Natural and Applied Sciences International Journal (NASIJ)
eISSN: 2788-4619 (online)
https://www.ideapublishers.org/index.php/nasij
Vol. 4, No. 1 (January-June), 2023, pp. 17-36
https://doi.org/10.47264/idea.nasij/4.1.2 Research Article

Estimation of general parameters for sensitive study variables using auxiliary information for finite population

Amjad Mahmood*[1, 2] | Nadia Mushtaq[3] | Muhammad Hanif[1] | Farhad Hussain[4]

1. Department of Statistics, National College of Business Administration and Economics, Lahore, Pakistan.
2. Hailey College of Commerce, University of the Punjab, Lahore, Pakistan.
3. Department of Statistics, Forman Christian College (A Chartered University), Lahore, Pakistan.
4. Department of Management Science and Engineering, Hebei University, Baoding, Hebei, China.

* Corresponding Author Email: amjadmahmood502@gmail.com

Received: 16-Mar-2023 | Revised: 7-Jun-2023 | Accepted: 10-Jun-2023 | Published: 28-Jun-2023

Abstract

The judgment of parameters about the populace is significant for drawing a sample from the population under study in the survey method. Innumerable statisticians introduced numerous estimators to make predictions about the parameters in a population with the application of auxiliary information for sensitive variables. In the current investigation, the researchers tried to depict the general parameter estimate for sensitive variables using randomized response models. The survey was the method in this paper, and a simple random without replacement (SRSWOR) was utilized to gather the sample. Overall, it presented the general ratio and exponential ratio of estimations for the sensitive variable using non-sensitive AV founded on an RRT. The biasness and MSE expressions above second category calculations appeared as outcomes. Many empirical works are replicated to prove the performance of projected estimators for the sensitive variables for the population under study. This proven model will benefit other researchers and statisticians working in the statistics field or data collection, for instance, population census, to take forward it and develop more advanced statistical general parameters, and also for advanced investigations.

Keywords: parameter estimators, sensitive variables, inconsistent, randomized response model, simulation studies, auxiliary information.

How to Cite: Mahmood, A., Mushtaq, N., Hanif, M., & Hussain, F. (2023). Estimation of general parameters for sensitive study variables using auxiliary information for finite population. *Natural and Applied Sciences International Journal (NASIJ), 4*(1), 17-36. https://doi.org/10.47264/idea.nasij/4.1.2

1. Introduction

It is a common understanding that it is not an easy task to choose a whole sample, construct data-gathering tools and receive accurate answers to the asked statements from respondents during the survey stage. The error occurs in the survey due to these issues. Errors in the survey may result from sampling technique use or sampling error (SE). SE happened when there is a disparity in the sample representation and such error can be lower if the researcher can increase the size of the sample (Saleem, 2017). Non-sampling errors were found due to multiple reasons, like during the process phase, non-responsiveness, measurement, and coverage. The coverage error appears when some population elements are by mistake eliminated or incorporated or declassified in the sampling frame. Such kind of errors will bring biasness in the process of estimation. This error will be removed easily when the researcher might use various information-gathering types.

The processing error (PE) found when data coding or segregation mechanism takes place, it may also cause either bias or higher the occurrence of the estimates. PE may be resolved by the use of the technique of quality control. The measurement error (ME) may be generated by the selected participants in the sample or because participants didn't understand any statement in the given tool, scale, or questionnaire. ME may be defined as the variance between the received replies to a meticulous question and the real assessment. The ME may only be lessening if the designed scale is structured in a way that participants could understand it, and by the recruitment of trained interviewers.

During survey conduction, researchers always hope to collect reliable data for the estimation of socio-demographic characteristics. But to gather data in a survey about the sensitive variable for instance income from all sources, consumption of alcohol, evasion of income tax, classless driving patterns, discrimination, abusive behaviour, domestic violence, and child abuse, is not an easy job. Accurate replies to sensitive queries are hard to get from respondents during private discussions with direct queries due to the violation of the privacy of individuals (Onyango *et al.*, 2022). The researchers like Gonzalez-Ocanto s*et al.* (2012) and Lyall *et al.* (2013) discovered that to avoid such errors; the technique of indirect questioning is the right method to resolve this. The indirect questioning method can be vital to eliminate biasness as compared to direct questions (Rosenfeld *et al.*, 2016; Van der Heijden *et al.*, 2000).

However, On the other hand, Wolter and Preisendorfer (2013) responded with more negative outcomes. Even though the indirect questioning technique is promising and popular but it has its demerits like the found estimators are not very efficient as compared with those that investigators got under direct reasoning. Hence, the increase in variance often underbalanced the reduction in biasness. Though much classical research discovered an innovative multiple regression method (Blair & Imai, 2012; Bullock *et al.*, 2011). On the other hand, few old research studies argued to unite manifold techniques of direct/indirect inquiring, these novel ways repeatedly not very successful to primarily surmount the difficulty of competence inbuilt error in indirect inquiring (Aronow *et al.*, 2015; Blair *et al.*, 2014). Resultantly, it becomes mandatory for a fairly great number of participants to attain exact estimates during indirect inquiring practices.

The argument is about the numerical scrutiny of not direct reasoning practices with the utilization of auxiliary information or attributes for the occurrence of susceptible elements in the populace under investigation. The knowledge of such sort can be gathered from the data involving various sources like population polls, organizational reports, and specialist assessments. Even some previous research studies claimed susceptible and personal topics such as attendance, elector vote casting (Karp & Brockington, 2005; Martin *et al.*, 2000), and infection occurrence, and get in touch with the unlawful activities dealing personals (Wolter & Preisendorfer, 2013).

Many participants are all time very reluctant to give factual facts to a touchy matter due to feelings of humiliation or loss of status. Resultantly that question was refused by the respondent to give a response or give a deliberately wrong answer. The Warner's randomized response technique (RRT) was introduced in (1965) and it is considered to be finest and mostly utilized scientific approach that was based on effectual probabilistic consideration to lower non-responsive rates by keeping respondents anonymous.

To counter RRT Warner (1965) and Eichhorn and Hayre (1983) developed an alternative technique called the method of scrambled randomized response (SRR). This technique engrosses the quantitative response of the participant to a question of sensitive nature that is to be multiplied by a random figure from an acknowledged division as produced by the participant himself through some given mechanism. In this way, a fair response was received by the researcher without knowing about any random number used for mixing up the real reply. This technique is a particular case being presented by Warner in 1971, and by Pollock and Bek in 1976.

Different classifications are given for such models as a forced model of SRR (FSRR), a model of partial SRR (PSRR), or a model of optional SRR (OSRR). According to Hussain et *al.* (2007), in the full scrambled randomized response FSRR model, the recruiter is enforced to respond to the random question by the randomized mechanism. As per the words of Ryu *et al.* (2005) in his research study, the method of PSRR occupies an identified participant's number to testify the jumbled answer employing one or more randomization plans. In the OSRR method, the investigator spoon-feeds the participants to testify either accurately the non-sensitive variable, or the response or jumbled reply to the sensitive variable response to the asked query.

The estimators can be prepared with more preciseness by taking into account the auxiliary information for partly or wholly well-known populations. According to Bolfarine and Zacks (1992) in a model-based approach method, techniques for clearing the estimate with auxiliary information were discussed in the usage of model assist or calibration methods (Samdal & Wright, 1984; Deville & Samdal, 1992).

The auxiliary information played an imperative part in the assumption of sample studies. It is used to enhance the accuracy of an estimator. There are many enlisted techniques to use auxiliary information to achieve an efficient performance of estimators. These techniques are product method, regression type, ratio and many others. Auxiliary information can be present in numerous forms like attributes or variable forms. It can be obtained by utilizing census statistics, surveys on the bases of population, results of experimental trialling, and expert outlooks as well as hypotheses on population issues. Kadilar and Cingi (2004), and Gupta and Shabbir (2008) in their studies projected many estimators for the unknown population value estimation.

In the current investigation, the researchers tried to depict the general parameter estimate for sensitive variables using a randomized response technique with the use of auxiliary information for selecting participants in the field. The development is continued in the form of a diverse variety of estimates for dissimilar situations under numerous sampling plans. The finite population means estimator was investigated and proposed. The survey was the method in this paper and a SRSWOR was used to draw the group of people. The proposed mean estimator would be more effective in generating more friendly-oriented estimators in later studies. The present study used an additive model for randomized response with solitary scrambling variables.

2. Literature review

Many empirical studies like Olkin (1958), Raj (1965), Rao and Mudholkar (1967), Shukla (1966), Solanki and Srivastava (1965), Singh (1967) and Srivastava (1967 a;b) have projected ratio-type estimations which consume information from numerous auxiliary variables. Such estimates implicated the utilization of unfamiliar masses, that have had to be predicted and for assumed facts about the means of population for the employed auxiliary information. The opinion of population mean for a study variable was utilized by many investigations (Chand, 1975; Mukerjee *et al.*, 1987; Srivastava *et al.*, 1990). A below fractional information of supplementary variable B was used by one of the studies, which reported three dissimilar states of the auxiliary information in single, 2 typed sampling known as, the case of complete data, case of partial data, and case of non-information (Samiuddin & Hanif, 2007).

Adichwal *et al.* (2022) introduced a novel estimation to predict universal parameter t (a,b) with the use of auxiliary information with SRSWOR. This estimator is useful to a conventional estimator to identify the constants of people, CV, populace mean, SD and mean square of the populace. In another study, Onyango *et al.* (2022) discussed the issue of estimation of a mean and non-responsive statement under a three-tier RRT form. Auxiliary information trait is utilized to predict a universal group of ratio type of exponential estimates.

Cochran (1940) started the ratio estimator by the consumption of auxiliary information and studied the association between the investigative and the auxiliary variable (AV). In stratified random sampling, Bowley (1926) utilized it. Watson established the concept of auxiliary data to raise the estimate method but it was Neyman (1938), who enlarged the consumption of auxiliary information.

Srivastava and Jhajj (1980; 1981; 1986) also anticipated a group of estimations with value of correlation coefficient and lower mean square error (MSE) that was consumed in the judgment of populace mean. One classical research proposed an estimator that can use auxiliary information as an attribute (Shabbir & Gupta, 2007). Ratio estimation for the mean of the susceptible variable (Y) that utilized data from a non-sensitive AV (X) was introduced by (Sousa *et al.*, 2010).

To guess the population factors, information on two auxiliary variables was used by (Tracy *et al.*, 1996), and many estimators were proposed by classical researchers (Singh & Singh, 2001; Khoshnevisan *et al.*, 2007; Singh *et al.*, 2008; Singh & Solanki, 2011; Sharma *et al.*, 2017; Adichwal *et al.*, 2016, 2017, 2019; Mishra & Singh, 2017). In another similar work by Hansen and Hurwitz (1943), for probability proportion with sampling, auxiliary information was used.

A difference estimate was found by Hansen *et al.* (1953) in the existence of auxiliary information. In a simple random technique, a fresh unbiased ratio estimator with a single auxiliary variable was introduced by Hartley and Ross (1954). The innovative way in the given estimator removes biasness in the ratio estimate.

Robson (1957) introduced the idea of product estimate and projected the estimator for such auxiliary information which portrays a pessimistic association with the key variable of concern. In one of the studies, the multi-auxiliary variable was first time introduced to predict the parameter of finite population and extend the ratio estimate (Olkin, 1958). Many of the previous studies used auxiliary information to plan ratio, product, and ratio-cum-product estimators respectively: for unlike conditions of connection under systematic sampling design (Swain, 1964; Shukla, 1971; Singh, 1967). Another study used linear regression and ratio estimates to introduce regression cum ratio estimator to predict finite population mean; with the use of 2 assisting variables for the success of more precise prediction (Mohanty, 1967).

Singh (1967) opined that during the occurrence of two AVs, a regression product estimator originated. One of the studies used four diverse ways to employ the auxiliary information in survey sampling techniques (Tripathi, 1970; 1973; 1976). Das and Tripathy (1978) employed auxiliary information to estimate the variance in a finite population. In another study, Srivenkataraman (1980) made use of the twin conversion on the AV to obtain the well-organized outcomes of the estimates. In single auxiliary information, a ratio estimator was first given by Isaki (1983), and an efficient variance class was suggested by Srivastava and Jhajj (1980). One of the research by Rueda and Cebrian (1996) by using auxiliary information projected an impartial variance estimate to predict the variance in finite population; also confirmed for ratio estimate, the lesser MSE (Isaki, 1983). By replacing the mean of the support variable with a single phase sampling estimator for both linear association involving the auxiliary variable and chief variable of concern was going by or not passing through the source, Kiregyra (1984) anticipated ratio-in-regression estimator.

Other empirical works like Tripathi *et al.* (1988), bestowed the common judgment theory for bivariate population, while on the other hand, a novel family of estimators by employing two auxiliary information was given by Srivastava *et al.* (1990). Roy (2003) anticipated regression-type estimates with a populace mean of one AV. Similarly, by using 2 AV, another study got an accurate estimation of the planned estimator (Kadilar & Cingi, 2004). The projected estimator balanced with the presented estimators and the optimal conditions were also obtained. To obtain better outcomes of the estimate, Kadilar and Cingi (2006) utilized the identified parameters besides AV complete data. One of the previous investigations offered a group of sequence estimates for the variance judgment for the restricted mean of people with the use of two AVs and balanced it with the on-hand estimators (Jhajj*et al.*, 2007). Along similar lines, Chand (1975) and Samiud din and Hanif (2007) have given the group of linked ratio kind estimates for the population means by employing two AVs in sole and 2^{nd} typed sampling by considering the linear relationship among the investigated variable and the AV.

Bahl and Tuteja (1991) discussed the three situational events such as whole, partial, and non-information. They have given the exponential ratio and exponential result estimators as extraordinary types of the planned group of estimates. A class of unbiased estimates with the application of exponential ratio and exponential product estimate was given by Singh *et al.* (2008) and also in numerical research, even projected the estimators established to be better than the present exponential estimators. A novel chain of exponential estimates for use in the sole auxiliary variable was introduced by Singh *et al.* (2009).

Shabbir and Gupta (2010) by employing a single AV established a general estimator for variance of fixed populace for sample drawn by stratified random. In one of the previous works, a class of strong ratio kind estimates with the use of auxiliary variables, SRSWOR, and tough highest probability estimators was given (Oral & Kadilar, 2011).

Another empirical work has given exponential product kind estimates with the use of linear transformation on auxiliary variables (Grover *et al.*, 2012). Bahl and Tuteja (1991) by employing the concept of an exponential ratio and exponential product estimators also used auxiliary information. Further Upadhyaya *et al.* (2011) anticipated a generalized exponential estimator. The optimal state was achieved which helped the general estimators to perform well than the model given by classical research (Singh*et al.*, 2011), they found two auxiliary variables to the considered ratio exponential and estimator of exponential product for the variance of the restricted populace and unmitigated it in twofold sampling. For the prediction of unknown variance of people, one of the studies by Subramani and Kumarapandiya (2012 a;b) employed the population parameters related to the AV to propose the groups of customized ratio type estimate. The best conditions were also scored under which the planned types were more competent than the normally employed kind of ratio-type estimators.

A lot of work has been done on the study variables but on the particular aspect of sensitive variables with the application of auxiliary variables, limited work is carried out. To fill in the research gap this study has been conducted with the main objective of estimating the general parameters for a sensitive variable under study with the application of auxiliary information for a finite population.

3. Methodology

In this research, through the method of survey and SRSWOR, researchers tried to obtain the terminology of estimated MSE for the designed estimations by using the well-known Taylor and exponential sequence. Experimental and theoretical studies were conducted using R-software to get the results and tried to evaluate the efficiency of the designed estimates over the previously occurred estimations.

3.1. Terminology

To deliberate a limited population with elements from $U = U_1, U_1, \ldots U_N$ that sample extent n is brought by the consumption of sampling technique of$(SRSWOR)$. Let it stand as the variable for investigation, an insightful variable that couldn't be detected as straight as a result of the biases of the participant. Let X be as a non-susceptible AV connected with Y. Let S be a scrambling variable i.e., independent of Y & X.

The participant desired to give a scrambled reply but was asked to present a factual response$X$. Let $(\bar{y}, \bar{x})$ be the sample means matching to$(\bar{Y}, \bar{X})$, the mean of the populace of Y and X, correspondingly.

To attain the *MSE* terms for the anticipated estimation, it can be presented as:

$$\epsilon_o = \frac{\bar{z}}{\bar{\bar{Z}}} - 1 \qquad \epsilon_1 = \frac{S_z^{\ 2}}{S_Z^{\ 2}} - 1 \qquad \epsilon_2 = \frac{\bar{x}}{\bar{\bar{X}}} - 1 \qquad \epsilon_3 = \frac{S_x^{\ 2}}{S_X^{\ 2}} - 1$$

$$E(\epsilon_o) = E(\epsilon_1) = E(\epsilon_2) = E(\epsilon_3) = 0$$

$$E(\epsilon_o^{\ 2}) = n^{-1} C_Z^{\ 2} \qquad E(\epsilon_1^{\ 2}) = n^{-1}(\delta_{40} - 1) \qquad E(\epsilon_2^{\ 2}) = n^{-1} C_X^{\ 2}$$

$$E(\epsilon_3^{\ 2}) = n^{-1}(\delta_{04} - 1)$$

$$E(\epsilon_o \epsilon_1) = n^{-1}\delta_{30} C_Z \qquad E(\epsilon_o \epsilon_2) = n^{-1}\rho_{ZX} C_X C_Z \qquad E(\epsilon_o \epsilon_3) = n^{-1}\delta_{12} C_Z$$

$$E(\epsilon_1 \epsilon_2) = n^{-1}\delta_{21} C_X \qquad E(\epsilon_1 \epsilon_3) = n^{-1}(\delta_{22} - 1) \qquad E(\epsilon_2 \epsilon_3) = n^{-1}\delta_{03} C_X$$

Whereas,

$$\mu_{rs} = \frac{1}{N}\sum^{N}\left(Z - \bar{Z}\right)^r \left(X - \bar{X}\right)^s ; \delta_{rs} = \frac{\mu_{rs}}{\mu_Z^{\ r/2} \mu_Z^{\ s/2}} \text{ and } (r, s) \text{ is non-negative integers}$$

$$\mu_{20} = S_Z^{\ 2}, \quad \mu_{02} = S_X^{\ 2}, \quad \mu_{11} = S_{XZ}, \quad C_Z^{\ 2} = \frac{S_Z^{\ 2}}{\bar{Z}^2} = \frac{\mu_{20}}{\bar{Z}^2}$$

$$C_X^{\ 2} = \frac{S_X^{\ 2}}{\bar{X}^2} = \frac{\mu_{11}}{\bar{X}^2} \quad \text{and} \quad \rho_{XZ} = \frac{S_{XZ}}{S_X S_Z} = \frac{\mu_{11}}{\sqrt{\mu_{02}}\sqrt{\mu_{20}}} = \rho_{zx} = \frac{\rho_{yx}}{\sqrt{1 + \frac{s_s^2}{s_y^2}}}$$

The procedure of estimation of general parameters for sensitive variables is described below such as:

$$\breve{t}_{(a,b)} = \bar{Z}^a (s_z^2 - \sigma_s^2)^{\frac{b}{2}} \tag{1}$$

$$= \bar{Z}^a (1 + \epsilon_0)^a (s_z^2 - \sigma_s^2 + s_z^2 \epsilon_1)^{\frac{b}{2}} \tag{2}$$

Multiplying and neglecting cube and higher powers of the ∈

$$= t_{(a,b)}(1 + \epsilon_0 a + \frac{a(a-1)}{2}\epsilon_0^2 + R_{zy}\frac{b}{2}\epsilon_1 + \frac{b(b-2)}{8}R_{zy}^2 \epsilon_1^2 + \frac{ab}{2}\epsilon_0 \epsilon_1 R_{zy}) \tag{3}$$

Take the square of (3) on both edges and disregard the cube and tall powers of ∈

$$(\breve{t}_{(a,b)} - t_{(a,b)})^2 = t_{(a,b)}^2 (\epsilon_0 a + \frac{a(a-1)}{2}\epsilon_0^2 + R_{zy}\frac{b}{2}\epsilon_1 + \frac{b(b-2)}{8}R_{zy}^2 \epsilon_1^2 + \frac{ab}{2}\epsilon_0 \epsilon_1 R_{zy})^2 \tag{4}$$

$$= \frac{t_{(a,b)}^2}{n}(a^2C_z^2 + \frac{ab}{2}C_z\delta_{30}R_{zy}^2 + \frac{b^2}{4}R_{zy}^2(\delta_{40}-1)) \quad (5)$$

$$MSE(\breve{t}_{(a,b)}) = \frac{t_{(a,b)}^2}{n} f_1(a,b) \quad (6)$$

Where,

$$f_1(a,b) = (a^2C_z^2 + \frac{ab}{2}C_z\delta_{30}R_{zy}^2 + \frac{b^2}{4}R_{zy}^2(\delta_{40}-1))$$

3.1.1. Proposed estimator

a) Ratio estimator

The universal appearance of ratio estimation is specified as:

$$t_r = \bar{Z}^a(\sigma_z^2 - \sigma_s^2)^{\frac{b}{2}}(\frac{\bar{X}}{\bar{x}}) \quad (7)$$

The procedure of MSE of ratio estimator for general parameter estimation is given as:

$$= t_{(a,b)}(1 + \in_0 a + \frac{a(a-1)}{2}\in_0^2 + R_{zy}\frac{b}{2}\in_1 + \frac{b(b-2)}{8}R_{zy}^2 \in_1^2 + \frac{ab}{2}\in_0\in_1 R_{zy})[1 - \in_2 + \in_2^2] \quad (8)$$

Taking square on both sides and ignoring cube and higher powers of elements ∈

$$(t_r - t_{(a,b)})^2 = t^2{}_{(a,b)}(\in_0^2 a^2 + R_{zy}^2\frac{b^2}{4}\in_1^2 + \in_2^2 + \frac{ab}{2}\in_0\in_1 R_{zy} - 2a\in_0\in_2 - b\in_1\in_2 R_{zy}] \quad (9)$$

Where,

$$f_3(a,b) = [2a\delta_{zx}C_z + b\delta_{21}R_{zy}] \quad (10)$$

$$MSE(t_r) = \frac{t^2{}_{(a,b)}}{n}(f_1(a,b) + C_x[C_x - f_3(a,b]) \quad (11)$$

b) Exponential ratio estimator

The common structure of the exponential ratio estimate is presented as:

$$t_{er} = \breve{t}_{(a,b)}\exp(\frac{\bar{X} - \bar{x}}{\bar{X} - \bar{x}}) \quad (12)$$

$$t_{er} = t_{(a,b)}(1+ \in_0 a + \frac{a(a-1)}{2} \in_0^2 + R_{zy} \frac{b}{2} \in_1 + \frac{b(b-2)}{8} R_{zy}^2 \in_1^2 + \frac{ab}{2} \in_0 \in_1 R_{zy}) \exp(\frac{\bar{X} - \bar{x}}{\bar{X} - \bar{x}}) \quad (13)$$

$$t_{er} - t_{(a,b)} = \frac{t_{(a,b)}}{n} (C_Z^2 a^2 + R_{zy}^2 \frac{b^2}{4} (\varsigma_{40} - 1) + ab\varsigma_{30} C_Z R_{zy} + \frac{C_X^2}{4} - a\varsigma_{ZX} C_X C_Z - \frac{b}{2} C_X \varsigma_{21} R_{zy}) \quad (14)$$

After simplification it becomes,

$$MSE(t_{er}) = \frac{t^2_{(a,b)}}{n} (f_1(a,b) + \frac{C_x}{2} (\frac{C_x}{2} - f_3(a,b) \quad (15)$$

3.1.2. Specific cases

Some special cases of the estimation of ratio and ratio exponential are specified below:

Table-1: Ratio estimators

a	b	R	Estimator	Mean Square Error
1	0	1	$t_{r1} = \bar{Z} \cdot \left(\frac{\bar{X}}{\bar{x}} \right)$	$\frac{t^2_{(a.b)}}{n} \left[C_Z^2 + C_X \left\{ C_X - 2C_Z \rho_{ZX} \right\} \right]$
0	1	1	$t_{r2} = \left(\sigma_Z^2 - \sigma_S^2 \right)^{1/2} \cdot \left(\frac{\bar{X}}{\bar{x}} \right)$	$\frac{t^2_{(a.b)}}{n} \left[\frac{\left(\delta_{40} - 1 \right)}{n} + C_X \left\{ C_X - \delta_{21} \right\} \right]$
1	1	1	$t_{r3} = \bar{Z} \cdot \left(\sigma_Z^2 - \sigma_S^2 \right)^{1/2} \cdot \left(\frac{\bar{X}}{\bar{x}} \right)$	$\frac{t^2_{(a.b)}}{n} \left[\left\{ C_Z^2 + C_Z \delta_{30} + \frac{\left(\delta_{40} - 1 \right)}{4} \right\} + C_X \left\{ C_X - 2C_Z \rho_{ZX} + \delta_{21} \right\} \right]$

Table-2: Exponential ratio estimators

a	b	R	Estimator	MSE
1	0	1	$t_{er1} = \bar{Z} \cdot \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	$\frac{t^2_{(a.b)}}{n} \left[C_Z^2 + \frac{C_X}{2} \left\{ \frac{C_X}{2} - 2C_Z \rho_{ZX} \right\} \right]$
0	1	1	$t_{er2} = \left(\sigma_Z^2 - \sigma_S^2 \right)^{1/2} \cdot \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	$\frac{t^2_{(a.b)}}{n} \left[\frac{\left(\delta_{40} - 1 \right)}{n} + \frac{C_X}{2} \left\{ \frac{C_X}{2} - \delta_{21} \right\} \right]$
1	1	1	$t_{er3} = \bar{Z} \cdot \left(\sigma_Z^2 - \sigma_S^2 \right)^{1/2} \cdot \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right)$	$\frac{t^2_{(a.b)}}{n} \left[\left\{ C_Z^2 + C_Z \delta_{30} + \frac{\left(\delta_{40} - 1 \right)}{4} \right\} + \frac{C_X}{2} \left\{ \frac{C_X}{2} - 2C_Z \rho_{ZX} + \delta_{21} \right\} \right]$

4. Simulation studies

In this section, an elaborative arithmetical investigation was done to access the efficiency of the estimation considered in this research. A replication of classical research was offered below to show the effectiveness of the projected estimates. It was tried to work on 3 finite population's size of 1000 each from a bivariate standard populace with diverse covariance templates to signify the spreading of (Y, X). The scrambling variable S is reserved to be standard variant with mean equivalent to nil and standard deviation equivalent to ten percent of the SD of X. The testified reply is prearranged as $Z = Y + S$.

Whole replicated populaces have the hypothetical mean of $[Y, X]$ as $\mu = [2 \quad 2]$

The covariance matrices are specified as under:

Population 1:

$$\Sigma = \begin{bmatrix} 9 & 1.9 \\ 1.9 & 4 \end{bmatrix} \qquad \rho_{XY} = 0.3209$$

Population 2:

$$\Sigma = \begin{bmatrix} 6 & 3 \\ 3 & 2 \end{bmatrix} \qquad \rho_{XY} = 0.8684$$

For both the populations, study considered three sample sizes:

$$n = 100, 200 \; \& \; 300$$

The % relative efficiency (PRE) is intended from the subsequent equations:

$$PRE = \frac{MSE\left(t_{(a,b)}\right)}{MSE\left(t_{\alpha}\right)}$$

Where,

$$\alpha = tr1, tr2, tr3, ter1, ter2, ter3$$

The outcomes for the estimators under inquiry are displayed in Table-3.

5. Results

Table-3 proposed an estimator along with the sample mean, ratio estimator and exponential ratio. It also showed that the amount of MSE declined and the PRE improved with increasing the size of the sample for all the estimators. The ratio and exponential ratio kind estimations executed well because of the strong optimistic relationship between the investigated inquiry and AVs. From this simulation study as summarized in Table-3 it is shown that the proposed estimate is more competent than the conventional estimations for all the population at various levels of correlation. The proposed estimators performed well then all the previously given estimators.

Table-3: Estimator result for 1st population

N	ρ_{XY}	n	*Estimation*	*MSE Estimation*	*PRE*
1000	0.3209	100	$t_{(a,b)}$	0.3898	100
			t_{r1}	0.2036	191.43
			t_{r2}	0.0203	1921.09
			t_{r3}	0.0561	695.09
			t_{er1}	0.2745	142.03
			t_{er2}	0.0332	1173.96
			t_{er3}	0.0334	1165.53
		200	$t_{(a,b)}$	0.1978	100
			t_{r1}	0.1014	195.11
			t_{r2}	0.0081	2437.07
			t_{r3}	0.0238	831.63
			t_{er1}	0.1382	143.139
			t_{er2}	0.0139	1413.63
			t_{er3}	0.0149	1320.20
		300	$t_{(a,b)}$	0.1070	100
			t_{r1}	0.0544	196.44
			t_{r2}	0.0048	2193.10
			t_{r3}	0.0159	671.26
			t_{er1}	0.0744	143.78
			t_{er2}	0.0084	1270.48
			t_{er3}	0.0094	1128.12

The table-4 proposed an estimator along with the mean of sample, ratio estimate exponential ratio. It also showed that the amount of MSE lower down and the PRE amplified with increasing the size of sample for all the estimators. The ratio and exponential ratio type estimations executed well because of the strong optimistic connection between the inquired space and AVs.

From this simulation study as illustrated in Tables-3 and 4, it is shown that the proposed estimation worked superior as unlike empirically proposed estimators for all the population at the diverse level of correlation. The projected estimators were superior and well-rated as compared to the previously proposed estimations that were introduced by different statisticians. According to the findings in Table-4, the projected ratio and exponential ratio kind estimators significantly performed well for the general parameter's estimator in case of sensitive variables and non-responsive.

Table-4: Estimator result for 2nd population

N	ρ_{XY}	n	*Estimation*	*MSE Estimation*	*PRE*
1000	0.8684	100	$t_{(a,b)}$	0.4254	100
			t_{r1}	0.2153	197.50
			t_{r2}	0.0193	2205.16
			t_{r3}	0.0596	713.75
			t_{er1}	0.2951	144.13
			t_{er2}	0.0325	1306.32
			t_{er3}	0.0357	1188.82
		200	$t_{(a,b)}$	0.1736	100
			t_{r1}	0.0945	183.62
			t_{r2}	0.0089	1937.03
			t_{r3}	0.0285	609.20
			t_{er1}	0.1248	139.07
			t_{er2}	0.0146	1185.96
			t_{er3}	0.0161	1074.06
		300	$t_{(a,b)}$	0.0956	100
			t_{r1}	0.0513	186.45
			t_{r2}	0.0051	1847.52
			t_{r3}	0.0149	637.90
			t_{er1}	0.0684	139.86
			t_{er2}	0.0084	1131.98
			t_{er3}	0.0087	1089.17

6. Discussion

In the present study, the general parameter estimators for sensitive variables with the use of randomized response techniques and auxiliary information in the sampling domain for people in the field were proposed. It also articulated a mean estimator for the finite population. An investigation with a SRSWOR was the best-opted method and technique to be used for the projection of new estimators. The proposed mean estimator would be more effective to generate more friendly-oriented estimators in later studies. The present study only used an additive model for randomized response with a single scrambling variable.

The characteristics of the projected estimates are to be derived to 2nd category guess by the application of well-known Taylor and exponential chains. The outcomes of the current investigation are entirely supported by the idea to apply auxiliary information to help in guessing the estimates for a given attribute; it was observed that the study meticulously utilized the pertinent data to greater the competence of the judgmental method.

Numerical progress is recognized with the SRSWOR as a method. Besides, the suggestion is useful for two normally utilized information groups to evaluate the practicality of the established classes. The assessment is following the previously done study by Haq *et al.* (2017), who also found the same class of estimators with the dual application of auxiliary information. The outcome is also in line with study that also anticipated a novel better estimator for restricted population mean Haq *et al.* (2017).

7. Conclusion

The present study proposed general type estimation with sensitive variables by the utilization of non-sensitive auxiliary variables. The MSE expression is resultant from it. It was noticed that the projected estimators are extra effective, and researchers demonstrated that the competence of the anticipated estimator can be pretty considerable if the association between researched and auxiliary variable is in height. These outcomes are also reinforced by the previous research done on the presented variables. This study concludes that an appropriate application of the auxiliary information assist in escalating the accuracy of an estimation equally at the scheming stage and at the judgment stage. In any type of surveys, the absolute auxiliary information is often accessible with the sample casing. It is concluded that the achieved outcomes could be made better by the use of multifaceted sampling designs like stratified and two phase sampling technique. It was foresight by researchers that an identical approach can be employed to make estimations for population variance but this can be taken up by other statisticians for their future studies. The statisticians could use different randomized response models with two scrambling variables in the company of multifaceted AVs.

ORCID iD

Amjad Mahmood https://orcid.org/0009-0006-6396-2305

Nadia Mushtaq https://orcid.org/0000-0002-0652-0029

Farhad Hussain https://orcid.org/0000-0003-1399-6399

References

Adichwal, N. K., Ahmadini, A. A. H., Raghav, Y. S., Singh, R., & Ali, I. (2022). Estimation of general parameters using auxiliary information in simple random sampling without replacement. *Journal of King Saud University-Science*, *34*, 101754. https://doi.org/10.1016/j.jksus.2021.101754

Adichwal, N. K., Mishra, P., Singh, P., Singh, R., & Yan, Z. (2016). A two parameter ratio-product-ratio type estimator for population coefficient of variation based on SRSWOR. *J. Adv. Res. Appl. Math Stat.*, *1*(3&4), 1-5. https://doi.org/10.1007/s40819-015-0073-3

Adichwal, N. K., Sharma, P., & Singh, R. (2017). Generalized class of estimators for population variance using information on two auxiliary variables. *IJACM*, *3*(2), 651-661. https://doi.org/10.1007/s40819-015-0119-6

Adichwal, N. K., Sharma, P., Raghav, Y. S., & Singh, R. (2019). A class of Estimators for population mean utilizing information on auxiliary variables using two phase sampling schemes in the presence of non-response when study variable is an attribute. *Pakistan J. Stat*., *35*(3),187-196.

Aronow, P. M., Coppock, A., Crawford, F. W., & Green, D. P. (2015). Combining list experiment and direct question estimates of sensitive behaviour prevalence. *Journal of Survey Statistics and Methodology*, *3*, 43-66. https://doi.org/10.1093/jssam/smu023

Bahl, S., & Tuteja, R.K. (1991). Ratio and product type exponential estimator. *Information and Optimization Sciences*, *12*, 159–163.

Blair, G., &Imai. K. (2012). Statistical analysis of list experiments. *Political Analysis*, *20*, 47-77. https://doi.org/10.1093/pan/mpr048

Blair, G., Imai, K., & Lyall. J. (2014). Comparing and combining list and endorsement experiments: Evidence from Afghanistan. *American Journal of Political Science*, *58,* 1043-1063. https://doi.org/10.1111/ajps.12086

Bolfarine, H., & Zacks, S. (1992). *Prediction theory for finite populations*. Springer-Verlag.

Bowley, A. L. (1926). Measurements of precision attained in sampling. *Bull. Inst. Inte. Statist*., *22*, 1-62.

Bullock, W., Imai, K., & Shapiro, J. N. (2011). Statistical analysis of endorsement experiments: measuring support for militant groups in Pakistan. *Political Analysis*, *19*, 363-84. https://doi.org/10.1093/pan/mpr032

Chand, L. (1975). *Some ratio-type estimators based on two or more auxiliary variables*. Ph.D. Thesis, Iowa State University, Ames, Iowa. https://core.ac.uk/download/pdf/38912792.pdf

Cochran, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. *Journal of Agricultural Science*, *30*, 262-275. http://dx.doi.org/10.1017/S0021859600048012

Das, A. K., & Tripathi, T. P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya C*, *40*, 139-148.

Deville, J. C., & Sarndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*. 87, 376. https://doi.org/10.1080/01621459.1992.10475217

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

Gonzalez-Ocantos, E., de Jonge, C. K., Mele´ndez, C., Osorio, J., & Nickerson, D. W. (2012). Vote buying and social desirability bias: Experimental evidence from Nicaragua. *American Journal of Political Science*, *56*, 202-17. https://doi.org/10.1111/j.1540-5907.2011.00540.x

Grover, L. K., Kaur, P., & Vishawkarma G. K. (2012). Product type exponential estimators of population mean under linear transformation of auxiliary variable in simple random sampling. *Applied Mathematics and Computation*, 219, 1937-1946. https://doi.org/10.1016/j.amc.2012.08.036

Gupta, S., & Shabbir, J. (2008). On estimating in estimating the population mean in simple random sampling. *J. Appl. Stat.*, *35*(5), 559–566. https://doi.org/10.1080/02664760701835839

Hartley, H., & Ross, A. (1954) Unbiased Ratio Estimates. *Nature*, 174, 270-271. https://doi.org/10.1038/174270a0

Hansen, M. H., Hurwitz, W.N., & Madow, W.G. (1953). *Sample survey methods and theory* (Volume I). John Wiley and Sons.

Hansen, M. M., & Hurwitz, W. N. (1943). On the theory of sampling from finite populations. *Annals of Mathematical Statistics*, *14*, 333-362.

Haq, A., Khan, M., & Hussain, Z. (2017). A new estimator of finite population mean based on the dual use of the auxiliary information. *Communications in Statistics - Theory and Methods*, *46*(9), 4425–4436. https://doi.org/10.1080/03610926.2015.1083112

Hussain, Z., Shabbir, J., & Gupta, S. (2007). An alternative to Ryu et al. randomized response model. *Journal of Statistics & Management Systems*, *10*(4), 511-517.

Isaki, C. T. (1983). Variance estimation using auxiliary information. *Journal of American Statistical Association*, *381*, 117-123.

Jhajj, H. S., Sharma, M. K., & Grover, L. K. (2007). A wide and efficient class of estimators of population variance under sub-sampling scheme. *Model Assisted Statistics and Applications*, *2*(1), 41–51. https://doi.org/10.3233/MAS-2007-2106

Kadilar, C., & Cingi, H. (2004) Ratio estimators in simple random sampling. *Applied Mathematics and Computation*, *151*, 893-902. https://doi.org/10.1016/S0096-3003(03)00803-8

Kadilar, C., & Cingi, H. (2006). Improvement in variance estimation using auxiliary information. *Hacettepe Journal of Mathematics and Statistics*, *35*(1), 111-115.

Karp, J. A., & Brockington, D. (2005). Social desirability and response validity: A comparative analysis of over reporting voter turnout in five countries. *Journal of Politics*, 67, 825-40. https://doi.org/10.1111/j.1468-2508.2005.00341.x

Khoshnevisan, M., Singh, R., Chauhan, P., Sawan, N., & Smarandache, F. (2007). A general family of estimators for estimating population means using known value of some population parameter(s). *Far East J. Stat*., *22*(2), 181-191

Kiregyra, B. (1984). A regression-type estimator using two auxiliary variables and model of double sampling from finite populations. *Metrika*, *31*, 215-226. https://doi.org/10.1007/BF01915203

Lyall, J., Blair, G., & Imai, K. (2013). Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan. *American Political Science Review*, *107*, 679-705.

Martin, L. M., Leff, M., Calonge, N., Garrett, C., and Nelson, D.E. (2000). Validation of self-reported chronic conditions and health services in a managed care population. *American Journal of Preventive Medicine*, 18, 215-221. https://doi.org/10.1016/s0749-3797(99)00158-0

Mishra, P., & Singh, R. (2017). A new log-product type estimator using auxiliary information. *J. Sci. Res.*, 61 (1&2), 179-183

Mohanty, S. (1967). Combination of regression and ratio estimate. *Journal of Indian Statistical Association*, 5, 16-19.

Mukerjee, R., Rao, T. J., & Vijayan, K. (1987). Regression-type estimators using multiple auxiliary information. *Australian Journal of Statistics*, *29*(3), 244-254. https://doi.org/10.1111/j.1467-842X.1987.tb00742.x

Neyman, J. (1938). Contribution to the theory of sampling human populations. *Journal of the American Statistical Association*, *33*, 101-116. http://dx.doi.org/10.1080/01621459.1938.10503378

Olkin, I. (1958). Multivariate ratio estimation for finite populations. *Biometrika*, *45*, 154-165.

Onyango, R., Oduor, B., Odundo, F. (2022). Mean estimation of a sensitive variable under nonresponse using three-stage rrt model in stratified two-phase sampling. *Journal of Probability and Statistics*, *14*, 4530120. https://doi.org/10.1155/2022/4530120

Oral, E., & Kadilar, C. (2011). Robust ratio-type estimators in simple random sampling. *Journal of the Korean Statistical Society*, *40*(4), 457–467. https://doi.org/10.1016/j.jkss.2011.04.001

Pollock, K. H., & Yuksel, B. (1976). A comparison of three randomized response models for quantitative data. *Journal of the American Statistical Association*, *71*, 356, 884-886. https://doi.org/10.1080/01621459.1976.10480963

Raj, D. (1965). On a method of using multi-auxiliary information in sample surveys. *Journal of American Statistical Association*, *60*(309), 270-277. http://dx.doi.org/10.1080/01621459.1965.10480789

Rao, P. S. R. S., & Mudholkar, G. S. (1967). Generalized multivariate estimators for thee mean of a finite population. *Journal American Statistical Association*, *62*, 1009-1012. https://doi.org/10.1016/S0169-7161(88)06020-1

Robson, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio-type estimation. *J. Amer. Statist. Assoc*., *52*, 511-522. http://dx.doi.org/10.1080/01621459.1957.10501407

Rosenfeld, B., Imai, K., & Shapiro, J. (2016). An empirical validation study of popular survey methodologies for sensitive questions. *American Journal of Political Science*, *60*, 783-802.

Roy, D. C. (2003). A regression-type estimator in two-phase sampling using two auxiliary variables. *Pakistan Journal of Statistics*, *19*(3), 281- 290.

Rueda, M. and Cebrian, A. A. (1996). Repeated substitution method the ratio estimator for the population variance. *Metrika*, *43*, 101-105

Ryu, C., Hu, C. H., Locy, R. D., & Kloepper, J. W. (2005). Study of Mechanism for Plant growth promoted by rhizobacteria in Arabidopsis thaliana. *Plant and Soil*, *268*, 285-292. https://doi.org/10.1007/s11104-004-0301-9

Saleem, I. (2017). *Some contributions to scramble randomized response technique using auxiliary variable for parameters in finite population*. PhD Thesis, National College of Business and Economics, Lahore, Pakistan.

Samiuddin M., & Hanif M. (2007). Estimation of population mean in single- and two-phase sampling with or without additional information. *Pak. J. Statist*., 23(2), 99-118.

Shabbir, J., & Gupta, S. (2010). On estimating finite population mean in simple and stratified random sampling. *Communications in Statistics Theory and Methods*, *40*(2), 199-212.

Shabbir, J., & Gupta, S. (2007). On estimating the finite population mean with known population proportion of an auxiliary variable. *Pakistan J. Stat*., *23*, 1-9.

Sharma, P., Verma, H. K., Adichwal, N. K., & Singh, R. (2017). Generalized estimators using characteristics of Poisson distribution. *Int. J. Appl. Comput. Math*, *3*(2), 745-755. https://doi.org/10.48550/arXiv.1502.02343

Shukla, N. D. (1971). Systematic sampling and product method of estimation. *Proceeding of the all-India Seminar on Demography and Statistics, Varanasi, India*.

Singh, H. P., & Solanki, R. S. (2011). Generalized ratio and product methods of estimation in survey sampling. *Pak. J. Stat. Oper. Res*., *7*, 245-264. https://doi.org/10.18187/pjsor.v7i2.189

Singh, M.P. (1967). Ratio cum product method of estimation. *Metrika*, *12*, 34-42.

Singh, R., Chauhan, P., Sawan, N. & Smarandache, F. (2011). Improved exponential estimator for population variance using two auxiliary variables. *Italian Journal Pure and Applied Mathematics*, *28*(N), 101-108. https://doi.org/10.48550/arXiv.0902.0126

Singh, R., Kumar, M., Chaudhary, O. K., Kadilar, C. (2009). Improved Exponential Estimator in Stratified Random Sampling. *Pak. J. Stat. Oper. Res*., *5*(2), 67-82.

Singh, R., Kumar, M., & Smarandache, F. (2008). Almost unbiased estimator for estimating population mean using known value of some population parameter(s). *Pak.J. Stat. Oper.Res*., *4*(2), 63-76. https://doi.org/10.18187/pjsor.v4i2.50

Singh, H.P., &Singh, R. (2001). Improved ratio-type estimator for variance using auxiliary information. *J.I.S.A.S.*, *54*(3), 276-287.

Solanki, R. S., & Srivastava, S. K. (1965). An estimation of mean of a finite population using several auxiliary variables. *Journal of Indian Statistical Association*, *3*, 189-194.

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. https://doi.org/10.1080/15598608.2010.10411999

Srivastava, S. K. (1967a). An estimator using auxiliary information in sample surveys. *Journal of Indian Statistical Association*, 16, 121-132.

Srivastava, S. K. (1967b). A generalized estimator for the mean of finite population using multi-auxiliary information. *Journal of Indian Statistical Association*, *16*, 121-132

Srivastava, S. K., & Jhajj, H. S. (1980). A class of estimators using auxiliary information for estimating finite population variance. *Sankhya C, 4*, 87–96

Srivastava, S. K., Khare, B. B., & Srivastava, S. R. (1990). A generalized chain ratio estimator for mean of finite population. *Journal of Indian Society of Agricultural Statistics*, *42*, 108-117.

Srivastava, S.K., & Jhajj, H. S. (1981). A Class of Estimators of the Population Mean in Survey Sampling Using Auxiliary Information. *Biometrika*, *68*, 341-343.

Srivastava, S. K., & Jhajj, H. S. (1986). On the estimation of finite population correlation coefficient. *J. Ind. Soc. Agric. Stat.*, *38*, 82-91

Srivenkataramana, T. (1980). A dual to ratio estimator in sample surveys. *Biometrika*, *67*(1), 199- 204. http://dx.doi.org/10.1093/biomet/67.1.199

Subramani, J., & Kumarapandiyan, G. (2012a). Variance estimation using median of the auxiliary variable. *International Journal of Probability and Statistics*, *1*(3), 36–40. https://doi.org/10.5923/j.ijps.20120103.02

Subramani, J., & Kumarapandiyan, G. (2012b). Variance estimation using quartiles and their functions of an auxiliary variable. *International Journal of Statistics and Applications*, *2*(5), 67–72. https://doi.org/10.5923/j.statistics.20120205.04

Swain, A.K.P.C. (1964). The use of systematic sampling ratio estimate, *Journal of the Indian Statistical Association*, *2*, 160–164.

Tracy, D. S., Singh, H. P., & Singh, R. (1996). An alternative to ratio-cum-product estimator in sample surveys. *J.S.P.I.*, *53*, 375-387. https://doi.org/10.1016/0378-3758(95)00136-0

Tripathi, T. P. (1970). *Contributions to the sampling theory in multivariate information*. Ph.D. Thesis, Punjabi University, Patiala, India.

Tripathi, T.P. (1973). Double sampling for inclusion probabilities and regression method of estimation. *J. Ind. Statist. Assoc.*, *10*, 33-46.

Tripathi, T. P. (1976). On double sampling for multivariate ratio and difference method of estimation. J. *Ind. Statist. Assoc.*, *33*, 33-54.

Tripathi, T. P., Singh, H. P., Upadhyaya, L. N. (1988). A generalized method of estimation in double sampling. Estimation of mean on second occasion using pps sampling and multivariate information. *J. Indian Soc. Agric. Stat.*, *26*, 91–101.

Upadhyaya, L. N., Singh, H. P. Chatterjee, S., & Yadav, R. (2011). Improved ratio and product exponential type estimators. *Journal of Statistical Theory and Practice*, *5*(2), 285-302. https://doi.org/10.1080/15598608.2011.10412029

Van der Heijden, P. G., van Gils, G., Bouts, J., & Hox. J. J. (2000). A comparison of randomized response, computer-assisted self-interview, and face-to-face direct questioning: Eliciting sensitive information in the context of welfare and unemployment benefit. *Sociological Methods and Research*, *28*, 505-37. https://doi.org/10.1177/0049124100028004005

Warner, S. L. (1971). Linear Randomized Response Models. *Journal of the American Statistical Association, 66*, 884-888.

Warner, S. L. (1965). Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, *60*(309), 63-69. https://doi.org/10.1080/01621459.1965.10480775

Wolter, F., & Preisendorfer, P. (2013). Asking sensitive questions: An evaluation of the randomized response technique versus direct questioning using individual validation data. *Sociological Methods and Research*, *42*, 321-53. https://doi.org/10.1177/0049124113500474