0.90 with an increment of 0.1. We observe small differences in efficiency with almost each value of the design parameter when an auxiliary variable is utilized in RRT models. Thus, simulation results are only presented for $P = 0.20$. That means 20 percent of the respondents gave direct answers; the rest of the respondents use the randomized devices. The performances of the estimators are measured by the simulated MSE:

$$\text{MSE}\left(\hat{\mu}\right) = \frac{1}{5000} \sum_{i=1}^{5000} \left(\hat{\mu}_i - \mu_y\right)^2$$

where $\hat{\mu}_i$ is the estimate of μ_y on the ith sample. Simulation results are summarized in Tables 1-4.

In Tables 1-4, theoretical and empirical MSE values of the estimators, according to degree of the correlation between the sensitive and non-sensitive variables, are given for the four specified models. In all circumstances, regardless of both degree of correlation and sample size, the proposed estimator is always more efficient than the Diana and Perri (2011) estimator $\hat{\mu}_{\text{DP}}$, the Sousa et al. (2010) estimator $\hat{\mu}_{\text{SR}}$, and the Gupta et al. (2012) estimator $\hat{\mu}_{\text{GRR}}$. The MSE values of the estimators are smaller when the sample size increases, and that is an expected result. However, additive models performed better than multiplicative models. When additive models are applied in RRT, more efficient estimates are obtained.

Application

To test the models and show the performance of the proposed estimator in comparison to other estimators, a survey was performed at the Hacettepe University Department of Statistics to estimate the grade point average (GPA) of students who graduated in 2016. One hundred and two students who graduated in 2016 are considered as our population. In this application, the study variable Y is the GPA of students, the auxiliary variable X is study hours per week. Four models for $P = 0.20$ were applied to the population. Twenty students were requested to report their true GPA, and 82 students used the randomized devices. To apply the randomized devices, random numbers were generated for scrambling variables W and T. For scrambling variable W, 82 random numbers were

generated from the normal distribution with mean equal to zero and standard deviation equal to 0.60. For scrambling variable *T*, 82 random numbers were generated from the normal distribution with mean equal to zero and standard deviation equal to 0.20.

The following are some characteristics of the population:

$$\bar{Y} = 2.51,\ \bar{X} = 7.16,\ S_y^2 = 0.1166,\ S_x^2 = 38.53,\ \rho_{yx} = 0.71$$

Table 5. Theoretical Bias and MSE values of the estimators by using non-sensitive auxiliary variable according to Models

Model	Estimators	n = 50		n = 100		n = 200	
		Bias	MSE	Bias	MSE	Bias	MSE
M_1	$\hat{\mu}_{DP}$	--	0.0155	--	0.0091	--	0.0039
	$\hat{\mu}_{SR}$	0.0563	0.1205	0.0329	0.0705	0.0143	0.0305
	$\hat{\mu}_{GRR}$	0.0061	0.0155	0.0035	0.0091	0.0015	0.0039
	$\hat{\mu}_{NH(exp)}$	0.0041	**0.0151**	0.0024	**0.0089**	0.0011	**0.0038**
M_2	$\hat{\mu}_{DP}$	--	2.3298	--	1.3638	--	0.5909
	$\hat{\mu}_{SR}$	0.1290	2.3519	0.0151	1.3767	0.0054	0.5966
	$\hat{\mu}_{GRR}$	0.5661	1.9313	0.3571	1.2184	0.0337	0.5622
	$\hat{\mu}_{NH(exp)}$	0.5548	**1.9212**	**0.3504**	1.2134	0.0330	**0.5609**
M_3	$\hat{\mu}_{DP}$	--	0.0143	--	0.0084	--	0.0091
	$\hat{\mu}_{SR}$	0.0604	0.1270	0.0354	0.0743	0.0153	0.0322
	$\hat{\mu}_{GRR}$	0.0056	0.0143	0.0033	0.0084	0.0014	0.0036
	$\hat{\mu}_{NH(exp)}$	0.0039	**0.0139**	0.0023	**0.0082**	0.0010	**0.0035**
M_4	$\hat{\mu}_{DP}$	--	2.0248	--	1.1852	--	0.5136
	$\hat{\mu}_{SR}$	0.0474	2.0475	0.0277	1.1985	0.0121	0.5194
	$\hat{\mu}_{GRR}$	0.4974	1.6970	0.3135	1.0696	0.1440	0.4914
	$\hat{\mu}_{NH(exp)}$	0.9625	**0.3203**	0.2991	**0.5629**	0.0276	**0.3902**

Note: Blank cells indicate unbiased estimators.

To compute the Bias and MSE values of the Diana and Perri (2011) estimator $\hat{\mu}_{\text{DP}}$, the Sousa et al. (2010) estimator $\hat{\mu}_{\text{SR}}$, the Gupta et al. (2012) estimator $\hat{\mu}_{\text{GRR}}$, and the proposed estimator $\hat{\mu}_{\text{NH(exp)}}$ for the four models based on different sample sizes: $n = 20$, 30, and 50, arbitrarily take $\alpha = 1$ and $\beta = -1$, that is

$$\gamma = \frac{\mu_x}{2(\mu_x - 1)}$$

for simplicity. The results are summarized in Table 5.

In the application study, the most efficient estimator was the proposed exponential-type estimator. It was always more efficient than the existing estimators in all RRT models for different sample sizes. From Table 5, it can be concluded that the additive models were more efficient than the multiplicative models and that the proposed estimator gave better results.

Conclusion

An exponential-type estimator was proposed, based on a non-sensitive auxiliary variable, for the population mean of a sensitive variable for Generalized Quantitative RRT models. The MSE equation is derived for all Quantitative RRT models. The proposed estimator was more efficient than other existing estimators in all circumstances, regardless of which model was applied. It was shown that the efficiency of the proposed estimator can be quite substantial if the correlation between the study and the auxiliary variables is high. Additionally, the additive models were more efficient than the multiplicative models. These results were supported by simulation and application studies. In a future work, an estimator will be developed for the population mean of the sensitive study variable by combining additive and multiplicative techniques based on Quantitative RTT using multi-sensitive auxiliary variables.

References

Bahl, S., & Tuteja, R. K. (1991). Ratio and product type exponential estimators. *Journal of Information and Optimization Sciences, 12*(1), 159-164. doi: 10.1080/02522667.1991.10699058

Bar-Lev, S. K., Bobovitch, E., & Boukai, B. (2004). A note on randomized response models for quantitative data. *Metrika, 60*(3), 255-260. doi: 10.1007/s001840300308

Cingi, H., & Kadilar, C. (2009). *Advances in sampling theory – Ratio method of estimation*. Bentham Science Publishers. doi: 10.2174/97816080501231090101

Diana, G. & Perri, P. F. (2011). A class of estimators for quantitative sensitive data. *Statistical Papers, 52*(3), 633-650. doi: 10.1007/s00362-009-0273-1

Eichhorn, B. H., & Hayre, L. S. (1983). Scrambled randomized response methods for obtaining sensitive quantitative data. *Journal of Statistical Planning and Inference, 7*(4), 307-316. doi: 10.1016/0378-3758(83)90002-2

Grover, L. K., & Kaur, P. (2014). A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable. *Communications in Statistics – Simulation and Computation, 43*(7), 1552-1574. doi: 10.1080/03610918.2012.736579

Gupta, S., Shabbir, J., Sousa, R., & Real, P. C. (2012). Estimation of the mean of a sensitive variable in the presence of auxiliary information. *Communications in Statistics – Theory and Methods, 41*(13-14), 1-12. doi: 10.1080/03610926.2011.641654

Özgül, N. (2013). *Proportion and mean estimators in randomized response models* (Unpublished doctoral thesis). Hacettepe University, Ankara, Turkey.

Özgül, N. & Cingi, H. (2014). A new class of exponential regression cum ratio estimator in two phase sampling. *Hacettepe Journal of Mathematics and Statistics, 43*(1), 131-140. Available from http://dergipark.ulakbim.gov.tr/hujms/article/view/5000017145

Shabbir, J., & Gupta, S. (2011). On estimating finite population mean in simple and stratified random sampling. *Communications in Statistics – Theory and Techniques, 40*(2), 199-212. doi: 10.1080/03610920903411259

Sousa, R., Shabbir, J., Real, P. C., & Gupta, S. (2010). Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information. *Journal of Statistical Theory and Practice, 4*(3), 495-507. 10.1080/15598608.2010.10411999

Thornton, B., & Gupta, S. N. (2004). Comparative validation of a partial (versus full) randomized response technique: Attempting to control for social

desirability response bias to sensitive questions. *Individual Differences Research, 2*(3), 214-224.

    Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association, 60*(309), 63-69. doi: 10.2307/2283137

    Warner, S. L. (1971). The linear randomized response model. *Journal of the American Statistical Association, 66*(336) 884-888. doi: 10.2307/2284247

Appendix A: Special Models for Generalized RTT

First Model (**M₁**): $S = Y + W$, $R = X$, Mean given by

$$\mu_z = \mu_y + (1-P)\mu_w$$
$$\mu_u = \mu_x$$
$$c = (1-P)\mu_w, \, h = 1$$

Variance and correlation equations to be used in (17) are given by

$$S_z^2 = S_y^2 + (1-P)\mu_w^2 \left(C_w^2 + P \right)$$
$$S_u^2 = S_x^2$$
$$\rho_{zu} = \frac{S_{yx}}{S_x \sqrt{S_y^2 + (1-P)\mu_w^2 \left(C_w^2 + P \right)}}$$

Second Model (**M₂**): $S = YW$, $R = X$, Mean given by

$$\mu_z = \left\{ P + (1-P)\mu_w \right\} \mu_y$$
$$\mu_u = \mu_x$$
$$c = 0, \, h = \left\{ P + (1-P)\mu_w \right\}$$

385

Variance and correlation equations to be used in (17) are given by

$$S_z^2 = \mu_y^2 \left(1+C_y^2\right)\left\{P+\left(1-P\right)\mu_w^2\left(1+C_w^2\right)\right\} - \mu_z^2$$

$$S_u^2 = S_x^2$$

$$\rho_{zu} = \frac{S_{yx}\left\{P+\left(1-P\right)\mu_w\right\}}{S_x\sqrt{\mu_y^2\left(1+C_y^2\right)\left\{P+\left(1-P\right)\mu_w^2\left(1+C_w^2\right)\right\} - \mu_z^2}}$$

Third Model (**M₃**): $S = Y + W$, $R = X + T$, Mean given by

$$\mu_z = \mu_y + \left(1-P\right)\mu_w$$

$$\mu_u = \mu_x + \left(1-P\right)\mu_t$$

$$c = \left(1-P\right)\mu_w, \; h = 1$$

Variance and correlation equations to be used in (17) are given by

$$S_z^2 = S_y^2 + \left(1-P\right)\mu_w^2\left(C_w^2 + P\right)$$

$$S_u^2 = S_x^2 + \left(1-P\right)\mu_t^2\left(C_t^2 + P\right)$$

$$\rho_{zu} = \frac{S_{yx}+P\left(1-P\right)\mu_w\mu_t}{S_x\sqrt{\left\{S_y^2+\left(1-P\right)\mu_w^2\left(C_w^2+P\right)\right\}\left\{S_x^2+\left(1-P\right)\mu_t^2\left(C_t^2+P\right)\right\}}}$$

Fourth Model (M₄): $S = YW$, $R = XT$, Mean given by

$$\mu_z = \left\{ P + (1-P)\mu_w \right\} \mu_y$$

$$\mu_u = \left\{ P + (1-P)\mu_t \right\} \mu_x$$

$$c = 0, \; h = P + (1-P)\mu_w$$

Variance and correlation equations to be used in (17) are given by

$$S_z^2 = \mu_y^2 \left(1 - C_y^2\right) \left[P + (1-P)\mu_w^2 \left(1 + C_w^2\right) \right] - \mu_z^2$$

$$S_u^2 = \mu_x^2 \left(1 + C_x^2\right) \left[P + (1-P)\mu_t^2 \left(1 + C_t^2\right) \right] - \mu_u^2$$

$$\rho_{zu} = \frac{\left\lfloor \sigma_{yx} \left\{ P + (1-P)\mu_t \mu_w \right\} + P(1-P)\mu_y \mu_x \left(1 - \mu_w\right)\left(1 - \mu_t\right) \right\rfloor}{\sqrt{\left[\mu_y^2 \left(1 - C_y^2\right) \left[P + (1-P)\mu_w^2 \left(1 + C_w^2\right) \right] - \mu_z^2 \right] \left[\mu_x^2 \left(1 + C_x^2\right) \left[P + (1-P)\mu_t^2 \left(1 + C_t^2\right) \right] - \mu_u^2 \right]}}$$