FINAL REPORT OF
RESEARCH AND SEED GRANT
[No. RA/23/7368-74]

FORCED SCRAMBLED OPTIONAL RANDOMIZED RESPONSE MODEL FOR CONCURRENT ESTIMATION OF MEANS OF SENSITIVE VARIABLE

A Project Submitted to the
Dean Research Studies
University of Jammu

Submitted by

DR. SUNIL KUMAR
Assistant Professor
Department of Statistics
University of Jammu
Jammu-180006, J&K, India

DEPARTMENT OF STATISTICS

NEW UNIVERSITY CAMPUS
UNIVERSITY OF JAMMU
JAMMU (J & K) – 180006 (INDIA)

Ref No. Stat./..24\744.......................   Dated ....10-01-24.....

10/1/24

Dr Sunil Kumar, M.Sc., M.Phil., Ph.D.
Assistant Professor
Department of Statistics, University of Jammu, Jammu

To
Dean Research Studies
University of Jammu, Jammu

Respected Madam,

Subject: Submission of Final Report for 'Research and Seed Grant'

Greetings, I express my heartfelt thanks for releasing and approving the Research and Seed Grant for Professor/Associate Professor/Assistant Professor for the session 2023-2024. Grant released vide order no.: RA/23/7368-74 dated 24/01/2023. As per the order no. RA/24/4700-4800 dated 02/01/2024. I am herewith submitting the detailed progress report through the departmental research committee.

Thanking you

Yours Sincerely

10/01/2024

Dr. Sunil Kumar
Principal Investigator
Department of Statistics
University of Jammu
Jammu

No. PGD/STAT/24/929
Date :- 14/03/2024

No. PGD/...
Date :- 14-03-2024

UTILIZATION CERTIFICATE

WITH UP-TO-DATE STATEMENT OF EXPENDITURE

(February 2023 to March 2024)

1. Sanction letter no.: **RA/23/7368-74; dated 24/01/2023**
2. Title of the Project: **Forced Scrambled Optional Randomized Response Model for Concurrent Estimation of Means of Sensitive Variable**
3. Name of the PU: **Dr Sunil Kumar**
4. Name of the Department: **Department of Statistics, University of Jammu**
5. Total Project Cost: **Rs. 2,00,000.00**
6. Statement of Expenditure:

S. No.	Sanction Heads	Funds Allocated (Rs)	Expenditure (Rs)	Balance (Rs)
1.	Purchase of Minor Equipment	70,000	51,993.00	18,007.00
2.	Consumables/Chemicals/Glassware etc.	80,000	0.00	80,000.00
3.	Contingency	50,000	31,255.00	18,745.00
	Grand Total	**2,00,000**	83248	**1,16,752**

Certified that out of **Rs. 2,00,000/-** (Two Lakhs only) of grant-in-aid, sanctioned vide order no. RA/23/7368-74, dated 24/01/2023 during the year of 2023 in favour of Dr. Sunil Kumar (PI), a sum of **Rs. 83,248.00** has been utilized for the purpose of research for which it was sanctioned and the balance of **Rs. 1,16,752.00** remaining unutilized.

Signature of PI
Date: 12/03/2024

Dr. Sunil Kumar
Principal Investigator
Department of Statistics
University of Jammu
Jammu

Signature of Joint Registrar
Date:

Deputy Registrar (Grants)
University of Jammu

15/3

Signature of Deputy Registrar
Date:

Joint Registrar (Finance)
University of Jammu

15/3

OFFICE OF DEAN RESEARCH STUDIES
UNIVERSITY OF JAMMU

ORDER

Based on the recommendations of the Committee constituted for the purpose vide order No. RA/3977-92 dated 05.12.2022 and also on the recommendations of the Dean of the Faculty concerned, sanction is hereby accorded to the payment of Rs. 200000/- as financial assistance in favour of Prof./Dr. Sunil Kumar, Department of Statistics as per the details given below out of the **Research & Seed Grant** for Professor / Associate / Assistant Professor, under the Head **'Quality Assurance Fund (DIQA)'** as per order No. Fin./2022-23/3338-42 dated 16.09.2022:-

a) Hiring of Services / Honorarium for experts	:	-
b) Equipment (Repair) or any accessory, if needed, to the existing equipment	:	-
c) Purchase of Minor Equipment	:	70,000/-
d) AMC's of existing Equipment	:	
e) Consumables/Chemicals/Glassware etc.	:	80,000/-
f) Contingency	:	50,000/-
g) Field work	:	-
h) Any other item		-
Total	:	200000/-

You are required to meet the said expenditure as per University norms. The Principal Investigator (PI) shall submit the bills for pass & payment as per the existing GFR/GeM guidelines to the Grant Section. The quantum of assistance sanctioned is required to be exhausted/utilized within a period of one year starting from the date of issue of order. Utilization certificate will be submitted after completion of the project.

A detailed report of project shall mandatorily be submitted by PI to the office of the Dean Research Studies with a clear statement on whether the said project has enabled the PI to put up a bigger proposal for funding to any national funding agency.

24/1/2023
Assistant Registrar (R&A)

No. No. RA/23/7368-74
Dated: 24/01/2023

24/1 24/01

Copy to:

1. Special Secretary to the Hon'ble Vice-Chancellor.
2. Sr. P.A. to DAA/DRS/Registrar/DIQA.
3. Dean of the Faculty concerned.
4. HOD concerned.
5. Principal Investigator.
6. Joint Registrar (Finance).
7. Deputy Registrar (Grants).

DEPARTMENT OF STATISTICS
NEW UNIVERSITY CAMPUS
UNIVERSITY OF JAMMU
JAMMU (J & K) – 180006 (INDIA)

Ref. No. Stat./ 24\743 Dated: 10-01-24

Minutes of the Departmental Research Committee (DRC)

The meeting of the Departmental Research Committee (DRC) was held on 10/01/2024 at 11:00 am in the office of the Head, Department of Statistics, University of Jammu. The following were present:

1.	Prof Parmil Kumar	Convener
2.	Prof Rahul Gupta	Member
3.	Prof J P Singh Joorel	- do -
4.	Prof Pawan Kumar	- do -
5.	Dr V K Shivgotra	- do -

Agenda: to discuss the detailed progress report of the research and seed grant sanctioned in favour of Dr Sunil Kumar

Dr. Sunil Kumar has presented the outcomes of the report and the deliberations on the outcome of the report were made and after detailed discussions it was resolved unanimously that Dr Sunil Kumar has carried out the work on the topic **"Forced Scrambled Optional Randomized Response Model for Concurrent Estimation of Means of Sensitive Variable"** and the work done has been carried out adhering to the guidelines issued by the Dean Research Studies, University of Jammu, Jammu. Also, the work done by Dr Sunil Kumar is of utmost importance and applicable to directly several practical situations. Thus, the report is considered to be of high satisfaction by the DRC and is forwarded to DRS for further assessment.

Head,
Prof Parmil Kumar
Department of Statistics
University of Jammu, Jammu

Head
Department of Statistics
University of Jammu
Jammu

Contents

List of Tables

List of Figures

Abbreviations

RRT: Randomized Response Technique

ORRT: Optional Randomized Response Technique

FQORRT: Forced Quantitative Optional Randomized Response Technique

FRORRT: Forced Re-scrambled Optional Randomized Response Technique

RE: Relative Efficiency

Chapter 1

Introduction

1.1 Introduction and Background

One of the primary purposes of statistical research is to determine the true value of population parameters. Sampling is frequently used in our everyday activities. For example, suppose in a store we evaluate the quality of rice, sugar, wheat, maize or any other product by taking a quantity from the bag and deciding to buy it or not. Sampling is the process, technique, or act of selecting a suitable sample or a representative section of a population in order to determine the parameters or characteristics of the entire population. Because of their low cost, speed, precision, and validity, sample surveys are utilized in business and industry, agriculture, scientific, social, and economic fields, among others. Regrettably, if the population is large, collecting data from every member of the population would be prohibitively expensive or time-consuming. Rather than conducting a census, we can collect data from a sample and use the sample statistics to draw conclusions about the target population. It can usually be reduced by increasing the sample size. Non-sampling errors can be caused by a variety of issues, including respondent errors, measurement errors, non-response, and so on. As a result, the conclusions will be acceptable only if the sample actually represents the population and the sample responses are truthful.

Otherwise, the sample is skewed, and the study's findings are untrustworthy.
Sometimes, we also have several types of survey methods that are widely employed, such as email surveys, phone surveys and personal interview surveys. Email and phone surveys are less expensive, however they have a significant non-response rate. The issue of non-response may result in some participation bias. People who are passionate about a problem, for example, are more likely to participate and their viewpoint may not be representative of the entire population. People are less likely to refuse a personal interview survey than the other two approaches, although it is more expensive. A personal face-to-face interview may also generate social desirability response bias if the survey question is sensitive. For illustration, if a survey question asks,"What is your salary", "Have you ever taken illicit drugs?","How frequently do you gossip about others close to you?" or"Have you ever lied about your income on your tax returns?" most individuals will try to present themselves in a socially desirable light, therefore their answers may be slanted toward what they believe is socially desirable.
In addition to participation bias and social desirability response bias, there are some other non-sampling errors that will affect parameter estimation, such as measurement errors caused by definition differences or misconceptions. As a result, dealing with these issues is critical when we estimate the population parameters of a sensitive variable.

1.2 Objectives

This study will be conducted to determine forced scrambled optional randomized response model for concurrent estimation of mean of sensitive variable. Specifically, this study aims to the following

1. To develop an estimator for the estimation of sensitive variable using forced quantitative optional randomized response (FQORR) models.

2. To develop an estimator for the estimation of sensitive variable using forcibly re-scrambled optional randomized response (FRORR) model.

3. To compare the developed estimators in 1 and 2 with existing or competing estimators using ORRT models.

4. To check the efficiency of the proposed estimators in realistic environment, simulation study will be carried out to support theoretic findings.

1.3 Review of Literature

An innovative approach to the problem of estimating the proportion of a sensitive characteristic, like induced abortion, drugs consumption, gambling, etc., by making use of a randomization device, which was first developed by Warner [38], dealt only with qualitative variable. The method developed by Warner [38] is inventive because it makes use of a simple randomization device, such as a deck of cards, spinner etc., and because its use is easy for both the interviewers and interviewees. Fox [11], Chaudhuri, Christofides and Rao [8], Chaudhuri and Christofides [7], Chaudhuri [6], Tracy and Mangat [37], and Fox and Tracy [12] provides detailed reviews of the research on RRT. Further, Warner [39] model which features the quantitative additive version and is further expanded by Pollock and Bek's [31].

Furthermore, Eichhorn and Hayre's [10] introduced a multiplicative scrambling RRT model for obtaining sensitive quantitative data. Later on, Gupta et al. [17] modified Eichhorn and Hayre's [10] scrambling RRT model and developed Optional Randomized Response Technique (ORRT) and showed that ORRT perform better than non-optional RRT models. Based on this result Gupta et al. [19] improved Sousa et al. [33] by using ORRT scrambling model. Additionally, this work is extended by Gupta et al. [21], Diana and Perri's [9], Gupta et al. [18, 20, 22, 23], Kalucha et al. [24], Zhang et al. [41], Zhang et al. [43] and among others.

Recently Kumar and Kour [27, 28] have studied a mean estimation of sensitive variable in the simultaneously presence of non-response and measurement error in simple and two-phase sampling. Likewise, Kumar et al. [29] also developed an improved ratio-cum-product estimator with non-response and measurement error by utilizing ORRT models: a sensitive estimation approach.

In view of the well-established research interest in the topic of sensitive variables, it is important to study the idea of simultaneously estimating the mean of a sensitive variable by using forced quantitative optional randomized response (FQORR) models by taking inspiration from Ahmed et al. [2] and the other refers to the simultaneous estimation of two means by making use of the forced quantitative randomized response (FRQRR) model of Gjestvang and Singh [13] but then re-scrambling the scrambled scores by following Ahmed et al. [2]. This concept of forcing and re-scrambling previously scrambled responses appears to be novel in the field of optional randomized response techniques. The proposed forced quantitative optional randomized response model and forced re-scrambled optional randomized response model performance has been studied both analytically and empirically.

1.4 Glimpse of Project

The project is comprises of 3 chapters.

Chapter 1 entitled **"Introduction"**. In this chapter, we have discussed the background of the research, concepts, objectives of the study and review of literature.

Chapter 2 entitled **"Forced Quantitative ORRT for the Estimation of Mean of a Sensitive variable"**. A general approach for eliciting an optional randomized response from a sample of individuals in order to estimate the population mean of a sensitive variable by using forced quantitative randomized response technique (FQORRT), which consists of a true response, two scrambling variable(s) and a fixed factor are presented in this chapter. The unbiasedness and variance properties of the proposed estimator are examined both theoretically and empirically. For improvement, we choose the proposed FQORR model as it offers an estimator of the mean and sensitivity level of a sensitive variable and outperforms all considered competitors i.e. Eichhorn and Hayre model [10], Bar-Lev, Bobovitch, and Boukai (BBD) model [5], Gjestvang and Singh model [13], Gupta et al. [17, 19, 21] models. The findings are confirmed through simulation study reveals that the suggested model is preferably chosen over several of the existing models that are considered in literature.

Chapter 3 entitled **"Mean Estimation of Sensitive Variable: A Forced Re-scrambled Optional Randomized Response (FRORR) Approach"**. This chapter introduces a novel approach termed as Forcibly re-scrambled optional randomized response technique (FRORRT) designed for estimating the mean of sensitive variable while protecting respondents privacy which consists of a true response, two scrambling variables and a fixed factor which is chosen by the investigator based on prior experiences but then re-scrambling the scrambled scores because the concept of re-scrambling responses that have already undergone scrambling appears to be a novel approach within the area of optional randomized response sampling. The unbiasedness and variance properties of proposed FRORR model are studied both theoretically as well as empirically. To enhance our approach, we opt for the FRORR model because it provides estimates for both the mean and sensitivity level of a sensitive variable and also outperforms the other considered models. A simulation study is also conducted which demonstrates that the outcomes of proposed model is favoured over various existing models under consideration in the literature.

CHAPTER

2

Forced Quantitative ORRT for the Estimation of Mean of a Sensitive variable

2.1 Introduction

Due to privacy concerns, a respondent may be hesitate to reveal the truth or provide erroneous information in order to obtain sensitive data such as induced abortions, drug addiction, HIV infection status, incidence of domestic violence, income, under-reported tax, the status relative to medical conditions etc. Under these scenarios, when using the direct technique of interview (asking questions directly to respondents),the respondents frequently make false responses or even refuse to respond due to social censure or fear (refer Arnab [3]) are likely to contain response bias. In such cases, Warner [38] introduced the randomized response method that can be utilized to collect more trustworthy data and protect respondent anonymity. Following the innovative work of Warner [38], many extensions are presented such as Fox and Tracy [12], Gjestvang and Singh [13, 14], Fox [11], Singh and Gorey [32] Tarray et al. [34], Ahmed et al. [1, 2] and among others, deals with only qualitative data.

For quantitative data such as income of person, tax dodging, no. of students cheat in an exam etc., Eichhorn and Hayre [10] developed a multiplicative randomized response model by using scrambling variable to estimate the population mean of sensitive quantitative variable. Further, Gupta et al. [17] proposed ORRT model which is based on a very simple concept that a query can be sensitive for one person but not for other. In ORRT, the investigator requested to the interviewee to give a scramble response if they thought the question is sensitive or else give truthful responses. Numerous authors including Gupta et al. [19, 21], Mushtaq et al. [30], Khalil et al. [25, 26] Zhang et al. [42, 43], Tiwari et al [35], Tiwari et al. [36] Kumar and Kour [28], Zapata et al. [40] Kumar et al. [29], Azeem et al. [4] and so on suggested an ORRT model for estimating the population mean of sensitive variable when the auxiliary variable is sensitive or non-sensitive.

As driven by previous discussions and Gjestvang and Singh [13], the main goal of this chapter is to develop an forced quantitative ORRT model for estimating the population mean of sensitive variable which is different from the other optional randomized response technique's (ORRT), in that, we use value of fixed factor which is chosen by the investigator based on prior experiences. The rest of the chapter are arranged in such a manner that review of relevant models are discussed in section 2.2. Section 2.3 describes the proposed FQORRT model. In section 2.4 and 2.5, an attempt has been made to the compare the proposed estimator with the existing estimators along with the simulation studies in support of the proposed theoretical results. The concluding remarks are then elaborated in section 2.6.

2.2 Review of relevant models and strategies

To predict the population mean of μ_y, we evaluated several of the existing models with their variances in the context of RRT and ORRT are mentioned below

2.2.1 Eichhorn and Hayre's [10] Model

Eichhorn and Hayre [10] proposed a scrambled randomised response approach for determining the mean 'μ_y' and variance 'σ_y^2' of sensitive study variable Y. After that, each sample respondent is encouraged to use a randomization device to create a random number, S (say), from some pre-assigned distribution such as Uniform, Poisson, Binomial and so on. The distribution of the random variable S, also called scrambling variable, is assumed to be known and the mean 'μ_y' and variance 'σ_y^2' of scrambling variables are also assumed to be known. The i^{th} respondent in a sample of size n is asked to report the value $Z_{i(EH)} = S_i Y_i$ as a scrambled response on the sensitive variable Y. An unbiased estimator of population mean of μ_y of Eichhorn and Hayre [10] model is given by

$$\mu_{y_{(EH)}} = \frac{1}{n}\sum_{i=1}^{n} Z_{i(EH)} \tag{2.2.1}$$

with variance

$$V(\mu_{y_{(EH)}}) = \frac{1}{n}\left[\sigma_y^2 + C_\gamma^2(\sigma_y^2 + \mu_y^2)\right] \tag{2.2.2}$$

where $C_\gamma^2 = \sigma_s^2/\mu_s^2$ denotes the known coefficient of variation of the scrambling variable S, (μ_s, σ_s^2) is the mean and variance of the scrambling variable and $C_y^2 = \sigma_y^2/\mu_y^2$.

2.2.2 Gupta et al. [17] Model

The optional RRT model of Gupta et al. [17] is a modification of the Eichhorn and Hayre [10] model where the reported response is given by

$$Z_{G_0} = \begin{cases} Y & \text{with probability 1-W} \\ SY & \text{with probability W,} \end{cases} \tag{2.2.3}$$

where W is the sensitivity level of the question and S is a scrambling variable, independent of Y, with unit mean. It can be proved easily that $E(Y) = E(Z)$ which suggests that μ_y can again be estimated by $\mu_{y_{(G_0)}} = \bar{Z}_{(G_0)}$. The variance of Gupta et al. [17] estimator is given by

$$V(\mu_{y_{(G_0)}}) = \frac{1}{n}\left[\sigma_y^2 + W\frac{\sigma_s^2}{\mu_y^2}(\sigma_y^2 + \mu_y^2)\right] \tag{2.2.4}$$

The value of W will be close to 1, if a question in the survey is more sensitive then more people will report scrambled responses. And if the value of W will be close to 0 then the question is not very sensitive. Thus W is a measure of the level of sensitivity of the question in the personal interview surveys.

2.2.3 Bar-Lev et al. [5] Model

Bar-Lev et al. [5] developed a modification on the Eichhorn and Hayre [10] randomized response technique, which is known as BBB method. The probability mass function of the responses in the BBB approach is given by

$$Z_{i(BBB)} = \begin{cases} Y_i S & \text{with probability (1-p)} \\ Y_i & \text{with probability p,} \end{cases} \tag{2.2.5}$$

Thus, under the BBB model, each respondent is requested to spin a spinner, the

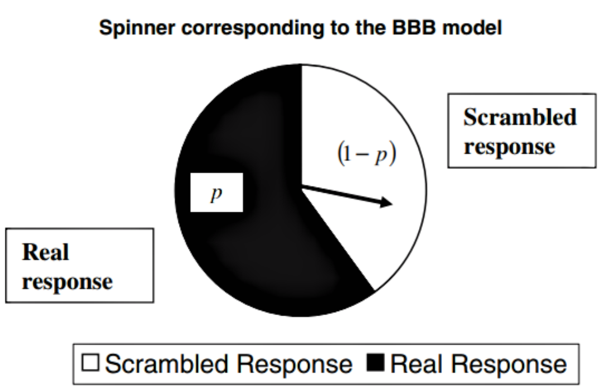

Figure 2.1: BBB randomized response device.

result of which is hidden to the interviewer. If the spinner falls with probability p in the shaded area, the respondent is asked to provide the genuine real response to the value of the sensitive variable, say Y_i. If the spinner falls with probability $(1-p)$ in the non-shaded area then the respondent is requested to respond with a scrambled response, say $Y_i S$, where S is a scrambling variable with a known distribution. Based on BBB model, an unbiased estimator of population mean μ_y is given by

$$\mu_{y(BBB)} = \frac{1}{n(1-p)\mu_s + p} \sum_{i=1}^{n} Z_{i(BBB)}$$

with variance

$$V(\mu_{y(BBB)}) = \frac{\mu_y^2}{n}\left[C_y^2 + (1 + C_y^2)C_s^2(p)\right] \tag{2.2.6}$$

where $C_s^2(p) = \frac{(1-p)\mu_s^2(1+C_\gamma^2)+p}{[(1-p)\mu_s+p]^2} - 1.$

2.2.4 Gjestvang and Singh [13] Model

A new model of forced quantitative randomised response (FQRR) is suggested by Gjestvang and Singh [13]. In their suggested FQRR model, each respondent is asked to use a randomization device, such as a spinner (or a deck of cards), which consists three types of statements:

1. Report the real value of the sensitive variable, say Y_i.
2. Report the scrambled response, say Y_iS and
3. Report a fixed value that is already written on the spinner (or card), say F, with proportions (p_1, p_2, p_3) such that $p_1 = p_2 = p_3 = 1$. The FQRR model is given by

$$Z_{i(GS)} = \begin{cases} Y_i & \text{with probability} \quad p_1 \\ Y_iS & \text{with probability} \quad p_2 \\ F & \text{with probability} \quad p_3, \end{cases} \tag{2.2.7}$$

where Y_i signifies the genuine value of the sensitive variable, Y_iS denotes the scrambled value, and F denotes the interviewer's fixed response.

An unbiased estimator of the population mean μ_y proposed by Gjestvang and Singh [13] is given by

$$\mu_{y(GS)} = \frac{\frac{1}{n}\sum_{i=1}^{n} Z_{i(GS)} - p_3F}{(p_1 + p_2\mu_s)}$$

with variance

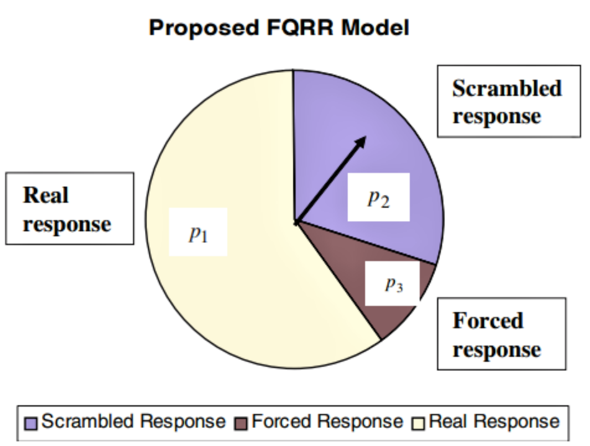

Figure 2.2: Gjestvang and Singh [13] proposed Forced quantitative randomized response (FQRR) model.

$$V(\mu_{y(GS)}) = \frac{1}{n(p_1 + p_2\mu_s)}\big[\{p_1 + p_2(\sigma_s^2 + \mu_s^2) - (p_1 + p_2\mu_s)^2\}(\sigma_y^2 + \mu_y^2) + p_3(1 - p_3)$$
$$F^2 - 2p_3F(p_1 + p_2\theta)\mu_y\big] + \frac{\sigma_y^2}{n} \tag{2.2.8}$$

2.2.5 Gupta et al. [21] Model

Under Gupta et al. [21] model, reported responses (Z_{G_i}; $i = 1, 2$), in the two sub-samples, are given by

$$Z_{(G_i)} = \begin{cases} Y & \text{with probability} \;\; T + (1-T)(1-W) \\ YS_i & \text{with probability} \;\; (1-T)W, \end{cases} \quad i = 1, 2 \tag{2.2.9}$$

where S_i, $i = 1, 2$ are independent scrambling variables (both independent of Y) with means μ_{s_i} and variances $\sigma_{s_i}^2$, respectively. In the model described in (8), note that a proportion (T) of the respondents provides truthful responses. From the remaining respondents, a proportion (W) provides scrambled responses and the rest provide truthful responses and $E(Z_{(G_i)}) = \mu_y + \mu_{s_i}W(1-T)$ where $\mu_{s_i} = E(S_i)$; $i = 1, 2$.

An unbiased estimator of Gupta et al. [21] model is given by

$$\mu_{y_{G_i}} = \frac{\mu_{s_1}\bar{Z}_{G_1} - \mu_{s_2}\bar{Z}_{G_2}}{\mu_{s_1} - \mu_{s_2}}$$

with variance

$$V(\mu_{y_{G_i}}) = \frac{1}{(\mu_{s_1} - \mu_{s_2})^2}\left(\mu_{s_2}^2 \frac{\sigma_{z_1}^2}{n_1} + \mu_{s_1}^2 \frac{\sigma_{z_2}^2}{n_2}\right) \tag{2.2.10}$$

2.2.6 Gupta et al. [19] Model

According to Gupta et al. [19] model, each selected respondent provides an additively scrambled response for Y if they consider the question sensitive, and a truthful response otherwise. The reported response for the study variable can be written as $Z_{(G_{yw})} = Y + ST$, where T is a Bernoulli random variable with parameter W and S is a scrambling variable with zero mean and known variance σ_s^2. The expected value of the observed response $Z_{(G_{yw})}$ is $E(Z_{(G_{yw})}) = E(Y + ST) = E(Y) = \mu_{yw}$. Therefore, an unbiased estimator of the study variable Y is defined as

$$\mu_{y(G_{yw})} = \frac{1}{n}\sum_{i=1}^{n} Z_{(G_{yw})}$$

with variance

$$V(\mu_{y(G_{yw})}) = \frac{1-f}{n}(\sigma_y^2 + W\sigma_s^2) \tag{2.2.11}$$

where $f = n/N$ is the sampling fraction.

2.3 Proposed Forced Quantitative ORRT Model

A modification of Gjestvang and Singh [13] model by adopting a new forced quantitative optional randomized response (FQORR) model is developed in this study in which each respondent is asked to use a randomization device, such as a spinner (or a deck of cards), which contains three sorts of statements i.e. Firstly, report the real value of the sensitive variable (Y_i). Secondly, report the scrambled response i.e $S_1Y + S_2$ and then report a fixed value which is already written on the spinner (or card), say F, with proportions p_1, p_2 and p_3 respectively, such that $p_1 = p_2 = p_3 = 1$ and (S_1, S_2) are two scrambling variables with variances $\sigma^2_{s_1}$ and $\sigma^2_{s_2}$, respectively and we take S_1 with a mean (μ_{s_1}) of 1 and S_2 with a mean (μ_{s_2}) of 0.

Analytically, the probability mass function for the i^{th} response in the FQORR model is given by

$$Z_{(F_i)} = \begin{cases} Y_i & \text{with probability } \; p_1 \\ S_1 + Y_i S_2 & \text{with probability } \; p_2 \\ F & \text{with probability } \; p_3, \end{cases} \tag{2.3.1}$$

where Y_i signifies the true value of the sensitive variable, $(S_1 + Y_i S_2)$ denotes the scrambled value, and F denotes the interviewer's forced or fixed response.

Taking the expected value of $Z_{(F_i)}$, we get

$$E(Z_{(F_i)}) = \mu_{y_i}(p_1 + p_2) + p_3 F$$

Then, the proposed unbiased estimator of the population mean μ_{y_i} is given as

$$\mu_{y_{(F_i)}} = \frac{\frac{1}{n}\sum_{i=1}^{n} Z_{(F_i)} - p_3 F}{p_1 + p_2} \tag{2.3.2}$$

The variance of the proposed unbiased estimator is given as

$$\begin{aligned} V(\mu_{y_{(F_i)}}) = E_1 V_2(\mu_{y_{(F_i)}}) + V_1 E_2(\mu_{y_{(F_i)}}) = \frac{1}{n(p_1+p_2)^2}\Big[\{p_1 + p_2(1 + \sigma^2_{s_1} + \sigma^2_{s_2}) \\ -(p_1+p_2)^2\}(\sigma^2_y + \mu^2_y) + p_3(1-p_3)F^2 - 2p_3 F(p_1+p_2)\sigma_y\Big] + \frac{\sigma^2_y}{n} \end{aligned} \tag{2.3.3}$$

which is optimal when

$$F_{opt.} = \frac{(p_1+p_2)\sigma_y}{(1-p_3)} \tag{2.3.4}$$

After substituting (2.3.4) in (2.3.3), we get the minimum variance of the proposed unbiased estimator which is given as

$$min.V(\mu_{y_{F_i}}) = \frac{\mu_y^2}{n}\left[\left\{\frac{p_1 + p_2(1+\sigma_{s_1}^2+\sigma_{s_2}^2)}{(p_1+p_2)^2} - 1\right\}(C_y^2+1) - \frac{p_3}{1-p_3}C_y^2 + 1\right] \quad (2.3.5)$$

2.4 Efficiency Comparisons

The effectiveness of the proposed FQORR model is compared with different existing models, including Eichhorn and Hayre's [10] model, Gupta et al. [17, 19, 21] model, Bar-Lev et al. [5] model and Gjestvang and Singh [13] model. Following are the conditions obtained by using the equations (2.2.2), (2.2.4), (2.2.6), (2.2.8), (2.2.10), (2.2.11) and (2.3.5), respectively.

(i) $min.V(\mu_{y_{(F_i)}}) < V(\mu_{y(EH)})$

$$if \quad \{\delta_1 - 1 - C_\gamma^2\}(C_y^2) - \left(\frac{p_3}{1-p_3} + 1\right)C_y^2 + 1 < 0 \quad (2.4.1)$$

(ii) $min.V(\mu_{y_{(F_i)}}) < V(\mu_{y(G_0)})$

$$if \quad \frac{\mu_y^2}{n}\left[\{(\delta_1 - 1)(C_y^2+1)\} - \frac{p_3}{1-p_3}C_y^2 + 1\right] - \frac{1}{(\theta_2-\theta_1)^2}\delta_2 < 0 \quad (2.4.2)$$

(iii) $min.V(\mu_{y_{(F_i)}}) < V(\mu_{y(BBB)})$

$$if \quad \left[\{(\delta_1 - 1)(C_y^2+1)\} - C_y^2\left(\frac{p_3}{1-p_3} + 1 - C_{s(p)}^2\right) + 1 - C_{s(p)}^2\right] < 0 \quad (2.4.3)$$

(iv) $min.V(\mu_{y_{(F_i)}}) < V(\mu_{y(GS)})$

$$if \quad \left[\left\{\delta_2 - \frac{p_3}{1-p_3}\right\}(C_y^2+1) + 2\right] < 0 \quad (2.4.4)$$

(v) $\quad min.V(\mu_{y_{(F_i)}}) < V(\mu_{y(G_i)})$

$$if \quad \frac{\mu_y^2}{n}\left[\{(\delta_1 - 1)(C_y^2 + 1)\} - \frac{p_3}{1-p_3}C_y^2 + 1\right] - \lambda(\sigma_y^2 + W\sigma_s^2) < 0 \qquad (2.4.5)$$

(v) $\quad min.V(\mu_{y_{(F_i)}}) < V(\mu_{y(G_{yw})})$

$$if \quad \left[\left\{\delta_1 - W^2 - \frac{p_3}{1-p_3} - 2\right\}(C_y^2 + 1)\right] < 0 \qquad (2.4.6)$$

Once the aforesaid conditions are met then it is obvious that the suggested forced ORRT estimator $\mu_{y_{(F_i)}}$ is efficient than the existing one. To verify the effectiveness of the aforementioned relationships, we conduct a simulation study with R software, which is detailed in the next section.

2.5 Simulation Study

To acquire a better grasp of the efficiency of the proposed model, we use R software to run a simulation study to test the effectiveness of our proposed model vs Eichhorn and Hayre's [10] model, Gupta et al. [17, 19, 21] model, Bar-Lev et al. [5] model and Gjestvang and Singh [13] model, we have generated a population of $N = 3000$, took a sample of size $n = 500$. The variables $X = rnorm(N, 0, 1)$ and Y, which is connected to X is defined as $Y = X + rnorm(N, 0, 1)$ also generated from normal distribution. The scrambling variable S_1 is also taken from normal distribution with mean 1 and varying variances i.e. $(0.5, 1, 1.5)$ and $S_2 = rnorm(N, 0, 2)$ is also taken from normal distribution and results are averaged over $3,000$ iterations.

To determine the amount of the percent relative efficiency, we calculated the ratio of the variance of existing model(s) i.e. Eichhorn and Hayre's [10], Gupta et al. [17], Bar-Lev et al. [5], Gjestvang and Singh [13], Gupta et al. [21] and Gupta et al. [19] to that of the suggested FQRR model i.e. $\mu_{y_{(F_i)}}$ as

$$RE(\mu_{y_{(j)}}, \mu_{y_{(F_i)}}) = \frac{V(\mu_{y_{(j)}})}{V(\mu_{y_{(F_i)}})} \times 100 \qquad (2.5.1)$$

where $(j) = (EH), (G_0), (BBB), (GS), (G_i), (G_{yw})$.

p_1	p_2	p_3	$RE(\mu_{y_{(EH)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_0)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{(y_{BBB})}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(GS)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_i)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_{yw})}}, \mu_{y_{(F_i)}})$
0.4	0.3	0.3	762.5346	106.6169	1336.9090	1279.2460	165.4933	195.1718
0.5	0.4	0.1	286.4923	40.0570	502.2910	1127.5480	62.1775	73.3281
0.6	0.2	0.2	286.4923	40.0570	502.2910	1127.5480	62.1775	73.3281
0.7	0.2	0.1	742.8762	103.8683	1302.4440	585.4211	161.2268	190.1402
0.8	0.1	0.1	1209.9520	169.1743	2121.3410	336.9696	262.5965	309.6889
0.3	0.3	0.4	923.3165	129.0973	1618.8000	1497.2330	200.3879	236.3242
0.3	0.2	0.5	2792.4520	390.4380	4895.8500	1423.5680	606.0473	714.7319
0.2	0.2	0.6	3315.1630	463.5230	5812.2920	1598.2340	719.4918	848.5208
0.1	0.2	0.7	3258.7950	455.6417	5713.4650	1569.3910	707.2582	834.0933
0.3	0.4	0.3	395.5167	55.3007	693.4374	1503.9290	85.8392	101.2331
0.3	0.5	0.2	199.4007	27.8800	349.5981	1483.0350	43.2760	51.0368
0.2	0.6	0.2	118.4566	16.5624	207.6833	1554.4260	25.7086	30.3191
0.4	0.2	0.4	2127.3610	297.4456	3729.7840	1184.7410	461.7024	544.5011
0.4	0.4	0.2	343.3730	48.0100	602.0168	1319.6490	74.5224	87.8868

Table 2.1: Relative Efficiency of the FQRR model with respect to other existing model(s) when $\sigma^2_{s_1} = 0.5$

p_1	p_2	p_3	$RE(\mu_{y_{(EH)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_0)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{(y_{BBB})}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(GS)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_i)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_{yw})}}, \mu_{y_{(F_i)}})$
0.4	0.3	0.3	541.5128	141.0462	949.4044	908.4551	117.5248	138.6010
0.5	0.4	0.1	205.2948	53.4725	359.9320	807.9793	44.5552	52.5455
0.6	0.2	0.2	869.7640	226.5447	1524.9090	606.4172	188.7654	222.6173
0.7	0.2	0.1	622.9618	162.2609	1092.2040	490.9230	135.2018	159.4480
0.8	0.1	0.1	1099.7000	286.4354	1928.0420	306.2646	238.6685	281.4697
0.3	0.3	0.4	623.1727	162.3159	1092.5740	1010.5250	135.2475	159.5019
0.3	0.2	0.5	1973.4630	514.0217	3459.9630	1006.0540	428.3018	505.1106
0.2	0.2	0.6	2219.6080	578.1344	3891.5160	1070.0690	481.7228	568.1119
0.1	0.2	0.7	2073.7890	540.1533	3635.8580	998.7083	450.0755	530.7891
0.3	0.4	0.3	259.0130	67.4643	454.1131	984.8815	56.2137	66.2947
0.3	0.5	0.2	127.9786	33.3342	224.3778	951.8362	27.7752	32.7563
0.2	0.6	0.2	72.0022	18.7542	126.2376	944.8373	15.6266	18.4290
0.4	0.2	0.4	1581.7640	411.9972	2773.2190	880.8945	343.2912	404.8548
0.4	0.4	0.2	235.7066	61.3937	413.2514	905.8666	51.1555	60.3294

Table 2.2: Relative Efficiency of the FQRR model with respect to other existing model(s) when $\sigma^2_{s_1} = 1$

p_1	p_2	p_3	$RE(\mu_{y_{(EH)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_0)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{(y_{BBB})}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(GS)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_i)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_{yw})}}, \mu_{y_{(F_i)}})$
0.4	0.3	0.3	335.9017	176.2597	588.9178	563.5169	72.9009	85.9745
0.5	0.4	0.1	128.6436	67.5039	225.5437	506.3029	27.9195	32.9265
0.6	0.2	0.2	642.4595	337.1216	1126.3890	447.9358	139.4333	164.4384
0.7	0.2	0.1	481.8819	252.8608	844.8567	379.7454	104.5831	123.3383
0.8	0.1	0.1	953.7748	500.4800	1672.2000	265.6248	206.9983	244.1200
0.3	0.3	0.4	363.0305	190.4952	636.4813	588.6835	78.7887	92.9181
0.3	0.2	0.5	1221.2390	640.8281	2141.1310	622.5771	265.0462	312.5779
0.2	0.2	0.6	1282.4660	672.9559	2248.4760	618.2741	278.3342	328.2489
0.1	0.2	0.7	1121.7370	588.6156	1966.6780	540.2130	243.4511	287.1100
0.3	0.4	0.3	145.4141	76.3040	254.9465	552.9286	31.5593	37.2189
0.3	0.5	0.2	70.1720	36.8218	123.0288	521.9020	15.2294	17.9606
0.2	0.6	0.2	37.1172	19.4767	65.0756	487.0650	8.0555	9.5002
0.4	0.2	0.4	1044.4260	548.0482	1831.1350	581.6477	226.6724	267.3224
0.4	0.4	0.2	140.0395	73.4837	245.5234	538.1990	30.3928	35.8433

Table 2.3: Relative Efficiency of the FQRR model with respect to other existing model(s) when $\sigma^2_{s_1} = 1.5$

Tables 2.1, 2.2 and 2.3 illustrates the values of Relative Efficiency of the suggested model as compared with existing model(s) with different values of variance of scrambling variable S_1 i.e. (0.5,1,1.5). It follows that the relative efficiency by using the suggested forced quantitative ORRT estimator $(\mu_{y_{(F_i)}})$ over Bar-Lev et al. [5] $(\mu_{y_{(BBB)}})$ estimator is larger as compared to the other considerable estimator's i.e. Eichhorn and Hayre [10] $(\mu_{y_{(EH)}})$, Gupta et al. [17, 19, 21] $(\mu_{y_{(G_0)}}, \mu_{y_{(G_i)}}, \mu_{y_{(G_{yw})}})$, Gjestvang and Singh [13] $(\mu_{y_{(GS)}})$ estimator(s). Thus, proposed forced quantitative ORRT estimator $(\mu_{y_{(F_i)}})$ over Bar-Lev et al. [5] $(\mu_{y_{(BBB)}})$ estimator is better than the other competing estimators and our recommendation is to prefer the proposed forced quantitative optional randomized response technique (FQORRT) in practice.

2.6 Conclusion

In this chapter, a revised forced quantitative optional randomised response (FQORR) model has been proposed and explored. The suggested FQORR model that is discovered is found to be more efficient than the other existing models developed by Eichhorn and Hayre's [10], Gupta et al. [17, 19, 21], Bar-Lev et al [5], Gjestvang and Singh [13], respectively and the characteristics up to the first order of approximation are also analysed. To back up the theoretical results, we have conducted a simulation study using R software and based on simulation results, it is clear that the suggested estimator $(\mu_{y_{(F_i)}})$ over Bar-Lev et al. estimator i.e. $(\mu_{y_{(BBB)}})$ is more efficient than the other competing estimator(s). As a result, the proposed forced optional randomized response (FQORR) model may be urged to survey practitioners in real-life problems whenever they intend to cope with stigmatized concerns.

CHAPTER 3

Mean Estimation of Sensitive Variable: A Forced Re-scrambled Optional Randomized Response (FRORR) Approach

3.1 Introduction

Gaining insight into sensitive aspects of life such as drugs consumption, support for terrorism, incidence of domestic violence, criminal activities, under-reported tax and so on, frequently depends on people being honest in their disclosures. In such situations, employing direct interview techniques (asking questions directly to respondents) often results in respondents providing inaccurate responses or opting not to respond, potentially influenced by social disapproval or fear introducing a likelihood of response bias. To address these challenges, Warner [38] initially introduced a method widely recognized as Randomized Response Technique (RRT). The purpose of RRT is to collect information about personal and socially undesirable behaviour. Building upon this legacy, Greenberg et al. [15] and Warner [39] have further advanced this approach. Moreover, Fox and Tracy [12] introduced a method for sensitive surveys by using RRT. Gjestvang et al. [13] developed forced quantitative randomized response model for estimating the mean of sensitive variable. Further, Gjestvang et al. [14],Fox [11], Tarray et al. [34] , Ahmed et al. [1] introduced an innovative work that deals with only qualitative data for the estimation of sensitive variable.

Certain variables possess a quantitative natures including income, expenditures, property, tax dodging etc., to deal with these problems, Eichhorn and Hayre [10] was the first to introduce scrambled RRT. Gupta et al. [17] proposed ORRT model which is based on a very simple concept that a question can be sensitive for one person but not for other. Further, Gupta and Shabbir [16] uses personal interview to address sensitive estimation. Using ORRT approach, various authors such as Gupta et al. [19, 21], Zhang et al. [43], Kumar and Kour [27, 28], Kumar e al. [29] and so on proposed a ORRT model for estimating the mean of sensitive variable.

Building upon prior discussions and work presented by Ahmed et al. [2], the primary goal of this research is to focussed on developing a Forced re-scrambled optional randomized response technique (FRORRT) model for estimating the mean of a sensitive variable that is distinct from other models as it incorporates a fixed factor value chosen by the investigator based on prior experiences and then re-scrambling the already scrambled scores. The rest of the chapter is structured in a way that review of relevant models are discussed in section 3.2. Section 3.3 describes the proposed FQORRT model. In section 3.4 and 3.5, an attempt has been made to the compare the proposed estimator with the existing estimators along with the simulation studies in support of the proposed theoretical results. The concluding remarks are then elaborated in section 3.6.

3.2 Some Existing Models

In the framework of RRT and ORRT, we have examined several existing models with their variances to estimate the population mean μ_y are listed in Table 3.1.

Models	Unbiased Estimator	Variances
Eichhorn and Hayre [10] $Z_{i(EH)} = S_i Y_i$	$\mu_{y_{(EH)}} = \frac{1}{n}\sum_{i=1}^{n} Z_{i(EH)}$	$V(\mu_{y_{(EH)}}) = \frac{1}{n}\left[\sigma_y^2 + C_\gamma^2(\sigma_y^2 + \mu_y^2)\right]$ where $C_\gamma^2 = \sigma_s^2/\mu_s^2$
Gupta et al. [17] $Z_{G_0} = \begin{cases} Y & \text{with probability 1-W} \\ SY & \text{with probability W,} \end{cases}$	$\mu_{y_{(G_0)}} = \bar{Z}_{(G_0)}$	$V(\mu_{y_{(G_0)}}) = \frac{1}{n}\left[\sigma_y^2 + W\frac{\sigma_s^2}{\mu_y^2}(\sigma_y^2 + \mu_y^2)\right]$
Bar-Lev et al. [5] $Z_{i(BBB)} = \begin{cases} Y_i S & \text{with probability (1-p)} \\ Y_i & \text{with probability p,} \end{cases}$	$\mu_{y(BBB)} = \frac{1}{n(1-p)\mu_s + p}\sum_{i=1}^{n} Z_{i(BBB)}$	$V(\mu_{y(BBB)}) = \frac{\mu_y^2}{n}\left[C_y^2 + (1 + C_y^2)C_s^2(p)\right]$ where $C_s^2(p) = \frac{(1-p)\mu_s^2(1+C_\gamma^2)+p}{[(1-p)\mu_s+p]^2} - 1$
Gjestvang and Singh [13] $Z_{i(GS)} = \begin{cases} Y_i & \text{with probability } p_1 \\ Y_i S & \text{with probability } p_2 \\ F & \text{with probability } p_3, \end{cases}$	$\mu_{y(GS)} = \frac{\frac{1}{n}\sum_{i=1}^{n} Z_{i(GS)} - p_3 F}{(p_1 + p_2\mu_s)}$	$V(\mu_{y(GS)}) = \frac{1}{n(p_1+p_2\mu_s)}\left[\{p_1 + p_2(\sigma_s^2 + \mu_s^2) - (p_1 + p_2\mu_s)^2\}(\sigma_y^2 + \mu_y^2) + \right.$ $\left. p_3(1 - p_3)F^2 - 2p_3F(p_1 + p_2\theta)\mu_y\right] + \frac{\sigma_y^2}{n}$ which is optimal when $F = \frac{(p_1+p_2\mu_s)\mu_y}{(1-p_3)}$ $min.V(\mu_{y(GS)}) = \frac{1}{n}\left[\frac{(p_1+p_2\mu_s^2(1+C_\gamma^2))(\sigma_y^2+\mu_y^2)}{(p_1+p_2\mu_s)^2} - 1 - \frac{p_3}{1-p_3}\right]$
Gupta et al. [21] $Z_{(G_i)} = \begin{cases} Y & \text{with probability } T + (1-T)(1-W) \\ YS_i & \text{with probability } (1-T)W, i = 1, 2 \end{cases}$	$\mu_{y_{G_i}} = \frac{\mu_{s_1}\bar{Z}_{G_1} - \mu_{s_2}\bar{Z}_{G_2}}{\mu_{s_1} - \mu_{s_2}}$	$V(\mu_{y_{G_i}}) = \frac{1}{(\mu_{s_1}-\mu_{s_2})^2}\left(\mu_{s_2}^2\frac{\sigma_{z_1}^2}{n_1} + \mu_{s_1}^2\frac{\sigma_{z_2}^2}{n_2}\right)$
Gupta et al. [19] $Z_{(G_{yw})} = Y + ST$	$\mu_{y(G_{yw})} = \frac{1}{n}\sum_{i=1}^{n} Z_{(G_{yw})}$	$V(\mu_{y(G_{yw})}) = \frac{1-f}{n}(\sigma_y^2 + W\sigma_s^2)$
Ahmed et al. [2] $Z_{1(A)} = Y_1S_1Y_2S_2 + F$ $Z_{2(A)} = \begin{cases} Y_1S_1^2 + Y_2S_1S_2 + FS_1 & \text{with probability } W \\ Y_1S_1S_2 + Y_2S_2^2 + FS_2 & \text{with probability } (1-W), \end{cases}$ ($Z_{2(A)}$ is re-scrambled response)	$\mu_{y_{1(A)}} = \frac{\{W\mu_{s_1}\mu_{s_2} + (1-W)(\sigma_{s_2}^2 + \mu_{s_2}^2)\}Z_{1A} - \mu_{s_2}Z_{2A} - (1-W)F\sigma_{s_2}^2}{(1-W)\mu_{s_1}\sigma_{s_2}^2 - W\mu_{s_2}\sigma_{s_2}^2}$ $\mu_{y_{2(A)}} = \frac{\mu_{s_1}Z_{2A} - Z_{1A}\{W(\sigma_{s_2}^2 + \mu_{s_2}^2) + (1-W)\mu_{s_1}\mu_{s_2}\}WF\sigma_{s_2}^2}{(1-W)\mu_{s_1}\sigma_{s_2}^2 - W\mu_{s_2}\sigma_{s_1}^2}$	$V(\mu_{y_{1(A)}}) = \frac{\{W\mu_{s_1}\mu_{s_2}+(1-W)(\sigma_{s_2}^2+\mu_{s_2}^2)\}^2\sigma_{z_1}^2 + \mu_{s_2}^2\sigma_{z_2}^2 - 2\mu_{s_2}\{W\mu_{s_1}\mu_{s_2}+(1-W)(\sigma_{s_2}^2+\mu_{s_2}^2)\}\sigma_{z_1z_2}}{\{(1-W)\mu_{s_1}\sigma_{s_2}^2 - W\mu_{s_2}\sigma_{s_1}^2\}^2}$ which is optimal when $F = \frac{\mu_{s_2}T_2 - \{W\mu_{s_1}\mu_{s_2}+(1-W)(\sigma_{s_2}^2+\mu_{s_2}^2)\}\{W\mu_{y_1}\sigma_{s_1}^2+(1-W)\mu_{y_2}\sigma_{s_2}^2\}^2}{T_1\mu_{s_2}}$ $min.V(\mu_{y_{1(A)}}) = V(\mu_{y_{1(A)}}) + \frac{\mu_{s_2}T_2 - \{W\mu_{s_1}\mu_{s_2}+(1-W)(\sigma_{s_2}^2+\mu_{s_2}^2)\}\{W\mu_{y_1}\sigma_{s_1}^2+(1-W)\mu_{y_2}\sigma_{s_2}^2\}^2}{nT_1\{(1-W)\mu_{s_1}\sigma_{s_2}^2 - W\mu_{s_2}\sigma_{s_1}^2\}^2\}^2}$ $V(\mu_{y_{2(A)}}) = \frac{\{W(\sigma_{s_2}^2+\mu_{s_2}^2)+(1-W)\mu_{s_1}\mu_{s_2}\}^2\sigma_{z_1}^2 + \mu_{s_1}^2\sigma_{z_2}^2 - 2\mu_{s_1}\{W(\sigma_{s_2}^2+\mu_{s_2}^2)+(1-W)\mu_{s_1}\mu_{s_2}\}\sigma_{z_1z_2}}{n\{(1-W)\mu_{s_1}\sigma_{s_2}^2 - W\mu_{s_2}\sigma_{s_1}^2\}^2}$ which is optimal when $F = \frac{\mu_{s_1}T_2 - \{W(\sigma_{s_2}^2+\mu_{s_2}^2)+(1-W)\mu_{s_1}\mu_{s_2}\}\{W\mu_{y_1}\sigma_{s_1}^2+(1-W)\mu_{y_2}\sigma_{s_2}^2\}^2}{T_1\mu_{s_1}}$ $min.V(\mu_{y_{2(A)}}) = V(\mu_{y_{2(A)}}) + \frac{\left[\mu_{s_1}T_2 - \{W(\sigma_{s_2}^2+\mu_{s_2}^2)+(1-W)\mu_{s_1}\mu_{s_2}\}\{W\mu_{y_1}\sigma_{s_1}^2+(1-W)\mu_{y_2}\sigma_{s_2}^2\}^2\right]^2}{nT_1\{(1-W)\mu_{s_1}\sigma_{s_2}^2 - W\mu_{s_2}\sigma_{s_1}^2\}^2\}^2}$

where $T_1 = W(\sigma_{s_1}^2 + \mu_{s_1}^2) + (1 - W)(\sigma_{s_2}^2 + \mu_{s_2}^2) - (W\mu_{s_1} + (1 - W)\mu_{s_2})^2$ and

$T_2 = W\{(\sigma_{s_1}^3 + 3\mu_{s_1}\sigma_{s_1}^2 + \mu_{s_1}^3)\mu_{y_1} + (\sigma_{s_1}^2 + \mu_{s_1}^2)\mu_{s_2}\mu_{y_2}\} + (1 - W)\{(\sigma_{s_2}^3 + 3\mu_{s_2}\sigma_{s_2}^2 + \mu_{s_2}^3)\mu_{y_2} + (\sigma_{s_2}^2 + \mu_{s_2}^2)\mu_{s_1}\mu_{y_1}\} - (W\mu_{s_1} + (1 - W)\mu_{s_2})\{W((\sigma_{s_1}^2 + \mu_{s_1}^2)\mu_{y_1} + \mu_{s_1}\mu_{s_2}\mu_{y_2}) + (1 - W)(\mu_{s_1}\mu_{s_2}\mu_{y1} + (\sigma_{s_2}^2 + \mu_{s_2}^2)mu_{y_2})$.

Table 3.1: Several Existing Models in the context of RRT and ORRT

3.3 Forced Re-scrambled ORRT Model

Recently, Ahmed et al. [2] have proposed a method for concurrently estimating the means of two sensitive variables by collecting one scrambled response and other new re-scrambled response by making use of the forced quantitative randomized response model of Gjestvang and Singh [13]. Driven by Ahmed et al. [2], we apply the forced re-scrambled optional randomized response model (FRORRT) to expand their concept to the mean estimation of sensitive variable and also re-scrambled the scrambled scores by using FRORRT. Suppose that there is one quantitative sensitive variables of interest Y in a finite population of size N units. Let $\bar{y}$ be the corresponding population mean, which we wish to estimate. Assume we select a simple random sample without replacement (SRSWOR) of size n from the population. Each respondent selected in the sample is asked to generate two scrambling variables S_1 and S_2 and the distributions of the scrambling variables S_1 and S_2 are assumed to be known. Let $E(S_1) = 1$, $V(S_1) = \sigma_{s_1}^2$ and $E(S_2) = 0$, $V(S_2) = \sigma_{s_2}^2$. Now, each respondent is asked to report a scrambled response given by

$$Z_1 = S_1 Y + S_2 \tag{3.3.1}$$

All respondents are also requested to draw a card from a deck consisting of two types of cards, similar to the Warner [38] device, but that has different types of outcomes. In the deck, let W be the proportion of cards bearing the statement, “Please report the value of the scrambling variable S_1 which you used in scrambling the response on two sensitive variables”. Let $(1 - W)$ be the proportion of cards bearing the statement,“Please report the value of the scrambling variable S_2 which you used in scrambling the response on two sensitive variables”. Thus the second response from the i^{th} respondent is given by

$$Z_2 = \begin{cases} S_1 & \text{with probability } W \\ S_2 & \text{with probability } 1 - W \end{cases} \tag{3.3.2}$$

Our approach distinguishes itself from the existing randomized response techniques found in the literature. Specifically, unlike existing methods, we introduce a novel step where, from the initial scrambled response in (3.3.1), the interviewer generates a re-scrambled response $Z_1^* = Z_1 + F$, where F is predetermined and chosen by the investigator based on past experience. In practical terms, we begin with equivalent information from a respondent as in the Ahmed et al. [1] model. However, during the estimation stage, the investigator introduces a re-scrambling process to the initial observed response by incorporating a constant F. It's worth noting that respondents can alternatively be instructed to add the constant before providing their response, with the same effect. The resulting modified first response, denoted as Z_1^*, is expressed as follows

$$Z_1^* = S_1 Y + S_2 + F \tag{3.3.3}$$

Taking expected value on both sides of (3), we get

$E(Z_1^*) = E(S_1 Y + S_2 + F) = Y + F$

From Z_1^* and Z_2, we create an additional re-scrambled response Z_2^* given as

$$Z_i^* = Z_1 Z_2 = \begin{cases} S_1^2 Y + S_1 S_2 + S_1 F & \text{with probability } W \\ S_1 S_2 Y + S_2^2 + S_2 F & \text{with probability } 1 - W \end{cases} \tag{3.3.4}$$

Now, Taking the expected value of Z_2^*, we get

$E(Z_2^*) = W\sigma_{s_1}^2 Y + WY + WF + \sigma_{s_2}^2 - W\sigma_{s_2}^2$

Then, we have

$E(S_1^3) = \sigma_{s_1}^3 + 3\sigma_{s_1}^2 + 1$, $E(S_2^3) = \sigma_{s_2}^3$, $E(S_1^4) = \sigma_{s_1}^4 + 4\sigma_{s_1}^3 + 6\sigma_{s_1}^2 + 1$, $E(S_2^4) = \sigma_{s_2}^4$;

$E(S_1^2S_2^2) = (\sigma_{s_1}^2 + 1)$, $E(S_1S_2^3) = \sigma_{s_2}^3$, $E(S_1^3S_2) = 0$

The proposed unbiased estimator of the population mean μ_y is given as

$$\mu_{y_{(RF)}} = \frac{\frac{1}{n}\sum_{i=1}^{n} Z_i^* - WF - \sigma_{s_2}^2(1-W)}{W(\sigma_{s_1}^2 + 1)} \tag{3.3.5}$$

The variance of the proposed unbiased estimator is given as

$$V(\mu_{y_{(RF)}}) = E_1V_2(\mu_{y_{(RF)}}) + V_1E_2(\mu_{y_{(RF)}}) = \frac{1}{n\left[W^2(\sigma_{s_1}^2+1)^2\right]}\Big[(WA_1 - W^2A_1 - WA_2 + W^2A_2)(\sigma_y^2 + \mu_y^2) + (WA_3 - W^2A_3 - W\sigma_{s_2}^2 - W^2\sigma_{s_2}^2)F^2 + W(A_2 - WA_2 - A_1 + WA_1) + 2FA_4\sigma_y^3 + (1-W)^2\sigma_{s_2}^3 + 2\sigma_y\sigma_{s_2}^3\Big] + \frac{\sigma_y^2}{n} \tag{3.3.6}$$

which is optimal when

$$F_{opt.} = \frac{W^2A_4\sigma_y - (1-W)^2\sigma_{s_2}^3}{2(WA_3 - W^2A_3 + W\sigma_{s_2}^2 - W^2\sigma_{s_2}^2)} \tag{3.3.7}$$

After substituting (3.3.7) in (3.3.6), we get the minimum variance of the proposed unbiased estimator which is given as

$$min.V(\mu_{y_{(RF)}}) = \frac{1}{n\left[W^2(\sigma_{s_1}^2+1)^2\right]}\Big[(WA_1 - W^2A_1 - WA_2 + W^2A_2)(\sigma_y^2 + \mu_y^2) + \frac{W^2A_4\sigma_y - (1-W)^2\sigma_{s_2}^3}{(WA_3 - W^2A_3 - W\sigma_{s_2}^2 - W^2\sigma_{s_2}^2)}(1 + 2\sigma_y) + 2\sigma_y\sigma_{s_2}^3 + \sigma_y^2\left[W^2(\sigma_{s_1}^2+1)^2\right]\Big] \tag{3.3.8}$$

3.4 Efficiency Comparisons

The performance of the proposed FRORR model is evaluated by comparing it with various existing models i.e., models by Eichhorn and Hayre [10], Gupta et al. [17, 19, 21], Bar-Lev et al. [5], Gjestvang and Singh [13] and Ahmed et al. [2]. The conditions derived in Table 1 and (8) are presented below for each respective model.

(i) $\quad min.V(\mu_{y_{(F_i)}}) < V(\mu_{y(EH)})$

$$if \quad \frac{1}{W^2A_3^2}\left[(\delta_1 - C_\gamma^2)(\sigma_y^2 + \mu_y^2) + \delta_2(1+2\sigma_y) + 2\sigma_y\sigma_{s_2}^2 + \sigma_y^2(W^2A_3^2+1)\right] < 0 \tag{3.4.1}$$

(ii) $\quad min.V(\mu_{y_{(RF)}}) < V(\mu_{y(G_0)})$

$$if \quad \left[\left(\frac{\delta_1}{W^2A_3^2} - W\frac{\sigma_s^2}{\mu_y^2}\right)(\sigma_y^2 + \mu_y^2) + \frac{\delta_2}{W^2A_3^2}(1+2\sigma_y) + 2\sigma_y\sigma_{s_2}^2 + \sigma_y^2(W^2A_3^2+1)\right] < 0 \tag{3.4.2}$$

(iii) $\quad min.V(\mu_{y_{(RF)}}) < V(\mu_{y(BBB)})$

$$if \quad \left[\left(\frac{\delta_1}{W^2A_3^2} - -C_s^2(p)\right)(\sigma_y^2 + \mu_y^2) + \frac{\delta_2}{W^2A_3^2}(1+2\sigma_y) + 2\sigma_y\sigma_{s_2}^2 + \sigma_y^2(W^2A_3^2+1)\right] < 0 \tag{3.4.3}$$

(iv) $\quad min.V(\mu_{y_{(RF)}}) < V(\mu_{y(GS)})$

$$if \quad \left[\left(\frac{\delta_1}{W^2A_3^2} - \left\{\frac{p_1 + p_2\mu_s^2(1+C_\gamma^2)}{(p_1+p_2\mu_s)^2}\right\}\right)(\sigma_y^2 + \mu_y^2) + \frac{\delta_2}{W^2A_3^2}(1+2\sigma_y) + 2\sigma_y\sigma_{s_2}^2 + \frac{p_3}{1-p_3} + \sigma_y^2(W^2A_3^2+1)\right] < 0 \tag{3.4.4}$$

(v) $\quad min.V(\mu_{y_{(RF)}}) < V(\mu_{y(G_i)})$

$$if \quad \left[\left(\frac{\delta_1}{W^2A_3^2}(\sigma_y^2 + \mu_y^2) + \frac{\delta_2}{W^2A_3^2}(1+2\sigma_y) + 2\sigma_y\sigma_{s_2}^2 + \sigma_y^2\right) - \frac{1}{(\mu_{s_1} - \mu_{s_1})^2}\delta_3\right] < 0 \tag{3.4.5}$$

(vi) $\quad min.V(\mu_{y_{(RF)}}) < V(\mu_{y(G_{yw})})$

$$if \quad \left[\left(\frac{\delta_1}{W^2A_3^2}(\sigma_y^2+\mu_y^2)+\frac{\delta_2}{W^2A_3^2}(1+2\sigma_y)+2\sigma_y\sigma_{s_2}^2+\sigma_y^2\right)-\left(1-\frac{n}{N}\right)\right.$$
$$\left.(\sigma_y^2-W\sigma_s^2)\right]<0 \quad (3.4.6)$$

(vii) $\quad min.V(\mu_{y_{(RF)}}) < min.V(\mu_{y(1(A))})$

$$if \quad \left[\left(\frac{\delta_1}{W^2A_3^2}(\sigma_y^2+\mu_y^2)+\frac{\delta_2}{W^2A_3^2}(1+2\sigma_y)+2\sigma_y\sigma_{s_2}^2+\sigma_y^2\right)-V(\mu_{1(A)})-\frac{\delta_4}{T_1\delta_5}\right]$$
$$<0$$
$$(3.4.7)$$

(viii) $\quad min.V(\mu_{y_{(RF)}}) < min.V(\mu_{y(2(A))})$

$$if \quad \left[\left(\frac{\delta_1}{W^2A_3^2}(\sigma_y^2+\mu_y^2)+\frac{\delta_2}{W^2A_3^2}(1+2\sigma_y)+2\sigma_y\sigma_{s_2}^2+\sigma_y^2\right)-V(\mu_{2(A)})-\frac{\delta_6}{T_1\delta_5}\right]$$
$$<0$$
$$(3.4.8)$$

where

$\delta_1 = WA_1 - W^2A_1 + WA_2 + W^2A_2,$

$\delta_2 = \frac{W^2A_4\sigma_y-(1-W)^2\sigma_{s_2}^3}{(WA_3-W^2A_3-W\sigma_{s_2}^2-W^2\sigma_{s_2}^2)},$

$\delta_3 = \left(\mu_{s_2}^2\frac{\sigma_{z_1}^2}{n_1}+\mu_{s_1}^2\frac{\sigma_{z_2}^2}{n_2}\right),$

$\delta_4 = \mu_{s_2}T_2 - \{W\mu_{s_1}\mu_{s_2}+(1-W)(\sigma_{s_2}^2+\mu_{s_2}^2)\}\{W\mu_{y_1}\sigma_{s_1}^2+(1-W)\mu_{y_2}\sigma_{s_2}^2\}^2,$

$\delta_5 = \{(1-W)\mu_{s_1}\sigma_{s_2}^2 - W\mu_{s_2}\sigma_{s_1}^2\}^2,$

$\delta_6 = \mu_{s_1}T_2 - \{W(\sigma_{s_2}^2+\mu_{s_2}^2)+(1-W)\mu_{s_1}\mu_{s_2}\}\{W\mu_{y_1}\sigma_{s_1}^2+(1-W)\mu_{y_2}\sigma_{s_2}^2\}^2.$

When the aforementioned conditions are satisfied then it is clear that the proposed forced re-scrambled ORRT estimator $mu_{y_{RF}}$ is more efficient than the existing one. To validate the effectiveness of the above relationships, we carried out a simulation study using R software, as described in the next section.

3.5 Simulation Study Using R Software

The use of simulation studies is to validate results has been established. The properties of the proposed FRORRT model have been investigated for various sample sizes and also test the effectiveness of our proposed model over existing ones. We have generated a population of $N = 3000$ with varying sample sizes. The variables $X = rnorm(N, 0, 1)$ and Y, which is connected to X is defined as $Y = X + rnorm(N, 0, 1)$ also generated from normal distribution. The scrambling variable $S_1 = rnorm(N, 1, 0.5)$ and $S_2 = rnorm(N, 0, 1)$ is also taken from normal distribution. The proportions (p_1, p_2, p_3) are fixed i.e., $p_1 = 0.2$, $p_2 = 0.3$ and $p_3 = 0.5$ such that $p_1 = p_2 = p_3 = 1$.

To determine the amount of the percent relative efficiency, we calculated the ratio of the variance of existing model(s) i.e. Eichhorn and Hayre [10], Gupta et al. [17], Bar-Lev et al. [5], Gjestvang and Singh [13], Gupta et al. [21], Gupta et al. [19] and Ahmed et al. [2] to that of the suggested FRORR model i.e. $\mu_{y_{RF}}$ as

$$RE(\mu_{y_j}, \mu_{y_{RF}}) = \frac{V(\mu_{y_j})}{V(\mu_{y_{RF}})} \times 100 \tag{3.5.1}$$

where $j = \big[(EH), (G_0), (BBB), (GS), (G_i), (G_{yw}), 1(A), 2(A)\big]$.

W	$RE(\mu_{y_{(EH)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_0)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{(y_{BBB})}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(GS)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_i)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_{yw})}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{1(A)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{2(A)}}, \mu_{y_{(F_i)}})$
				F=0.2				
0.1	1327.4059	314.0923	1103.4895	393.6199	98.4114	42.0383	153.4019	197.0759
0.2	1244.5467	484.7345	876.3364	369.0494	91.3820	43.6256	145.0711	202.4839
0.3	936.9299	477.1954	557.6821	277.8308	68.1345	36.2340	110.7019	175.0679
0.4	539.8316	327.1314	271.2355	160.0780	38.8805	23.1533	65.2149	123.8414
0.5	238.8176	163.9246	100.9776	70.81735	17.0356	11.3442	29.9851	73.8162
				F=0.4				
0.1	1306.3551	309.1113	1085.9897	387.3614	96.8507	41.3716	150.9485	193.4999
0.2	1214.3752	472.9831	855.0914	360.0874	89.1666	42.5679	141.5336	197.1286
0.3	904.8667	460.8650	538.5974	268.3117	65.8029	34.9941	106.8970	168.7177
0.4	516.6197	313.0652	259.5727	153.1884	37.2087	22.1577	62.4004	118.2913
0.5	228.0659	156.5447	96.4316	67.6262	16.2686	10.8334	28.6300	70.3812
				F=0.6				
0.1	1264.6893	299.2523	1051.3524	375.3098	93.7617	40.0521	146.1173	186.9640
0.2	1167.6799	454.7959	822.2114	346.5212	85.7380	40.9311	136.0747	189.1860
0.3	862.5356	439.3050	513.4009	255.9665	62.7245	33.3570	101.8827	160.5322
0.4	488.8204	296.2192	245.6052	145.0626	35.2065	20.9654	59.03417	111.7415
0.5	215.8026	148.1271	91.2463	64.0416	15.3939	10.2509	27.0863	66.5038

Table 3.2: Relative Efficiency of the FQRR model with respect to other existing model(s) when $n = 650$

W	$RE(\mu_{y_{(EH)}},\mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_0)}},\mu_{y_{(F_i)}})$	$RE(\mu_{(y_{BBB})},\mu_{y_{(F_i)}})$	$RE(\mu_{y_{(GS)}},\mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_i)}},\mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_{yw})}},\mu_{y_{(F_i)}})$	$RE(\mu_{y_{1(A)}},\mu_{y_{(F_i)}})$	$RE(\mu_{y_{2(A)}},\mu_{y_{(F_i)}})$
				F=0.2				
0.1	1145.9503	171.6243	976.5035	359.7750	78.3251	39.8975	139.4900	178.9593
0.2	1050.5009	275.4779	757.7887	329.8084	71.0949	39.6585	129.1328	181.1427
0.3	763.4961	289.6565	465.1755	239.7022	51.1630	30.6434	95.3168	152.8919
0.4	421.2022	211.8892	216.3606	132.2379	27.9479	17.5077	53.9392	105.1083
0.5	178.9069	113.6493	77.1969	56.1684	11.7542	8.0729	23.9468	61.2395
				F=0.4				
0.1	1128.8284	169.0600	961.9134	354.9071	77.1548	39.3014	137.4001	176.1651
0.2	1024.0953	268.5534	738.7407	321.9787	69.3078	38.6616	125.8804	176.4532
0.3	735.2022	278.9223	447.9369	231.1498	49.2670	29.5078	91.7787	147.1039
0.4	401.0684	201.7607	206.0184	126.0971	26.6119	16.6708	51.3568	100.0005
0.5	169.7094	107.8067	73.2282	53.3571	11.1500	7.6578	22.7135	58.0464
				F=0.6				
0.1	1095.010	163.9952	933.0962	345.0592	74.8434	38.1240	133.2817	170.8424
0.2	984.9321	258.2835	710.4900	310.3713	66.6574	37.1831	121.0634	169.6392
0.3	699.6251	265.4250	426.2609	220.4655	46.8829	28.0799	87.3341	139.9158
0.4	378.0862	190.1993	194.2130	119.1423	25.0870	15.7155	48.4114	94.2170
0.5	159.7050	101.4514	68.9114	50.3261	10.4927	7.2064	21.3731	54.5934

Table 3.3: Relative Efficiency of the FQRR model with respect to other existing model(s) when $n = 700$

W	$RE(\mu_{y_{(EH)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_0)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{(y_{BBB})}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(GS)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_i)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{(G_{yw})}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{1(A)}}, \mu_{y_{(F_i)}})$	$RE(\mu_{y_{2(A)}}, \mu_{y_{(F_i)}})$
				F=0.2				
0.1	1320.6529	946.7111	1056.6353	388.5980	89.6534	40.0348	141.6644	195.2950
0.2	1216.3969	2087.2438	823.7572	357.9210	81.7725	40.0102	131.6755	196.6427
0.3	893.7486	2896.1367	511.0112	262.9827	59.4983	32.5326	98.1597	165.7100
0.4	498.8604	2820.8704	240.2975	146.7881	32.8873	19.8204	56.1180	113.6039
0.5	212.3589	2051.0250	85.8231	62.4859	13.8639	9.14581	24.9110	65.3375
				F=0.4				
0.1	1317.5662	944.4984	1054.1656	388.1072	89.4438	39.9412	141.3377	194.9315
0.2	1198.4110	2056.3814	811.5770	353.0085	80.5634	39.4186	129.7315	193.7972
0.3	868.1037	2813.0360	496.3485	255.7119	57.7910	31.5991	95.3443	160.9781
0.4	478.4089	2705.2247	230.4462	140.9219	31.5391	19.0078	53.8172	108.9419
0.5	202.6959	1957.6971	81.9179	59.7068	13.2330	8.7296	23.7769	62.3513
				F=0.6				
0.1	1290.7599	925.2823	1032.7183	380.9504	87.6241	39.12868	138.4700	191.1303
0.2	1161.4035	1992.8793	786.5151	342.7726	78.0755	38.201384	125.7314	187.9394
0.3	830.8653	2692.3674	475.0570	245.2187	55.3120	30.24364	91.2577	154.1426
0.4	453.0425	2561.7869	218.2273	133.7093	29.8668	18.000045	50.9647	103.1876
0.5	191.5483	1850.0299	77.4127	56.5329	12.5053	8.2495478	22.4692	58.9212

Table 3.4: Relative Efficiency of the FQRR model with respect to other existing model(s) when $n = 800$

The discussions of simulation results presented in Tables 3.2-3.4 is as follows The values of Relative Efficiency of the suggested model as compared with existing model(s) with different values of sample size i.e., $n = 650, 700, 800$ are presented in Tables 3.2-3.4.

It is clear from Table 3.2 and 3.3 that the relative efficiency by using the suggested forced re-scrambled ORRT estimator ($\mu_{y_{RF}}$) over Eichhorn and Hayre [10] ($\mu_{y_{(EH)}}$) estimator is larger as compared to the other considerable estimator's i.e. Gupta et al. [17, 19, 21] ($\mu_{y_{(G_0)}}, \mu_{y_{(G_i)}}, \mu_{y_{(G_{yw})}}$), Bar-Lev et al. [5] ($\mu_{y_{(BBB)}}$), Gjestvang and Singh [13] ($\mu_{y_{(GS)}}$) and Ahmed et al. [2] ($\mu_{y_{1(A)}}$, $\mu_{y_{2(A)}}$) estimator(s). Thus, proposed FRORRT estimator ($\mu_{y_{RF}}$) over Eichhorn and Hayre [10] ($\mu_{y_{(EH)}}$) estimator is better than the other competing estimators.

From Table 3.4, it is obvious that the relative efficiency by using the suggested FRORRT estimator ($\mu_{y_{RF}}$) over Bar-Lev et al. [5] ($\mu_{y_{(BBB)}}$) estimator is larger and according to Table 4. the proposed FRORRT estimator ($\mu_{y_{RF}}$) over Bar-Lev et al. [5] ($\mu_{y_{(BBB)}}$) estimator is better than the other competing estimators. Thus, our recommendation is to prefer the proposed forced re-scrambled optional randomized response technique (FQORRT) in practice.

3.6 Conclusion

In this chapter, a novel Forced re-scrambled ORR model has been proposed and explored. The discovered FRORR model is determined to be more efficient than other existing models. The properties up to the first order of approximation are also studied. A simulation study using R software is also carried out to support up the theoretical results and from the simulation results, it is clear that the suggested model ($\mu_{y_{RF}}$) is more efficient than the other competing models. As a result, the proposed FRORRT approach may be forced to survey practitioners in real-world difficulties anytime they want to deal with stigmatized problems.

Bibliography

[1] Ahmed, S., Sedory, S. A. and Singh, S., *Simultaneous estimation of means of two sensitive variables*, Communications in Statistics–Theory and Methods **47** (2018), no. 2, 324–343.

[2] Ahmed, S., Sedory, S. and Singh, S., *Forcibly re-scrambling randomized response model for simultaneous estimation of means of two sensitive attributes*, Communications in Mathematics and Statistics **8** (2020), 23–45.

[3] Arnab, R., *Optional randomized response techniques for complex survey designs*, Biometrical Journal **46** (2004), 114–124.

[4] Azeem, M., *Using the exponential function of scrambling variable in quantitative randomized response models*, Mathematical Methods in the Applied Sciences (2023), 1–12, DOI 10.1002/mma.9295.

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

[6] Chaudhuri, A., *Randomized response and indirect questioning techniques in surveys*, Chapman and Hall–Boca Raton (2011).

[7] Chaudhuri, A and Christofides, T. C., *Indirect questioning in sample surveys*, Springer–Berlin (2013).

[8] Chaudhuri, A., Christofides, T. C. and Rao, C. R., *Data gathering, analysis and protection of privacy through randomized response techniques–qualitative and quantitative human traits*, Handbook of Statistics, Elsevier: North Holland (2016).

[9] Diana, G. and Perri, P. F., *A class of estimators for quantitative sensitive data, statistical papers*, Statistical Papers **52** (2011), 633–650.

[10] Eichhorn, B. H. and Hayre, L. S., *Scrambled randomized response methods for obtaining sensitive quantitative data*, Journal of Statistical Planning and Inference **7** (1983), no. 4, 307–316.

[11] Fox, J. A., *Randomized response and related methods*, SAGE, Los Angeles (ISBN 978-1-4833-8103-9.) (2016).

[12] Fox, J. A. and Tracy, P. E., *Randomized response: A method for sensitive surveys*, SAGE Publications (1986).

[13] Gjestvang, C. R. and Singh, S., *Forced quantitative randomized response model: A new device*, Metrika **66** (2007), 243–257.

[14] Gjestvang, C. R. and Singh, S., *An improved randomized response model: estimation of mean*, Journal of Applied Statistics **36** (2009), 1361–1367.

[15] Greenberg, B. G., Abdel-Latif A. A., Walt R. S. and Daniel G. H., *The unrelated question randomized response model: Theoretical framework*, Journal of the American Statistical Association **64** (1969), 520–539.

[16] Gupta, S. and Shabbir, J., *Sensitivity estimation for personal interview survey questions*, Statistica **64** (2004), 643–653.

[17] Gupta, S., Gupta, B. and Singh, S., *Estimation of sensitivity level of personal interview survey questions*, Journal of Statistical Planning and inference **100** (2002), 239–247.

[18] Gupta, S., Kalucha, G. and Shabbir, J., *A regression estimator for finite population mean of a sensitive variable using an optional randomized response model*, Communications in Statistics–Simulation and Computation **46** (2015), no. 3, 2393–2405.

[19] Gupta, S., Kalucha, J., Shabbir, J. and Dass, B. K., *Estimation of finite population mean using optional rrt models in the presence of non-sensitive auxiliary information*, American Journal of Mathematical and Management Sciences **33** (2014), no. 2, 147–159.

[20] Gupta, S., Mehta, S., Shabbir, J. and Khalil, S., *A unified measure of respondent privacy and model efficiency in quantitative rrt models*, Journal of Statistical Theory and Practice **12** (2018), no. 3, 506–511.

[21] Gupta, S., Shabbir, J. and Sehra, S. J., *Mean and sensitivity estimation in optional randomized response models*, Journal of Statistical Planning and Inference **140** (2010), no. 10, 2870–2874.

[22] Gupta, S., Shabbir. J., Sousa, R. and Corte-Real, P., *Improved exponential type estimators of the mean of a sensitive variable in the presence of nonsensitve auxiliary information*, Communications in Statistics–Simulation and Computation **45** (2016), no. 9, 3317–3328.

[23] Gupta, S., Shabbir, J., Sousa, R. and Corte-Real, P., *Estimation of the mean of a sensitive variable in the presence of auxiliary information*, Communications in Statistics–Theory and Methods **41** (2012), 13–14.

[24] Kalucha G., Gupta S. and Dass B. K., *Ratio estimation of finite population mean using optional randomized response models*, Journal of Statistical Theory and Practice **9**, no. 3.

[25] Khalil, S., Gupta, S. and Hanif, M., *Estimation of finite population mean in stratified sampling using scrambled responses in the presence of measurement errors*, Communications in Statistics–Theory and Methods **48** (2019), no. 6, 1553–1561.

[26] Khalil, S., Zhang, Q. and Gupta, S., *Mean estimation of sensitive variables under measurement errors using optional rrt models*, Communications in Statistics–Simulation and Computation **50** (2021), no. 5, 1417–1426.

[27] Kumar, S. and Kour, S. P., *Estimation of sensitive variable in two-phase sampling under measurement error and non-response using orrt models*, Sri Lankan Journal of Applied Statistics **22** (2021), 95—122.

[28] Kumar, S. and Kour, S. P., *The joint influence of estimation of sensitive variable under measurement error and non-response using orrt models*, Journal of Statistical Computation and Simulation **92** (2022), no. 17, 3583–3604.

[29] Kumar, S., Kour, S. P. and Zhang, Q., *An enhanced ratio-cum-product estimator with non-response and observational error by utilizing orrt models: a sensitive estimation approach*, Journal of Statistical Computation and Simulation **93** (2023), no. 5, 818–836.

[30] Mushtaq, N., Noor-Ul-Amin, M. and Hanif, M., *A family of estimators of a sensitive variable using auxiliary information in stratified random sampling*, Pakistan Journal of Operation Research **13** (2017), no. 1, 141–155.

[31] Pollock, K. and Bek, Y., *A comparison of three randomized response models for quantitative data*, Journal of the American Statistical Association **71** (1976), no. 356, 884–886.

[32] Singh, H. P. and Gorey, S. M., *An alternative to odumade and singh's generalized forced quantitative randomized response mode: a unified approach*, Model Assisted Statistics and Applications **12** (2017), no. 2.

[33] Sousa , R., Shabbir , J., Real , P. C. and Gupta , S., *Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information*, Journal of Statistical Theory and Practice **4** (2010), no. 3, 495–507.

[34] Tarray, T. A., Singh, H. P. and Yan, Z., *A dexterous optional randomized response model*, Sociological Methods and Research **46** (2017), 565–585.

[35] Tiwari, K. K., Bhougal, S., Kumar, S. and Rather, K. U. I., *Using randomized response to estimate the population mean of a sensitive variable under the influence of measurement error*, Journal of Statistical Theory and Practice **16** (2022).

[36] Tiwari, N. and Pandey, T. K., *An improved two-stage forced randomized response model for estimating the proportion of sensitive attribute*, Journal of the Indian Society for Probability and Statistics **23** (2022), 451–464.

[37] Tracy, D. S. and Mangat, N. S., *Some developments in randomized response sampling during the last decade–a follow up of review by chaudhuri and mukerjee*, Journal of Applied Statistical Science **4** (1996), 147–158.

[38] Warner, S. L., *Randomized response: a survey technique for eliminating evasive answer bias*, Journal of the American Statistical Association **60** (1965), no. 309, 63–69.

[39] Warner, S. L., *The linear randomized response model*, Journal of the American Statistical Association **66** (1971), no. 336, 884–888.

[40] Zapata, Z., Sedory, A. S. and Singh, S., *An innovative improvement in warner's randomized response device for evasive answer bias*, Journal of Statistical Computation and Simulation **93** (2022), no. 2.

[41] Zhang, Q., Kalucha, G., Gupta, S. and Khalil, S., *Ratio estimation of the mean under rrt models*, Journal of Statistics and Managements Systems **22** (2018), no. 1, 97–113.

[42] Zhang, Q., Khalil, S. and Gupta, S., *Mean estimation in the simultaneous presence of measurement errors and non-response using optional rrt models under stratified sampling*, Journal of Statistical Computation and Simulation **91** (2021), no. 17, 3492–3504.

[43] Zhang, Q., Khalil, S. and Gupta, S., *Mean estimation of sensitive variables under non-response and measurement errors using optional rrt models*, Journal of Statistical Theory and Practice **15** (2021), no. 3, 1–15.