JIRSS (2022)
Vol. 21, No. 02, pp 89-110
DOI:10.22034/jirss.2022.707639

An Improved Two-Stage Randomized Response Model for Estimating the Mean of a Quantitative Sensitive Random Variable

Rawan Arafa [1], **Reda Mazloum** [1]

[1] Department of Statistics, Faculty of Economics and Political Science, Cairo University, Cairo, Egypt.

Received: 09/04/2022, Accepted: 15/09/2023, Published online: 14/01/2024

Abstract. This paper introduces a new two-stage randomized response model for estimating the mean of a sensitive quantitative random variable. The proposed model is obtained for both simple and stratified random sampling. The efficiency of the proposed estimator, under both sampling schemes, is investigated with respect to various estimators and it is found to be more efficient. Moreover, a new measure for evaluating the performance of any randomized response estimator is introduced. The measure considers the relative efficiency of the randomized response estimator, and the privacy protection it offers. The performance of the proposed estimator is examined using the new measure and it is found to have an overall better performance than its rival estimators. A real data example is also examined using the proposed model and various models from literature.

Keywords. Efficiency, Privacy Protection, Scrambling Variable, Simple Random Sampling, Stratified Random Sampling.

MSC: 62D05.

1 Introduction

The first randomized response model was designed by Warner (1965). His aim was to encourage truthful responses to sensitive questions such as those relating to tax evasion, sexual tendencies, or drug usage. Warner designed his model to estimate the percentage of population belonging to a sensitive group, for qualitative binary

Corresponding Author :Rawan Arafa (rawan.ebrahim@feps.edu.eg),
Reda Mazloum (rmazloum@feps.edu.eg).

variables.

Shortly after Warner (1965), many researchers extended the randomized response technique to cover different data types and/or different sampling schemes. Greenberg et al. (1971) developed a randomized response model to estimate the mean of a quantitative random variable using an unrelated question for randomization. Eichhorn and Hayre (1983) studied the multiplicative model, where the true response is multiplied by an independent scrambling variable to conceal the true value. Bar-Lev et al. (2004) used Warner's (1965) randomization device along with the multiplicative model. Their model gives a probability p to answering the sensitive question truthfully and $1 - p$ to answering with the concealed value of the multiplicative model.

Ryu et al. (2005) suggested a two-stage randomized response model. According to their model, each respondent in a simple random sample with replacement (SRSWR) of size n is provided with two random devices R_1 and R_2. The first randomization device, R_1, has two statements (i) report your true response X for the sensitive question, and (ii) go to R_2, with probabilities p and $1-p$, respectively. The second randomization device, R_2, has two statements (i) report your true response X for the sensitive question, and (ii) report the scrambled response XS, with probabilities t and $1-t$, respectively. Given that $X \geq 0$, $S > 0$, and $P(S = 1) = 0$, the respondents will never have to report their true value for the sensitive question. Let μ_S and σ_S^2 be the known mean and variance of S, and let μ_X and σ_X^2 be the mean and variance of X, respectively.

Assuming that $\mu_S = 1$, they obtained an unbiased estimator, $\hat{\mu}_R$, of μ_X with variance given by:

$$V(\hat{\mu}_R) = \frac{\mu_X^2}{n}[C_X^2 + (1-p)(1-t)\sigma_S^2(1 + C_X^2)], \tag{1.1}$$

where $C_X = \sigma_X/\mu_X$ is the coefficient of variation of X.

They compared their estimator to that of Greenberg et al. (1971) and Gupta et al. (2002) and found theirs to be more efficient.

They also extended their model to the stratified random sampling under the condition $\mu_{S_h} = 1$, where $h = 1, ..., k$. They obtained an unbiased estimator, $\hat{\mu}_R^S$, of μ_X with variance given by:

$$V(\hat{\mu}_R^S) = \sum_{h=1}^{k} \frac{w_h^2 \mu_{X_h}^2}{n_h}\Big[C_{X_h}^2 + (1-p_h)(1-t_h)\sigma_{S_h}^2(1 + C_{X_h}^2)\Big], \tag{1.2}$$

where $w_h = N_h/N$ is the weight of the h^{th} stratum in the population; N_h is the size of stratum h; N is the total population size, and $C_{X_h}^2$, p_h, t_h, μ_{X_h}, n_h, $\sigma_{S_h}^2$ are as defined before but for stratum h.

They also obtained the following variance in the case of Neyman's optimal allocation:

$$V_{Ney}(\hat{\mu}_R^S) = \frac{1}{n}\Big(\sum_{h=1}^{k} w_h \mu_{X_h} \sqrt{C_{X_h}^2 + (1-p_h)(1-t_h)\sigma_{S_h}^2(1 + C_{X_h}^2)}\,\Big)^2. \tag{1.3}$$

Singh and Gorey (2019) developed a model whose response is given by:

$$Z = \begin{cases} X(S - p(\mu_S - 1)), & \text{with probability } 1 - p, \\ X((1 - p)\mu_S + p), & \text{with probability } p. \end{cases} \quad (1.4)$$

They obtained an unbiased estimator of μ_X, $\hat{\mu}_{SG}$, with variance equal to:

$$V(\hat{\mu}_{SG}) = \frac{\mu_X^2}{n}\left[C_X^2 + (C_S^*(p) - \psi)(1 + C_X^2)\right], \quad (1.5)$$

where

$$C_S^*(p) = \frac{p + (1 - p)(\mu_S^2 + \sigma_S^2)}{(p + \mu_S(1 - p))^2} - 1, \quad (1.6)$$

and

$$\psi = \frac{p(1 - p)(\mu_S - 1)^2}{[p + (1 - p)\mu_S]^2}. \quad (1.7)$$

They proved that their estimator is more efficient than Odumade and Singh (2009) estimator. Odumade and Singh (2009) proved that their estimator was more efficient than that of Bar-Lev et al. (2004).

In the next section, the proposed model for the estimation of the mean of a sensitive quantitative random variable in case of simple random sampling is introduced and its estimator's efficiency is investigated. Particularly, in subsection 2.1, the model and its estimator are introduced and applied on a real secondary data. In subsection 2.3, Ryu et al. (2005) model is generalized for the case of $\mu_S > 0$ and the proposed model proves to be more efficient for $\mu_S > 1$. In subsection 2.4, the proposed estimator is shown to be more efficient than the estimators of Singh and Gorey (2019), Odumade and Singh (2009), and Bar-Lev et al. (2004) and more efficient than Eichhorn and Hayre (1983) estimator under an achievable condition.

In section 3, the proposed model is extended to stratified random sampling. Specifically, subsection 3.1 introduces the proposed stratified model in the general case and in case of Neyman allocation. In subsection 3.2, the stratified Ryu et al. (2005) estimator is generalized for the case of $\mu_{S_h} > 0$, and the proposed stratified estimator is shown to be more efficient at $\mu_{S_h} > 1$. The proposed stratified estimator is compared to that of Singh and Gorey (2019) and it is found to be more efficient, in subsection 3.3.

Section 4 presents a new measure of the overall performance of the estimators, in terms of efficiency and privacy protection. Numerical comparisons between the proposed estimator and each of its competitors were made once using variations of the different models' parameters and another time using real data. The proposed estimator proves to have better performance than its competitors. Section 5 provides a conclusion for the effectiveness of the proposed estimator.

2 Proposed Model in Case of Simple Random Sampling

2.1 Model Layout and Estimator Properties

The proposed model modifies Singh and Gorey (2019) model using a two-stage quantitative randomized response model. According to the proposed model, each respondent in a SRSWR of size n is provided with two randomization devices R_1 and R_2. The first randomization device, R_1, has two statements (i) report the response (Xb), and (ii) go to R_2, with probabilities t and $1-t$, respectively. The second randomization device, R_2, has two statements (i) report the response (Xb), and (ii) report the scrambled response (XY), with probabilities p and $1-p$, respectively. The responses of respondents can be presented as follows:

$$Z = \begin{cases} Xb, & \text{with probability } t, \\ \text{go to } R_2, & \text{with probability } 1-t, \quad \begin{cases} Xb, & \text{with probability } p, \\ XY, & \text{with probability } 1-p, \end{cases} \end{cases} \tag{2.1}$$

where $p, t \in (0,1)$, and

$$Y = S - p(\mu_s - 1), \tag{2.2}$$

is a scrambling variable, independent from the sensitive variable X, with a known constant mean, b, given by:

$$b = (1-p)\mu_S + p. \tag{2.3}$$

The model reduces to that of Ryu et al. (2005) at $\mu_S = 1$, to Singh and Gorey (2019) at $p = 0$ or $t = 0$, and to a direct response question at $p = 1$.

The proposed estimator for the mean of the sensitive variable, μ_X, is given by:

$$\hat{\mu}_P = \frac{\bar{Z}}{b}, \tag{2.4}$$

where $\bar{Z}$ is the moment estimator of the mean of the responses.

Theorem 2.1. *$\hat{\mu}_P$, given in Equation 2.4, is an unbiased estimator for μ_X with variance equal to:*

$$V(\hat{\mu}_P) = \frac{\mu_X^2}{n}[C_X^2 + (1 + C_X^2)\gamma], \tag{2.5}$$

where

$$\gamma = \frac{(1-t)(1-p)\sigma_S^2}{b^2}, \tag{2.6}$$

and b is as defined in Equation 2.3.

Proof.

$$E(\hat{\mu}_P) = \frac{E(Z)}{b} = \frac{b\mu_X}{b} = \mu_X,$$

$\therefore$ $\hat{\mu}_P$ is unbiased.

$$V(\hat{\mu}_P) = \frac{V(Z)}{nb^2}. \tag{2.7}$$

It is easy to show that:

$$E(Z^2) = E(X^2)(b^2 + (1-t)(1-p)\sigma_S^2).$$

Consequently,

$$\begin{aligned} V(Z) &= E(X^2)\Big((b^2 + (1-t)(1-p)\sigma_S^2\Big) - b^2\mu_X^2, \\ &= \mu_X^2\Big((1 + C_X^2)[b^2 + (1-t)(1-p)\sigma_S^2] - b^2\Big). \end{aligned} \tag{2.8}$$

Substituting Equation 2.8 into Equation 2.7, yields the variance in Equation 2.5. □

2.2 Real Data Example

To show that the proposed model is practical, we applied the proposed model on a real data set from the U.S. Census Bureau's 2017 Current Population Survey Annual Social and Economic Supplement (CPS ASEC). Among other variables, the survey contains information on the annual total wage and salary earnings for individuals (WSAL_VAL), given as a dollar amount. Questions about finances are usually sensitive in nature and can illicit untruthful answers. Therefore, the variable WSAL_VAL is treated as our sensitive variable X.

The survey contains data on 87,689 persons whose sources of earnings are wages and/or salary. We treated the data from the 87,689 persons as a population with $\mu_X = \$51390.83$. A SRSWR of size $n = 100$ was drawn from that population. For the i^{th} person in the sample, $i = 1, ..., 100$, the randomized response, z_i, based on the proposed model was generated as follows:

(i) 100 random observations, s_i, were generated from the $F(1, 5)$ distribution. The $F(1, 5)$ distribution has mean $\mu_S = 5/3$.

(ii) 100 observations were generated from two Bernoulli random variables, L and Q, with probability of success, $p = t = 0.5$. These two variables mimic the job of the randomization devices R_1 and R_2.

$$L = \begin{cases} 1, & \text{w.p} \quad 1/2, \\ 0, & \text{w.p} \quad 1/2, \end{cases} \tag{2.9}$$

$$Q = \begin{cases} 1, & \text{w.p} \quad 1/2, \\ 0, & \text{w.p} \quad 1/2. \end{cases} \tag{2.10}$$

In real surveys, the two randomization devices can be as simple as a fair coin.

(iii) The constant $b = p + (1-p)\mu_S = 0.5 + 0.5(5/3) = 4/3$.

(iv) The observations of the scrambling variable Y were calculated, as in Equation 2.2.

(v) The response for the i^{th} person was obtained as:

$$z_i = q_i\, x_i b + (1 - q_i)\Big(l_i\, x_i b + (1 - l_i)\, x_i\, y_i\Big). \tag{2.11}$$

Using Equation 2.4, the estimate of the mean of the sensitive variable, WSAL_VAL, is $\hat{\mu}_P$ = \$53506.4.

2.3 Efficiency Comparison with Generalized Ryu et al. Estimator

As mentioned before, Ryu et al. (2005) obtained their estimator and its variance for the case of $\mu_S = 1$ only. The proposed model reduces to that of Ryu et al. (2005) for $\mu_S = 1$. To compare the efficiency of the proposed estimator to that of Ryu et al. (2005) in the general case, we obtain their estimator for any $\mu_S > 0$ along with its variance. We call the estimator in this case generalized Ryu et al. and denote it by $\hat{\mu}^*_R$.

The variance of the generalized Ryu et al. estimator is:

$$V(\hat{\mu}^*_R) = \frac{\mu_X^2}{n}[C_X^2 + (1 + C_X^2)C_{S,p,t}^2], \tag{2.12}$$

where

$$C_{S,p,t}^2 = \frac{p + (1-p)t + (1-p)(1-t)(\mu_S^2 + \sigma_S^2)}{c^2} - 1, \tag{2.13}$$

and

$$c = p + (1-p)t + (1-p)(1-t)\mu_S. \tag{2.14}$$

Theorem 2.2. *The proposed estimator is more efficient than that of the generalized Ryu et al. estimator for* $\mu_S > 1$.

Proof. The difference between the variance of the generalized Ryu et al. estimator, given in Equation 2.12, and the variance of the proposed estimator, given in Equation 2.5, is as follows:

$$V(\hat{\mu}^*_R) - V(\hat{\mu}_P) = \frac{\mu_X^2(1 + C_X^2)}{n}[C_{S,p,t}^2 - \gamma], \tag{2.15}$$

where $C_{S,p,t}^2$ is as given in Equation 2.13 and γ is as defined in Equation 2.6. It suffices to show that $C_{S,p,t}^2 - \gamma > 0$ for $\mu_S > 1$ as follows:

$$\begin{aligned}
C_{S,p,t}^2 - \gamma =& \frac{p + (1-p)t + (1-p)(1-t)(\mu_S^2 + \sigma_S^2)}{c^2} - 1 - \frac{(1-p)(1-t)\sigma_S^2}{b^2} \\
=& \frac{(1-p)(1-t)\Big[\sigma_S^2 + (1-\mu_S)^2(p + (1-p)t)\Big]}{c^2} - \frac{(1-p)(1-t)\sigma_S^2}{b^2} \\
=& (1-p)(1-t)\sigma_S^2\Big[\frac{1}{c^2} - \frac{1}{b^2}\Big] + \frac{(1-p)(1-t)\sigma_S^2(1-\mu_S)^2(p + (1-p)t)}{c^2} \\
=& (1-p)(1-t)\sigma_S^2\Big[\frac{(1-p)t(\mu_S - 1)(b + c)}{c^2 b^2}\Big] + \frac{(1-p)(1-t)\sigma_S^2(1-\mu_S)^2(p + (1-p)t)}{c^2},
\end{aligned}$$

where c and b are as defined in Equation 2.14 and Equation 2.3. Therefore, the difference is guaranteed to be positive for $\mu_S > 1$ and it is zero for $\mu_S = 1$. □

The behavior for $\mu_S < 1$ was investigated through numerical calculations. We considered the case where $p = t$; μ_S and C_S both range between 0.1 and 0.9 with a step of 0.2, i.e., μ_S and $C_S \in \{0.1, (0.2), 0.9\}$.

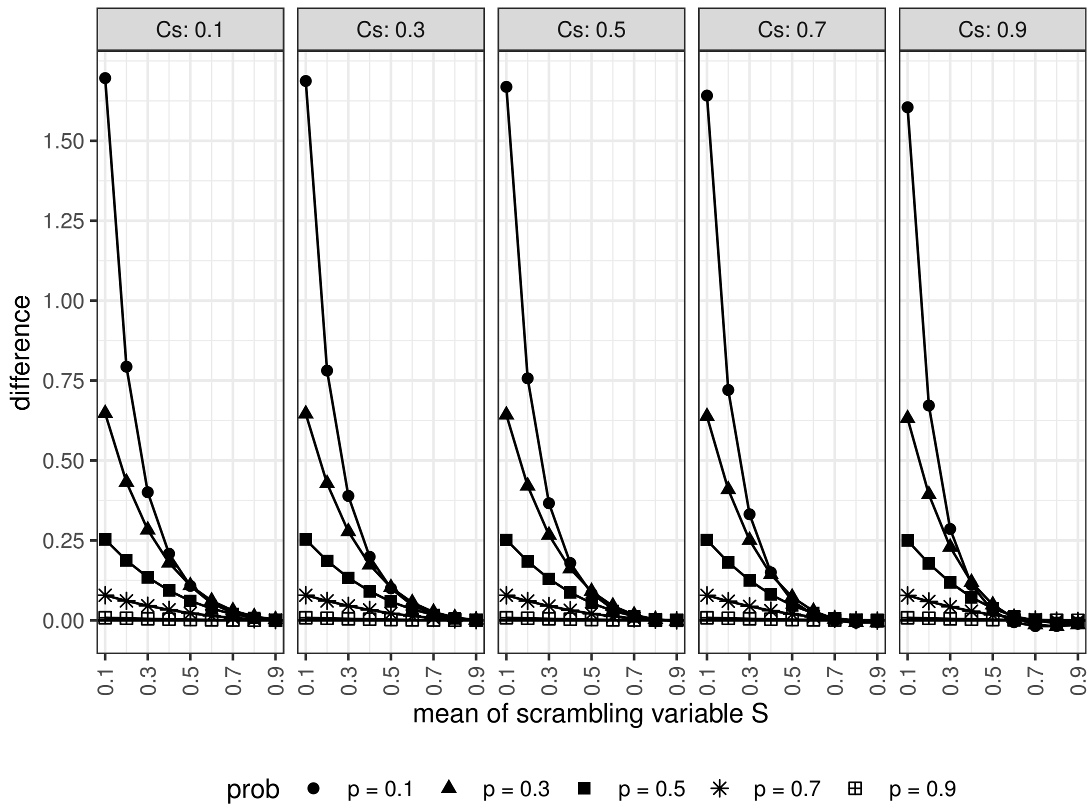

Figure 1: The difference $C^2_{S,p,t} - \gamma$ versus the mean of the scrambling variable S, μ_S, and p, at different values of C_S.

Figure 1 shows the calculated differences $C^2_{S,p,t} - \gamma$. The differences are positive at the start with $C_S = 0.1$ and 0.3. For $C_S = 0.5$, the differences are mostly positive except for values of $\mu_S > 0.8$ where we observe negative values that are very close to 0. On the other hand, for $C_S = 0.7$ and 0.9, we observe negative values for μ_S as little as 0.55. However, these negative values are very small and can be seen, in italics, in Table 2. As μ_S approaches 1, the differences converge to 0, which is expected since the difference at $\mu_S = 1$ is 0.

2.4 Efficiency Comparison with Singh and Gorey (2019) Estimator

As noted, the proposed model reduces to the Singh and Gorey (2019) model at $t = 0$, and to a direct response question at $t = 1$. Therefore, to show that the proposed estimator is more efficient than that of Singh and Gorey (2019) it is sufficient to prove that $V(\hat{\mu}_P)$ is a decreasing function in t for $t \in (0, 1)$. To do so we show that the first derivative with respect to t of $V(\hat{\mu}_P)$ is negative for all $t \in (0, 1)$. Consequently, we have the following theorem:

Theorem 2.3. *The proposed estimator is more efficient than that of Singh and Gorey (2019).*

Proof. The first derivative of $V(\hat{\mu}_P)$, given in Equation 2.5, with respect to t reduces to:

$$\frac{d\gamma}{dt} = \frac{d}{dt}\Big(\frac{(1-t)(1-p)\sigma_S^2}{b^2}\Big) = \frac{-(1-p)\sigma_S^2}{b^2}, \tag{2.16}$$

which is negative for all $t \in (0, 1)$ and $p \in (0, 1)$. □

Since the proposed estimator is more efficient than that of Singh and Gorey (2019), it is, consequently, more efficient than those of Bar-Lev et al. (2004) and Odumade and Singh (2009). Moreover, Bar-Lev et al. (2004) proved that their estimator is more efficient than that of Eichhorn and Hayre (1983) under an achievable condition. Therefore, it can be concluded that the proposed estimator is more efficient than all previously discussed estimators.

3 Proposed Model under Stratified Random Sampling

3.1 Model Layout and Estimator Properties

Suppose that a population, of size N, is divided into k non-overlapping strata, each of size N_h, where $h = 1, 2, ..., k$, and $\sum N_h = N$. A simple random sample with replacement of size n_h, is independently selected from each stratum, where $\sum_{h=1}^{k} n_h = n$ is the total sample size. The respondent's response from the h^{th} stratum is given by:

$$Z_h = \begin{cases} X_h b_h, & \text{with probability } t_h, \\ \text{go to } R_{2_h}, & \text{with probability } 1-t_h, \quad \begin{cases} X_h b_h, & \text{with probability } p_h, \\ X_h Y_h, & \text{with probability } 1-p_h, \end{cases} \end{cases} \tag{3.1}$$

where $p_h, t_h \in (0, 1)$ and

$$Y_h = S_h - p_h(\mu_{S_h} - 1), \tag{3.2}$$

is a scrambling variable independent from the sensitive variable and

$$b_h = (1-p_h)\mu_{S_h} + p_h, \tag{3.3}$$

is its known mean.

Following the line of argument used in the case of simple random sampling, we get the following unbiased estimator for the mean of the sensitive variable in stratum h, μ_{X_h}:

$$\hat{\mu}_{P_h} = \frac{\overline{Z}_h}{b_h}, \tag{3.4}$$

where $\overline{Z}_h$ is the moment estimator of the mean of the responses in the h^{th} stratum, with variance equal to:

$$V(\hat{\mu}_{P_h}) = \frac{\mu^2_{X_h}}{n_h}[C^2_{X_h} + (1 + C^2_{X_h})\gamma_h], \tag{3.5}$$

where for stratum h,

$$\gamma_h = \frac{(1 - p_h)(1 - t_h)\sigma^2_{S_h}}{b^2_h}, \tag{3.6}$$

C_{X_h} is the coefficient of variation of the sensitive variable, $\sigma^2_{S_h}$ is the variance of the scrambling variable, and b_h is as defined in Equation 3.3.

Consider the following estimator for the mean of the sensitive variable for the whole population, μ_X:

$$\hat{\mu}^s_P = \sum_{h=1}^{k} w_h \hat{\mu}_{P_h}, \tag{3.7}$$

where $w_h = N_h/N$, such that $\sum_{h=1}^{k} w_h = 1$.

Theorem 3.1. *$\hat{\mu}^s_P$ is an unbiased estimator for μ_X with variance equal to*

$$V(\hat{\mu}^s_P) = \sum_{h=1}^{k} \frac{w^2_h \mu^2_{X_h}}{n_h}[C^2_{X_h} + (1 + C^2_{X_h})\gamma_h], \tag{3.8}$$

where γ_h is as defined in Equation 3.6.

Proof.

$$E(\hat{\mu}^s_P) = E\left(\sum_{h=1}^{k} w_h \hat{\mu}_{P_h}\right) = \sum_{h=1}^{k} w_h \frac{E(Z_h)}{b_h} = \sum_{h=1}^{k} w_h \frac{b_h \mu_{X_h}}{b_h} = \mu_X,$$

$\therefore$ $\hat{\mu}^s_P$ is unbiased.

Using the assumption that the selections in different strata are made independently and substituting $V(\hat{\mu}_{P_h})$ from Equation 3.5, we get:

$$V(\hat{\mu}^s_P) = V\left(\sum_{h=1}^{k} w_h \hat{\mu}_{P_h}\right) = \sum_{h=1}^{k} w^2_h V(\hat{\mu}_{P_h}) = \sum_{h=1}^{k} \frac{w^2_h \mu^2_{X_h}}{n_h}[C^2_{X_h} + (1 + C^2_{X_h})\gamma_h].$$

□

Under Neyman's allocation, n_h is taken as $n_h = n\frac{w_h s_h}{\sum w_h s_h}$, where s_h is the standard deviation of the responses in the h^{th} stratum. The variance of $\hat{\mu}_P^s$ in this case is given as:

$$V_{Ney}(\hat{\mu}_P^s) = \sum_{h=1}^{k} \frac{w_h^2 V(Z_h)}{n_h b_h} \tag{3.9}$$

$$= \frac{1}{n} \sum_{h=1}^{k} \frac{w_h \sqrt{V(Z_h)}}{b_h} \sum_{h=1}^{k} w_h \sqrt{V(Z_h)}, \tag{3.10}$$

where

$$V(Z_h) = \mu_{X_h}^2 b_h^2 [C_{X_h}^2 + (1 + C_{X_h}^2 \gamma_h)]. \tag{3.11}$$

3.2 Efficiency Comparison with Generalized Stratified Ryu et al. Estimator

The proposed estimator under stratified random sampling reduces to that of Ryu et al. (2005) at $\mu_{S_h} = 1$ for all $h = 1, ..., k$. For a more comprehensive comparison, we extend the stratified Ryu et al. (2005) to the case of $\mu_{S_h} > 0$. We call the estimator the generalized stratified Ryu et al. estimator and denote it by $\hat{\mu}_R^{*s}$. Its variance is as follows:

$$V(\hat{\mu}_R^{*s}) = \sum_{h=1}^{k} \frac{w_h^2 \mu_{X_h}^2}{n_h} [C_{X_h}^2 + (1 + C_{X_h}^2) C_{S_h,p_h,t_h}^2], \tag{3.12}$$

where $w_h = N_h/N$, such that $\sum_{h=1}^{k} w_h = 1$, and

$$C_{S_h,p_h,t_h}^2 = \frac{p_h + (1-p_h)t_h + (1-p_h)(1 - t_h(\mu_{S_h}^2 + \sigma_{S_h}^2))}{c_h^2} - 1, \tag{3.13}$$

$$c_h = p_h + (1-p_h)t_h + (1-p_h)(1-t_h)\mu_{S_h}. \tag{3.14}$$

Comparing the variance of the proposed stratified estimator for each stratum, $\hat{\mu}_{P_h}$, to that of the generalized stratified Ryu et al. estimator gives us the following theorem:

Theorem 3.2. *The proposed stratified estimator is more efficient than that of the generalized stratified Ryu et al. estimator for* $\mu_{S_h} > 1$, $h = 1, ..., k$.

Proof. The difference between the variance of the generalized stratified Ryu et al. estimator, given in Equation 3.12, and the variance of the proposed stratified estimator given in Equation 3.8, is given by:

$$V(\hat{\mu}_R^{*s}) - V(\hat{\mu}_P^s) = \sum_{h=1}^{k} \frac{w_h^2 \mu_{X_h}^2}{n_h} (1 + C_{X_h}^2)[C_{S_h,p_h,t_h}^2 - \gamma_h], \tag{3.15}$$

where $C^2_{S_h,p_h,t_h}$ and γ_h are as defined in Equations 3.13 and 3.6, respectively. It suffices to show that $C^2_{S_h,p_h,t_h} - \gamma_h > 0$ for $\mu_{S_h} > 1$ as follows:

$$\begin{aligned}
C^2_{S_h,p_h,t_h} - \gamma_h =& \frac{p_h + (1-p_h)t + (1-p_h)(1-t_h)(\mu^2_{S_h} + \sigma^2_{S_h})}{c^2_h} - 1 - \frac{(1-p_h)(1-t_h)\sigma^2_{S_h}}{b^2_h} \\
=& \frac{(1-p_h)(1-t_h)\Big[\sigma^2_{S_h} + (1-\mu_{S_h})^2(p_h + (1-p_h)t_h)\Big]}{c^2_h} - \frac{(1-p_h)(1-t_h)\sigma^2_{S_h}}{b^2_h} \\
=& (1-p_h)(1-t_h)\sigma^2_{S_h}\Big[\frac{1}{c^2_h} - \frac{1}{b^2_h}\Big] + \frac{(1-p_h)(1-t_h)\sigma^2_{S_h}(1-\mu_{S_h})^2(p_h + (1-p_h)t_h)}{c^2_h} \\
=& (1-p_h)(1-t_h)\sigma^2_{S_h}\Big[\frac{(1-p_h)t_h(\mu_{S_h} - 1)(b_h + c_h)}{c^2_h b^2_h}\Big] \\
&+ \frac{(1-p_h)(1-t_h)\sigma^2_{S_h}(1-\mu_{S_h})^2(p_h + (1-p_h)t_h)}{c^2_h},
\end{aligned}$$

where c_h and b_h are as defined in Equation 3.14 and Equation 3.3.

Therefore, the difference is guaranteed to be positive for $\mu_{S_h} > 1$, $h = 1, ..., k$. □

3.3 Efficiency Comparison with Singh and Gorey (2019) Estimator

The variance of the Singh and Gorey (2019) estimator under stratified random sampling is given by:

$$V(\hat{\mu}^s_{SG}) = \sum_{h=1}^{k} \frac{w^2_h \mu^2_{X_h}}{n_h}\Big(C^2_{X_h} + (1 + C^2_{X_h})(C^*_{S_h}(p_h) - \psi_h)\Big), \tag{3.16}$$

where $C^*_{S_h}(p_h)$ and ψ_h are as defined in Equation 1.6 and Equation 1.7 but for stratum h.

Conducting a similar analogy to that of the simple random sample case, the variance of the proposed estimator reduces to that of Singh and Gorey (2019) at $t_h = 0$ for all $h = 1, ..., k$, and reduces to that of a direct questioning at $t_h = 1$ for all $h = 1, ..., k$. Therefore, it is only important to show that the variance of the proposed estimator is a decreasing function in t_h for $t_h \in (0, 1)$. Hence, we have the following theorem:

Theorem 3.3. *The proposed estimator is more efficient than that of Singh and Gorey (2019).*

Proof. The first derivative of $V(\hat{\mu}_{P_h})$ with respect to t_h, $h = 1, ..., k$, is always negative as it reduces to the following:

$$\frac{d\gamma_h}{dt_h} = \frac{d}{dt_h}\Big(\frac{(1-t_h)(1-p_h)\sigma^2_{S_h}}{b^2_h}\Big) = \frac{-(1-p_h)\sigma^2_{S_h}}{b^2_h}, \tag{3.17}$$

where b_h is as defined in Equation 3.3, and $p_h \in (0, 1)$. □

4 Privacy and Protection

The main goal of randomized response models is to provide the respondents with the needed privacy necessary for them to cooperate and provide truthful answers. Many measures have been proposed to quantify the privacy of randomized response models. In this paper, the measure $\tau = 1 - \rho_{x,z}^2$, suggested by Diana and Perri (2011), is used to measure the privacy of the models, where $\rho_{x,z}$ is the coefficient of correlation between the true response X and the randomized response Z. The advantage of this measure is that it ranges between 0 and 1.

Diana and Perri (2011) noted that the measure τ reduces to a ratio of variances as follows:

$$\tau = 1 - \rho_{x,z}^2 = 1 - \frac{V(\bar{x})}{V(\hat{\mu})} = 1 - \frac{1}{R.E(\bar{x}, \hat{\mu})}, \tag{4.1}$$

where $\hat{\mu}$ is any randomized response model unbiased estimator of μ_X.

Therefore, as the relative efficiency of the direct response estimator, $\bar{x}$, to the randomized response estimator, $\hat{\mu}$, increases, the privacy protection increases. Hence, it is not necessary to calculate the measure τ to know that the most efficient randomized response estimator has the least protection.

4.1 Measure of Performance

To study the overall performance of the randomized response estimators considering their efficiency and protection degree at the same time, we propose using a new measure of performance, Λ, which is given by:

$$\Lambda = \frac{R.E.(\bar{x}, \hat{\mu})}{\tau}. \tag{4.2}$$

The smaller the value of Λ the better the overall performance of the estimator whether due to small relative efficiency of $\bar{x}$ over it or the greater privacy it provides.

In the following subsection we carried out numerical comparison, in terms of Λ, between the proposed estimator and its competitors. The comparisons were made once using variations of the different models parameters and another time using real data.

4.2 Performance Comparisons

To carry out the numerical comparisons between the proposed estimator and its competitors, some assumptions were imposed. First, we considered the case $p = t$ where $p \in \{0.1, (0.1), 0.9\}$. It is easy to imagine that we can have as much liberty in the second stage of the model as we have in the first stage. Second, the $F(1, 5)$ distribution was used for the scrambling variable S. Consequently, $\mu_S = 5/3$, $\sigma_S^2 = 200/9$, and $C_S = 2\sqrt{2}$. This distribution satisfies Bar-Lev et al. (2004)'s condition that guarantees their estimator's efficiency over that of Eichhorn and Hayre (1983). Third, to compare the proposed estimator to that of Singh and Gorey (2019), the same values of C_X that were suggested in their paper were used, mainly $C_X \in \{0.1, (0.1), 1\}$. Fourth, to include Ryu et al. (2005) in the comparison, we used the generalized form obtained in subsection 2.3. The variances for the estimators of Eichhorn and Hayre (1983), Bar-Lev et al. (2004), and Odumade and Singh (2009) used in the comparison are given respectively as follows:

$$V(\hat{\mu}_E) = \frac{\mu_X^2}{n}\Big[C_X^2 + C_S^2(1 + C_X^2)\Big], \tag{4.3}$$

$$V(\hat{\mu}_B) = \frac{\mu_X^2}{n}\Big[C_X^2 + C_S^*(p)(1 + C_X^2)\Big], \tag{4.4}$$

$$V(\hat{\mu}_{OS}) = \frac{\mu_X}{n}\Big[C_X^2 + (1 + C_X^2)C_S^*(p) - \psi\Big]. \tag{4.5}$$

Figures 2 and 3 showcase the comparison between the different estimators, in terms of Λ, for various combinations of parameters.

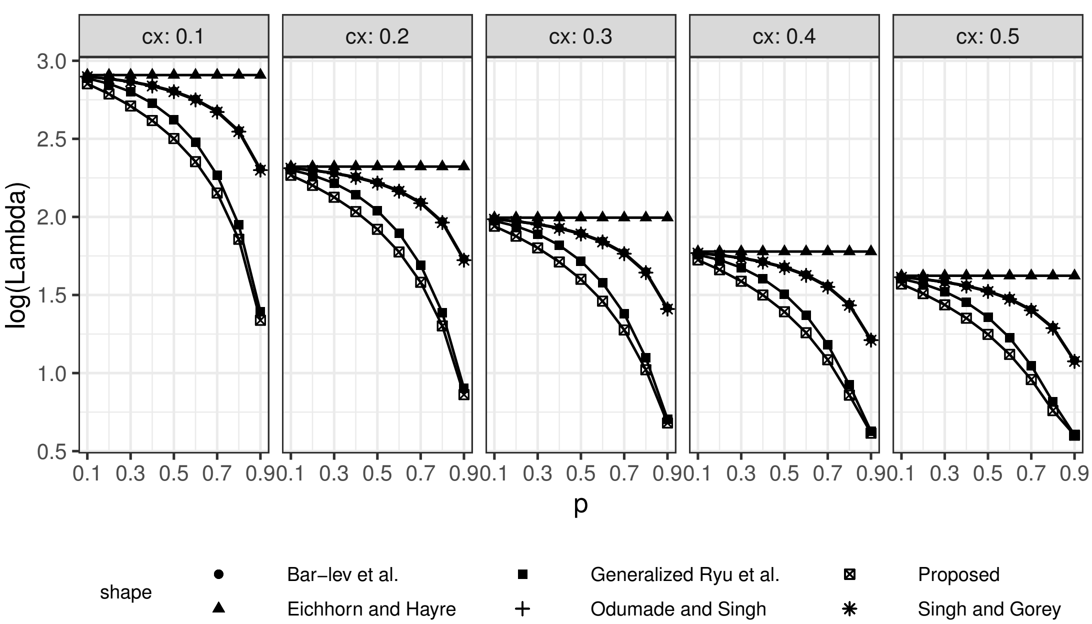

Figure 2: Calculated $\log(\Lambda)$ for the different estimators, at $C_X \in \{0.1, (0.1), 0.5\}$.

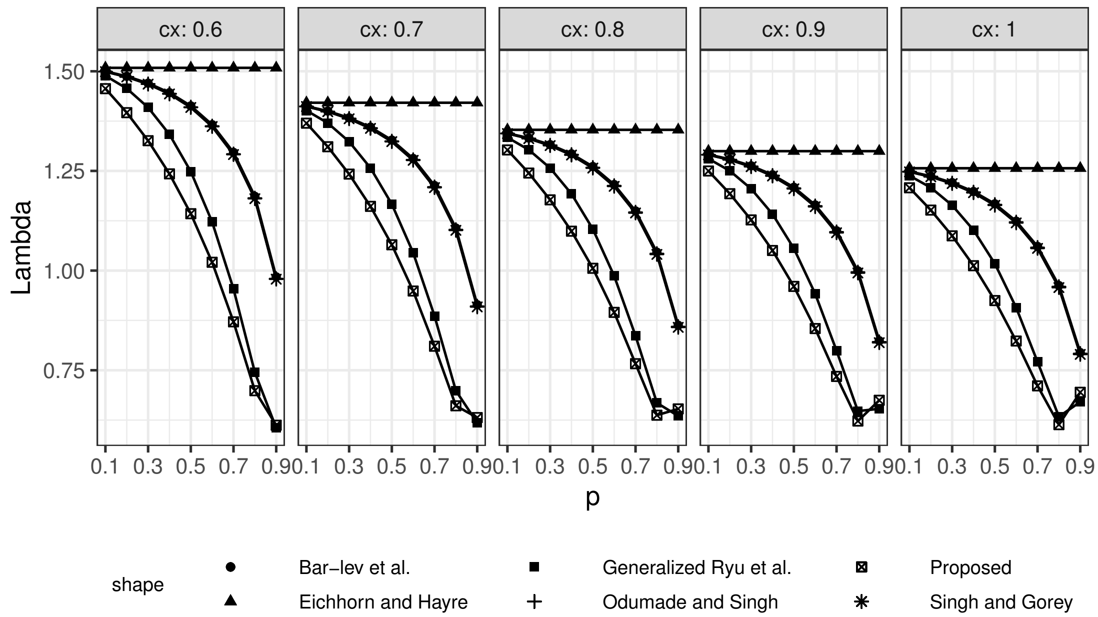

Figure 3: Calculated Λ for the different estimators, for $C_X \in \{0.6, (0.1), 1\}$.

As we can see in Figure 2, the proposed estimator has the lowest value of $\log(\Lambda)$ for all levels of p and $C_X \in \{0.1, (0.1), 0.5\}$. Additionally, all the estimators perform better at higher values of p and C_X. The differences between the estimators seem to decrease as C_X increases. Note that the graph shows $\log(\Lambda)$ for better readability of the data.

Figure 3 shows that the proposed estimator has the lowest Λ for $C_X \in \{0.6, (0.1), 1\}$ and $p \neq 0.9$. For $p = 0.9$, the generalized Ryu et al. estimator has the best performance. However, it is important to note that it will be hard to carry out a two-stage randomization model with $p = t$ at values of p as high as 0.9 since it will not earn the confidence of the respondent.

The values of Λ are provided in Table 3 in the Appendix where the values in bold refer to the smallest value of Λ in each row.

4.3 Real Data Performance Comparison

The same secondary data set used in subsection 2.1, 2017 CPS ASEC, was used to obtain estimated values of Λ for the proposed estimator and its competitors. Ten thousand samples each of size $n = 100$ were obtained from the observations of the study variable WSAL_VAL whose $C_X = 1.35$. We restricted the simulation to reasonable values of $p = t$, mainly $p = t \in \{0.3, (0.1), 0.7\}$. The response, z_i, of the i^{th} person was obtained under each of the competing randomized response models. To do so, we adopted a similar procedure to the one used in subsection 2.1 to obtain z_i for the proposed model, in the real data example, to each of the models under study.

After obtaining the responses, the estimators of the different models were obtained for each of the 10,000 samples. The variance was then calculated for each estimator and for $\bar{x}$, and the relative efficiencies, $RE(\bar{x}, \hat{\mu})$, were obtained. Finally, Λ was calculated as in Equation 4.2.

The results are presented in Table 1. The proposed model obtained the lowest values of Λ, in bold, for all the values of $p = t$ considered. This shows that the proposed estimator retained the best performance even for $C_X = 1.35$, which is greater than the maximum value of C_X used in the previous numerical comparisons.

Table 1: Λ for the different estimators

n	p	Λ_E	Λ_B	Λ_R^*	Λ_{OS}	Λ_{SG}	Λ_P
100	0.3	8.971	8.227	6.847	8.227	8.222	**5.794**
100	0.4	8.918	8.011	6.413	8.011	7.974	**5.227**
100	0.5	8.283	6.584	5.355	6.584	6.570	**4.325**
100	0.6	8.703	6.822	4.470	6.822	6.743	**3.605**
100	0.7	8.886	5.644	3.210	5.644	5.589	**2.707**

5 Conclusion

In this paper, we introduced a new randomized response model for estimating the mean of a sensitive quantitative random variable. An unbiased estimator for the mean of the sensitive variable was obtained and its efficiency was compared to different randomized response estimators. Through our analysis, the proposed estimator proved to be more efficient than Eichhorn and Hayre (1983), Bar-Lev et al. (2004), Odumade and Singh (2009), and Singh and Gorey (2019) estimators. The model was applied on a real secondary data set to showcase how the model can be applied in practice.

Additionally, Ryu et al. (2005) estimator can be considered as a special case of the proposed estimator when $\mu_S = 1$. We obtained a generalized form of Ryu et al. (2005) estimator for $\mu_S > 0$. The proposed estimator was more efficient than that of generalized Ryu et al. for $\mu_S > 1$. For $\mu_S \in (0, 1)$, the efficiency was investigated numerically at different values of C_S, μ_S, and at $p = t$. The results were in favor of the proposed estimator in most cases. Therefore, the proposed estimator is more (or equally as) efficient than all the previously discussed estimators.

The proposed model was extended to the stratified random sampling with replacement and an unbiased estimator of μ_X was obtained. The resulting stratified estimator reduces to that of Ryu et al. (2005) at $\mu_{S_h} = 1$ for all $h = 1, ..., k$. In case of the generalized stratified Ryu et al. model, the proposed stratified estimator is more efficient for

$\mu_{S_h} > 1 \ \forall \ h = 1, .., k$. When compared to the stratified estimator of Singh and Gorey (2019), the proposed estimator was more efficient.

In terms of privacy and efficiency, the performance of the proposed estimator was examined using the proposed measure of performance Λ. The proposed estimator had a better performance than the rival estimators at $p = t$ and $S \sim F(1, 5)$, for all values of p and $C_X \in \{0.1, (0.1), 0.6\}$. For $C_X \in \{0.7, (0.1), 1\}$, the proposed estimator had the best performance for all values of p except 0.9. It is worth mentioning that $p = 0.9$ is not a good choice for the randomization device since it raises the suspicion of the respondents. Also, it is clear from Table 3 that the loss in performance does not exceed 0.27 for all cases. Therefore, the increased efficiency offsets the decrease in privacy protection for almost all values of p and $C_X \leq 1$.

A real secondary data set was used to implement the model and compare its performance to the rest of the considered estimators. The 2017 CPS ASEC data were used with the variable of study being the total wage and salary earnings for individuals (WSAL_VAL). The proposed estimator had the best performance in terms of Λ for values of $p = t \in \{0.3, (0.1), 0.7\}$, and for $C_X = 1.35$ (>1).

Acknowledgments

We would like to thank the referees for taking the time to carefully revise the manuscript and for their valuable input and suggestions that greatly enriched the paper.

References

Bar-Lev, S. K., Bobovitch, E., and Benzion, B. (2004), A note on randomized response models for quantitative data, *Metrika*, **60**(3), 255-260.

Diana, G. and Perri, P. F. (2011), A class of estimators for quantitative sensitive data,*Statistical Papers*, **52**, 633-650.

Eichhorn, B. H., and Lakhbir S. Hayre, L. S. (1983), Scrambled randomized response methods for obtaining sensitive quantitative data, *Journal of Statistical Planning and Inference*, **7**(4), 307 - 316.

Greenberg, B. G, Kuebler Jr, R. R, Abernathy,J. R., and Horvitz, D. G. (1971), Application of the Randomized Response Technique in Obtaining Quantitative Data, *Journal of the American Statistical Association*, **66**(334), 243-250.

Gupta, S., Gupta , B. and Singh, S. (2002), Estimation of sensitivity level of personal interview survey questions,*Journal of Statistical Planning and Inference*, **100**(2), 239-247.

Odumade, O. and Singh, S. (2009), Improved Bar-Lev, Bobovitch, and Boukai Randomized Response Models,*Communications in Statistics - Simulation and Computation*, **38**(3), 473-502.

Ryu, J. B., Kim, J. M., Heo, T. Y., and Park, C. G. (2005), On stratified randomized response sampling, *Model Assisted Statistics and Applications*, **1**(1), 31-36.

Singh, H. P. and Gorey, S. (2019), A modified efficient randomized response model,*Investigacion Operacional*, **40**(3), 374-385.

Warner, S. L. (1965), Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias, *Journal of the American Statistical Association*, **60**(309), 63-69.

Appendix

Table 2: The difference $C^2_{S,p,t} - \gamma$

	$C_S = 0.1$								
	μ_S								
$p = t$	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.10	1.696263	0.793420	0.400779	0.208489	0.107704	0.053147	0.023649	0.008439	0.001668
0.20	1.037632	0.618288	0.369426	0.217498	0.123583	0.065777	0.031136	0.011702	0.002422
0.30	0.647579	0.432170	0.283008	0.179735	0.108850	0.061181	0.030364	0.011900	0.002563
0.40	0.408298	0.290624	0.201395	0.134475	0.085165	0.049835	0.025655	0.010402	0.002317
0.50	0.252821	0.187378	0.134784	0.093163	0.060930	0.036742	0.019459	0.008108	0.001859
0.60	0.148555	0.113068	0.083421	0.059076	0.039545	0.024386	0.013199	0.005620	0.001320
0.70	0.078542	0.060845	0.045671	0.032893	0.022385	0.014031	0.007719	0.003343	0.000801
0.80	0.033469	0.026222	0.019905	0.014497	0.009977	0.006324	0.003519	0.001543	0.000376
0.90	0.008165	0.006438	0.004918	0.003605	0.002497	0.001594	0.000894	0.000395	0.000098
	$C_S = 0.3$								
0.10	1.687136	0.781278	0.389284	0.198734	0.099910	0.047242	0.019474	0.005817	0.000432
0.20	1.033949	0.611085	0.360747	0.208791	0.115709	0.059214	0.026128	0.008356	0.000761
0.30	0.645970	0.428313	0.277617	0.173674	0.102847	0.055790	0.025985	0.008815	0.000959
0.40	0.407567	0.288636	0.198322	0.130722	0.081179	0.046035	0.022403	0.008003	0.001019
0.50	0.252492	0.186406	0.133169	0.091061	0.058571	0.034381	0.017348	0.006488	0.000950
0.60	0.148417	0.112634	0.082659	0.058036	0.038326	0.023115	0.012020	0.004684	0.000778
0.70	0.078492	0.060680	0.045371	0.032467	0.021868	0.013473	0.007185	0.002906	0.000542
0.80	0.033456	0.026177	0.019821	0.014373	0.009822	0.006153	0.003351	0.001402	0.000290
0.90	0.008164	0.006433	0.004908	0.003590	0.002478	0.001572	0.000871	0.000376	0.000086
	$C_S = 0.5$								
0.10	1.668882	0.756995	0.366295	0.179225	0.084322	0.035433	0.011125	0.000575	*-0.002038*
0.20	1.026584	0.596680	0.343389	0.191377	0.099962	0.046088	0.016114	0.001664	*-0.002560*
0.30	0.642752	0.420598	0.266836	0.161552	0.090841	0.045008	0.017229	0.002644	*-0.002248*
0.40	0.406105	0.284660	0.192177	0.123215	0.073207	0.038434	0.015901	0.003206	*-0.001578*
0.50	0.251836	0.184462	0.129937	0.086858	0.053855	0.029659	0.013127	0.003249	*-0.000867*

0.60	0.148141	0.111766	0.081137	0.055957	0.035887	0.020574	0.009663	0.002812	*-0.000306*
0.70	0.078393	0.060352	0.044771	0.031615	0.020832	0.012358	0.006118	0.002033	0.000022
0.80	0.033429	0.026088	0.019651	0.014126	0.009512	0.005810	0.003015	0.001120	0.000118
0.90	0.008161	0.006422	0.004888	0.003559	0.002439	0.001527	0.000827	0.000338	0.000062
				$C_S = 0.7$					
0.10	1.641502	0.720569	0.331811	0.149962	0.060939	0.017720	*-0.001398*	*-0.007289*	*-0.005744*
0.20	1.015536	0.575072	0.317353	0.165256	0.076340	0.026399	0.001092	*-0.008373*	*-0.007543*
0.30	0.637926	0.409025	0.250664	0.143368	0.072832	0.028835	0.004094	*-0.006612*	*-0.007058*
0.40	0.403913	0.278696	0.182960	0.111956	0.061249	0.027032	0.006147	*-0.003990*	*-0.005472*
0.50	0.250852	0.181545	0.125090	0.080554	0.046780	0.022575	0.006796	*-0.001610*	*-0.003594*
0.60	0.147728	0.110464	0.078853	0.052838	0.032229	0.016761	0.006127	0.000004	*-0.001932*
0.70	0.078243	0.059859	0.043870	0.030337	0.019279	0.010684	0.004516	0.000723	*-0.000758*
0.80	0.033390	0.025953	0.019398	0.013755	0.009048	0.005296	0.002510	0.000697	*-0.000140*
0.90	0.008156	0.006406	0.004857	0.003513	0.002380	0.001461	0.000760	0.000281	0.000027
				$C_S = 0.9$					
0.10	1.604995	0.672002	0.285833	0.110943	0.029762	*-0.005898*	*-0.018096*	*-0.017773*	*-0.010686*
0.20	1.000805	0.546262	0.282638	0.130428	0.044844	0.000147	*-0.018937*	*-0.021757*	*-0.014186*
0.30	0.631490	0.393595	0.229101	0.119124	0.048820	0.007270	*-0.013420*	*-0.018954*	*-0.013472*
0.40	0.400989	0.270745	0.170670	0.096943	0.045305	0.011831	*-0.006859*	*-0.013584*	*-0.010665*
0.50	0.249539	0.177656	0.118627	0.072147	0.037347	0.013131	*-0.001645*	*-0.008089*	*-0.007228*
0.60	0.147177	0.108728	0.075808	0.048680	0.027352	0.011677	0.001412	*-0.003740*	*-0.004101*
0.70	0.078044	0.059203	0.042669	0.028633	0.017208	0.008452	0.002381	*-0.001023*	*-0.001798*
0.80	0.033337	0.025774	0.019059	0.013260	0.008430	0.004611	0.001837	0.000133	*-0.000484*
0.90	0.008150	0.006385	0.004816	0.003452	0.002302	0.001372	0.000671	0.000205	*-0.000021*

Table 3: Λ at $S \sim F(1,5)$

C_X	p	$\hat{\mu}_E$	$\hat{\mu}_B$	$\hat{\mu}_R^*$	$\hat{\mu}_{OS}$	$\hat{\mu}_{SG}$	$\hat{\mu}_P$
0.10	0.10	810.00	792.64	771.49	791.08	791.06	**712.16**
	0.20	810.00	768.76	712.82	765.74	765.71	**612.97**
	0.30	810.00	736.76	633.23	732.42	732.37	**513.26**
	0.40	810.00	694.57	534.22	689.13	689.08	**414.25**
	0.50	810.00	639.56	420.43	633.31	633.25	**317.63**
	0.60	810.00	568.27	300.15	561.63	561.56	**225.83**
	0.70	810.00	476.14	185.06	469.66	469.59	**142.28**
	0.80	810.00	357.08	88.82	351.54	351.48	**71.91**
	0.90	810.00	202.82	24.63	199.31	199.27	**21.78**
0.20	0.10	210.00	205.54	200.09	205.15	205.13	**184.82**
	0.20	210.00	199.39	184.99	198.63	198.60	**159.28**
	0.30	210.00	191.15	164.50	190.07	190.02	**133.62**
	0.40	210.00	180.29	139.01	178.93	178.88	**108.13**
	0.50	210.00	166.13	109.72	164.57	164.51	**83.26**
	0.60	210.00	147.78	78.76	146.12	146.05	**59.64**
	0.70	210.00	124.06	49.14	122.44	122.38	**38.14**
	0.80	210.00	93.42	24.39	92.03	91.98	**20.05**
	0.90	210.00	53.71	7.99	52.84	52.80	**7.28**
0.30	0.10	98.90	96.82	94.28	96.64	96.63	**87.17**
	0.20	98.90	93.95	87.25	93.62	93.59	**75.28**
	0.30	98.90	90.12	77.71	89.64	89.59	**63.32**
	0.40	98.90	85.06	65.84	84.46	84.40	**51.45**
	0.50	98.90	78.46	52.19	77.77	77.71	**39.87**
	0.60	98.90	69.92	37.78	69.18	69.11	**28.88**
	0.70	98.90	58.87	24.00	58.15	58.09	**18.88**
	0.80	98.90	44.60	12.51	43.99	43.93	**10.50**
	0.90	98.90	26.12	5.08	25.73	25.70	**4.79**
0.40	0.10	60.02	58.77	57.25	58.67	58.66	**53.00**
	0.20	60.02	57.06	53.04	56.87	56.84	**45.88**
	0.30	60.02	54.76	47.33	54.49	54.45	**38.73**
	0.40	60.02	51.73	40.23	51.39	51.34	**31.63**
	0.50	60.02	47.79	32.07	47.40	47.33	**24.70**
	0.60	60.02	42.67	23.45	42.26	42.19	**18.13**
	0.70	60.02	36.06	15.22	35.66	35.59	**12.17**
	0.80	60.02	27.53	8.39	27.18	27.13	**7.22**
	0.90	60.02	16.48	4.24	16.27	16.23	**4.12**
0.50	0.10	42.03	41.17	40.12	41.10	41.09	**37.18**
	0.20	42.03	39.98	37.22	39.86	39.83	**32.28**
	0.30	42.03	38.40	33.28	38.23	38.18	**27.35**
	0.40	42.03	36.31	28.39	36.10	36.04	**22.46**

continued ...

... continued

C_X	p	$\hat{\mu}_E$	$\hat{\mu}_B$	$\hat{\mu}_R$	$\hat{\mu}_{OS}$	$\hat{\mu}_{SG}$	$\hat{\mu}_P$
	0.50	42.03	33.59	22.76	33.34	33.28	**17.69**
	0.60	42.03	30.07	16.83	29.80	29.74	**13.17**
	0.70	42.03	25.51	11.17	25.26	25.19	**9.09**
	0.80	42.03	19.63	6.53	19.41	19.36	**5.75**
	0.90	42.03	12.04	4.01	11.90	11.87	**4.00**
0.60	0.10	32.26	31.61	30.82	31.56	31.55	**28.60**
	0.20	32.26	30.71	28.62	30.63	30.60	**24.90**
	0.30	32.26	29.52	25.65	29.40	29.36	**21.18**
	0.40	32.26	27.94	21.96	27.79	27.74	**17.48**
	0.50	32.26	25.89	17.71	25.72	25.65	**13.89**
	0.60	32.26	23.23	13.24	23.04	22.98	**10.49**
	0.70	32.26	19.79	8.99	19.61	19.55	**7.44**
	0.80	32.26	15.36	5.55	15.20	15.15	**5.00**
	0.90	32.26	9.64	**4.03**	9.55	9.51	4.09
0.70	0.10	26.37	25.85	25.21	25.81	25.80	**23.43**
	0.20	26.37	25.13	23.45	25.07	25.04	**20.45**
	0.30	26.37	24.17	21.06	24.08	24.03	**17.46**
	0.40	26.37	22.90	18.09	22.79	22.73	**14.49**
	0.50	26.37	21.25	14.68	21.12	21.06	**11.61**
	0.60	26.37	19.11	11.09	18.97	18.91	**8.89**
	0.70	26.37	16.34	7.69	16.21	16.15	**6.46**
	0.80	26.37	12.78	5.00	12.67	12.62	**4.58**
	0.90	26.37	8.21	**4.15**	8.14	8.11	4.28
0.80	0.10	22.55	22.11	21.57	22.09	22.07	**20.07**
	0.20	22.55	21.51	20.09	21.46	21.43	**17.57**
	0.30	22.55	20.70	18.08	20.63	20.58	**15.05**
	0.40	22.55	19.63	15.58	19.54	19.49	**12.55**
	0.50	22.55	18.24	12.71	18.14	18.08	**10.13**
	0.60	22.55	16.44	9.70	16.33	16.27	**7.85**
	0.70	22.55	14.11	6.86	14.01	13.95	**5.84**
	0.80	22.55	11.12	4.66	11.03	10.98	**4.34**
	0.90	22.55	7.29	**4.32**	7.24	7.20	4.50
0.90	0.10	19.93	19.55	19.08	19.53	19.51	**17.78**
	0.20	19.93	19.02	17.79	18.99	18.96	**15.59**
	0.30	19.93	18.32	16.04	18.26	18.22	**13.40**
	0.40	19.93	17.39	13.86	17.32	17.27	**11.23**
	0.50	19.93	16.18	11.37	16.10	16.04	**9.13**
	0.60	19.93	14.61	8.75	14.53	14.46	**7.15**
	0.70	19.93	12.59	6.30	12.51	12.44	**5.43**
	0.80	19.93	9.98	4.44	9.92	9.86	**4.19**

continued ...

. . . continued

C_X	p	$\hat{\mu}_E$	$\hat{\mu}_B$	$\hat{\mu}_R$	$\hat{\mu}_{OS}$	$\hat{\mu}_{SG}$	$\hat{\mu}_P$
	0.90	19.93	6.67	**4.50**	6.63	6.59	4.73
1.00	0.10	18.06	17.72	17.30	17.70	17.69	**16.13**
	0.20	18.06	17.25	16.15	17.22	17.19	**14.18**
	0.30	18.06	16.62	14.58	16.58	16.53	**12.22**
	0.40	18.06	15.79	12.63	15.73	15.68	**10.29**
	0.50	18.06	14.70	10.41	14.64	14.58	**8.41**
	0.60	18.06	13.30	8.07	13.24	13.17	**6.66**
	0.70	18.06	11.50	5.90	11.43	11.37	**5.14**
	0.80	18.06	9.17	4.30	9.12	9.06	**4.11**
	0.90	18.06	6.23	**4.68**	6.20	6.16	4.95

Note: The values in **bold** represent the smallest value for Λ for each row. The smaller the value of Λ, the better the quality of the estimator.