# A Stratified Randomized Response Technique for Addressing Sensitive Questions in Survey 

Olayiwola O. M, Apantaku F.S and Sohe E.M<br>${ }^{1}$ Department of Statistics, College of Physical Sciences, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria.


#### Abstract

A Stratified Randomized Response Technique (SRRT) with three randomization devices which focused on classification of respondents in a heterogeneous population into non-overlapping homogenous was developed. A simple random sampling scheme was used to select 500 students among students of Federal University of Agriculture, Abeokuta, Ogun State. These students were stratified into two homogenous strata and proportional allocation procedure was used to select 199 male students and 101 female students. An interviewer administered questionnaire was used to collect data on related sensitive questions such as sexually transmitted disease, rape, abortion, cultism,murder, stealing, examination malpractice and bribery of lecturers for marks from selected students. An estimator was proposed for estimating the proportion of students with stigmatized and sensitive behaviour and its statistical properties were examined. The proportion of male students who had contracted sexually transmitted disease, involved in rape, abortion, cultism,murder, stealing and examination malpractice is higher than that for the females. The proposed estimator showed that the variance of proportion for both male and female students involved in the sensitive and stigmatized characteristics is lower than that for the existing estimator. Hence, the proposed estimator is efficient.


Keywords: Randomize Response, Sensitive Questions, Stratified Random Sampling.

### 1.0 Introduction

The randomized response technique (RRT) is useful for reducing response error problems when potentially sensitive questions such as the illegal use of drugs, sexual practice, illegal earning, or incidence of acts of domestic violence are included in surveys of human populations. Direct questioning of respondents about sensitive issues often results in either refusal or falsification of the answers. Social stigma and fear of reprisals sometimes result in untruthful, exaggerated, or misleading responses by respondents when approached with conventional survey methods. An ingenious method of counteracting fears in response to sensitive questionswas the first to suggested by warner [1]
Several authors had developed an efficient Randomized Response Techniques (RRTs), the developed techniques only considered a two-option of "yes" and "no" response [2]. In design, the two randomization devices are the same as that of Warner's device but with different probabilities of choosing the stigmatize question. The basic idea behind this suggestion is to reduce considerably the suspicion among the respondents by providing them choice to randomly select the randomization device itself. Consequently, respondents may reveal their true status. As a result, a new Re-parameterized Randomized Response Technique (RRRTs) that will be based on the random use of one of the three randomization devices was proposed [3].This work is proposing a modified Stratified Randomized Response Model based on the random use of one of the three randomization devices.
The basic idea behind this suggestion is to reduce considerably the suspicion among the respondents by providing each stratified group choice to randomly select the randomization device itself, which will increase response rate from the respondent.

Corresponding author: Olayiwola O.M, E-mail: laniyimathew @ yahoo.com, Tel.: +2348034305353

Direct interrogation about a stigmatized attribute such as induced abortion, use of drugs, tax evasion, etc. in a human population survey is a different exercise. A sampler may receive wrong answers from the survey respondent when he/she uses direct interrogating approach. Due to many reasons, information about prevalence of stigmatized attributes, in the population is essential. A complicated method of survey to gather information in relation to stigmatized attributes by ensuring confidentiality and anonymity to the respondents was proposed [1]. Up till now, a vast number of developmental improvements on Warner's Randomized Response Techniques (RRT) have been developed by several researchers [3-10] are some of the many to be listed. In some situations, prior information about the unknown parameter may be available and can be used along with the sample auxiliary information for determination of that unknown parameter known as the Bayesian approach of estimation. Work done by researchers on Bayesian analysis of randomized response model are not very much, nonetheless, attempts have been made on the Bayesian analysis of Randomized Response techniques [2,8,11-21] are the major references on the Bayesian analysis of the Randomized Response Techniques.
Existing Reparametized Randomized Response Model:
Despite the success achieved by many authors in developing an efficient Randomized Response Models (RRMs), the developed Models only considered a dichotomous option of "yes" and "no" response. In view of this, they propose a new Reparametized Randomized Response Model (RRRM) that was be based on the random use of one of the three randomization devices, In design, the three randomization devices are identical to that of Warner's device but with different probabilities of selection. In addition to $\alpha$ and $\beta$ proposed earlier by Hussain and Shabbir, we introduce $\delta$, a positive real number such that $\frac{\alpha}{\alpha+\beta+\delta}, \alpha \neq \beta \neq \delta$ is the probability of usingfirst randomization device which consists of the two statements of Warner's device and the new introduce device also with preset probabilities $P_{1}, P_{2}$ and $P_{3}$ respectively. By using Hussain and Shabbir's probability of a "yes" response for the $i^{t h}$ respondent, the probability of a "yes" response when the third option "undecided" is included is given by

$$
\varphi=\frac{\alpha}{\alpha+\beta+\delta}\left[p_{1} \pi+\left(1-p_{1}\right)(1-\pi)\right]+\frac{\beta}{\alpha+\beta+\delta}\left[p_{2} \pi+\left(1-p_{2}\right)(1-\pi)\right]+\frac{\delta}{\alpha+\beta+\delta}\left[p_{3} \pi+\left(1-p_{3}\right)(1-\pi)\right]
$$

The unbiased moment estimator of true probability of yes response (response rate) was given by

$$
\begin{equation*}
\pi=\frac{\varphi(\alpha+\beta+\delta)-\left[(\alpha+\beta+\delta)-p_{1} \alpha-p_{2} \beta-p_{3} \delta\right]}{2 p_{1} \alpha+2 p_{2} \beta+2 p_{3} \delta-\alpha-\beta-\delta} \tag{1}
\end{equation*}
$$

The variance of the estimator is given then by

$$
V\left(\hat{\pi}_{h}\right)_{\text {prop }}=\frac{\pi(1-\pi)}{n}+\frac{\left(p_{1} \alpha+p_{2} \beta+p_{3} \delta\right)\left(p_{3} \alpha+p_{2} \beta+p_{1} \delta\right)}{n\left[2 p_{1}(\alpha-\delta)+2 p_{2}(\beta-\delta)-(\alpha+\beta-\delta)\right]^{2}(\alpha+\beta+\delta)^{2}}
$$

### 2.0 Methodology

### 2.1 Proposed Modified Stratified Randomization Response Method

## Theorem 1

The probability of yes $Q($ yes $)$ for each stratum denoted as $\varphi_{h}$ is

$$
\begin{align*}
Q(\text { yes })=\varphi_{h} & =\frac{\alpha_{h}}{\alpha_{h}+\beta_{h}+\delta_{h}}\left[p_{1 h} \pi_{h}+\left(1-p_{1 h}\right)\left(1-\pi_{h}\right)\right] \\
& +\frac{\beta_{h}}{\alpha_{h}+\beta_{h}+\delta_{h}}\left[p_{2 h} \pi_{h}+\left(1-p_{2 h}\right)\left(1-\pi_{h}\right)\right]  \tag{2}\\
& +\frac{\delta_{h}}{\alpha_{h}+\beta_{h}+\delta_{h}}\left[p_{3 h} \pi_{h}+\left(1-p_{3 h}\right)\left(1-\pi_{h}\right)\right]
\end{align*}
$$

So that,
$\hat{\varphi}=\sum W_{h} \varphi_{h}$
Where $W_{h}$ is the stratum weights and $W_{h}=\frac{N_{h}}{N}$ for $h=1,2, \ldots, L$ so that $W=\sum_{h=1}^{L} W_{h}=1(\mathrm{~N}$ is the number of unit in the population and $N_{h}$ is the total number of units in stratum h).

In order to ensure equal privacy protection in the three randomization devices $\mathrm{X}_{1}, \mathrm{X}_{2}$, and $\mathrm{X}_{3}$, we put $p_{1}=1-p_{2}-p_{3}$ into equation (1), to get
$\pi_{h}=\frac{\varphi_{h}\left(\alpha_{h}+\beta_{h}+\delta_{h}\right)-\left[\left(\alpha_{h}+\beta_{h}+\delta_{h}\right)-p_{1 h} \alpha_{h}-p_{2 h} \beta_{h}-p_{3 h} \delta_{h}\right]}{2 p_{1 h} \alpha_{h}+2 p_{2 h} \beta_{h}+2 p_{3 h} \delta_{h}-\alpha_{h}-\beta_{h}-\delta_{h}}$
Hence, the unbiased sample estimate of $\pi_{h}$ is given as

$$
\begin{equation*}
\hat{\pi}_{h}=\frac{\hat{\varphi}_{h}\left(\alpha_{h}+\beta_{h}+\delta_{h}\right)-\left[\left(\alpha_{h}+\beta_{h}+\delta_{h}\right)-p_{1 h} \alpha_{h}-p_{2 h} \beta_{h}-p_{3 h} \delta_{h}\right]}{2 p_{1 h} \alpha_{h}+2 p_{2 h} \beta_{h}+2 p_{3 h} \delta_{h}-\alpha_{h}-\beta_{h}-\delta_{h}} \tag{4}
\end{equation*}
$$

Where $\hat{\varphi}=\frac{x}{n}$ and x is the number of respondents reporting a "yes" answer
Therefore,
$\hat{\pi}=\sum W_{h} \pi_{h}$
Where $W_{h}$ is the stratum weights and $W_{h}=\frac{N_{h}}{N}$ for $h=1,2, \ldots, L$ so that $W=\sum_{h=1}^{L} W_{h}=1(\mathrm{~N}$ is the number of unit in the population and $N_{h}$ is the total number of units in stratum h)
Setting $p_{1}=1-p_{2}-p_{3}$, the variance of the estimator for each stratum is then given by

$$
\left.\begin{array}{l}
{\left[\begin{array}{c}
\alpha_{h}+\beta_{h} \\
+\delta_{h}
\end{array}\right]^{2}\left[\begin{array}{c}
\pi_{h}\left(4 \alpha_{h}^{2} p_{1 h}^{2}-4 \alpha_{h}^{2} p_{1 h}-4 \alpha_{h} \beta_{h} p_{1 h}\right. \\
+8 \alpha_{h} \beta_{h} p_{1 h} p_{2 h}-4 \alpha_{h} \delta_{h} p_{1 h}+8 \alpha_{h} \delta_{h} p_{1 h} p_{3 h} \\
+\alpha_{h}^{2}+2 \alpha_{h} \beta_{h}-4 \alpha_{h} \beta_{h} p_{2 h}+2 \alpha_{h} \delta_{h} \\
-4 \alpha_{h} \delta_{h} p_{3 h}-4 \beta_{h}^{2} p_{2 h}+4 \beta_{h}^{2} p_{2 h}^{2} \\
-4 \beta_{h} \delta_{h} p_{2 h}+8 \beta_{h} \delta_{h} p_{2 h} p_{3 h}+\beta_{h}^{2} \\
+2 \beta_{h} \delta_{h}-4 \beta_{h} \delta_{h} p_{3 h}-4 \delta_{h}^{2} p_{3 h}+4 \delta_{h}^{2} p_{3 h}^{2} \\
\left.+\delta_{h}^{2}\right)-\pi_{h}^{2}\left(4 \alpha_{h}^{2} p_{1 h}^{2}-4 \alpha_{h}^{2} p_{1 h}\right. \\
-4 \alpha_{h} \beta_{h} p_{1 h}+8 \alpha_{h} \beta_{h} p_{1 h} p_{2 h}-4 \alpha_{h} \delta_{h} p_{1 h} \\
+8 \alpha_{h} \delta_{h} p_{1 h} p_{3 h}+\alpha_{h}^{2}+2 \alpha_{h} \beta_{h}-4 \alpha_{h} \beta_{h} p_{2 h} \\
+2 \alpha_{h} \delta_{h}-4 \alpha_{h} \delta_{h} p_{3 h}-4 \beta_{h}^{2} p_{2 h} \\
+4 \beta_{h}^{2} p_{2 h}^{2}-4 \beta_{h} \delta_{h} p_{2 h}+8 \beta_{h} \delta_{h} p_{2 h} p_{3 h} \\
+\beta_{h}^{2}+2 \beta_{h} \delta_{h}-4 \beta_{h} \delta_{h} p_{3 h}-4 \delta_{h}^{2} p_{3 h} \\
\left.+4 \delta_{h}^{2} p_{3 h}^{2}+\delta_{h}^{2}\right)
\end{array}\right]+\left[\begin{array}{c}
\alpha_{h}^{2} p_{1 h}+\alpha_{h} \beta_{h} p_{2 h} \\
+\alpha_{h} \delta_{h} p_{3 h}-\alpha_{h}^{2} p_{1 h}^{2} \\
+\alpha_{h} \beta_{h} p_{1 h}-2 \alpha_{h} \beta_{h} p_{1 h} p_{2 h} \\
+\alpha_{h} \delta_{h} p_{1 h}-2 \alpha \delta p_{1} p_{3} \\
+\beta_{h}^{2} p_{2 h}+\beta_{h} \delta_{h} p_{3 h} \\
-\beta_{h}^{2} p_{2 h}^{2}+\beta_{h} \delta_{h} p_{2 h} \\
-2 \beta_{h} \delta_{h} p_{2 h} p_{3 h}+\delta_{h}^{2} p_{3 h} \\
-\delta_{h}^{2} p_{3 h}^{2}
\end{array}\right]}
\end{array}\right] \begin{gathered}
\\
n_{h}\left[\begin{array}{c}
\left.2 p_{1 h}\left(\alpha_{h}-\delta_{h}\right)+2 p_{2 h}\left(\beta_{h}-\delta_{h}\right)-\left(\alpha_{h}+\beta_{h}-\delta_{h}\right)\right]^{2}\left(\alpha_{h}+\beta_{h}+\delta_{h}\right)^{2}
\end{array}\right]
\end{gathered}
$$

On simplification, we have
$V\left(\hat{\pi}_{h}\right)_{\text {prop }}=\frac{\pi_{h}\left(1-\pi_{h}\right)}{n_{h}}+\frac{\left(p_{1 h} \alpha_{h}+p_{2 h} \beta_{h}+p_{3 h} \delta_{h}\right)\left(p_{3 h} \alpha_{h}+p_{2 h} \beta_{h}+p_{1 h} \delta_{h}\right)}{n_{h}\left[2 p_{1 h}\left(\alpha_{h}-\delta_{h}\right)+2 p_{2 h}\left(\beta_{h}-\delta_{h}\right)-\left(\alpha_{h}+\beta_{h}-\delta_{h}\right)\right]^{2}\left(\alpha_{h}+\beta_{h}+\delta_{h}\right)^{2}}$
The variance of the estimator $\hat{\pi}$ is then given as;

$$
V(\hat{\pi})_{\text {prop }}=\frac{1}{n} \sum_{h=1}^{l} W_{h}\left[\pi_{h}\left(1-\pi_{h}\right)+\frac{\left(p_{1 h} \alpha_{h}+p_{2 h} \beta_{h}+p_{3 h} \delta_{h}\right)\left(p_{3 h} \alpha_{h}+p_{2 h} \beta_{h}+p_{1 h} \delta_{h}\right)}{\left[2 p_{1 h}\left(\alpha_{h}-\delta_{h}\right)+2 p_{2 h}\left(\beta_{h}-\delta_{h}\right)-\left(\alpha_{h}+\beta_{h}-\delta_{h}\right)\right]^{2}\left(\alpha_{h}+\beta_{h}+\delta_{h}\right)^{2}}\right]
$$

### 2.2 Sampling Strategy

For the purpose of this study, five hundred (500) questionnaires were constructed to know some of the wrong vices practiced among students in which eight different sensitive questions were administered. The students of Federal University of Agriculture, Abeokuta, (FUNAAB) Ogun state, Nigeria were considered. The questionnaire consists of two sections. Section A consists of the Socio- Demographic Data and section B consists of the some selected wrong vices practiced among
university students. Three randomization devices were used. The randomization devices used were coin, dice and deck of cards. The possible outcomes for the tossing of coins is head or tail, for the dice the possible outcome is $1,2,3,4,5$ and 6 while for the deck of card which has 54 cards with 13 hearts, 13 clubs, 13 spade and 13 diamond shapes. Each student were given an opportunity of picking any of the three devices randomly. Any student who picked the coin and dice was asked to toss privately and note the outcome, while those who picked the deck of card were asked to shuffle the cards and pick their choice, this was also done privately. For those outcome of the rolling of the dice was $1,2,3$ and 4 were required to answer the questions in the questionnaire privately and truthfully while those whose outcome of the rolling of the dice was 4 and 5 were required to tick 'yes' option for all the sensitive questions. For those whose outcome of the tossing of the coins was head were required to answer the questions privately and truthfully otherwise was required to tick 'yes' option for the questions. For those who picked the deck of cards, those whose outcome after shuffling and picking was a heart, spade and diamond were required to answer the questionnaire privately and truthfully otherwise they were instructed to tick the 'yes' option in the questionnaire. After each respondent had answered the questionnaire as applicable to him/her, they were then asked to submit the questionnaire.

### 3.0 Results and Discussion

The following section deals with the descriptive statistics of the administered sensitive questions. The first question as shown in table 1 revealed that $209(41.8 \%$ ) students ticked the 'Yes', $285(57.0 \%)$ tick ' $N$ ' while $6(1.2 \%)$ tick 'undecided', second question shows that $207(41.8 \%$ ) students ticked 'Yes' answers, $290(58.0 \%)$ ticked ' $N o$ ' and $3(0.6 \%)$ ticked 'undecided', third question shows that $203(40.9 \%$ ) students ticked 'Yes' answers, 292(58.4\%) ticked ' No ' and $5(1 \%)$ ticked undecided, fourth question shows that $199(39.8 \%$ ) students ticked 'Yes' answers and $301(60.2 \%$ ) ticked ' No ', fifth question shows that $196(39.2 \%)$ students ticked 'Yes' answers, 302(60.4\%) ticked ' No ' and $2(0.4 \%$ ) ticked undecided, sixth question shows that $200(40.0 \%)$ students ticked 'Yes' answers, $292(58.4 \%)$ ticked ' $N o$ ' and $8(1.6 \%)$ ticked undecided, seventh question shows that $268(53.6 \%)$ students ticked 'Yes' answers, $225(45.0 \%)$ ticked 'No' and $7(1.4 \%)$ ticked undecided, question eight shows that 214(42.8\%) students ticked 'Yes' answers, 284(56.8\%) ticked 'No' and 2(0.4\%) ticked undecided,

Table 1:Frequency of Respondents for the 500 Questionnaires

| Question | Sample size |  |  |  |  |  |  | Yes |  |  | No |  | Undecided |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Frequency | Percent | Frequency | Percent | Frequency | Percent |  |  |  |  |  |  |  |  |
| Have you contracted sexually <br> transmitted disease before? | 209 | 41.8 | 285 | 57.0 | 6 | 1.2 | 500 |  |  |  |  |  |  |  |
| Have you raped or being raped <br> before? | 207 | 41.4 | 290 | 58.0 | 3 | 0.6 | 500 |  |  |  |  |  |  |  |
| Have you aborted or aborted for <br> someone before1? | 203 | 40.9 | 292 | 58.4 | 5 | 1.0 | 500 |  |  |  |  |  |  |  |
| Are you a member of a secret cult? | 199 | 39.8 | 301 | 60.2 | 0 | 0 | 500 |  |  |  |  |  |  |  |
| Have you killed before? | 196 | 39.2 | 302 | 60.4 | 2 | 0.4 | 500 |  |  |  |  |  |  |  |
| Have you robbed a fellow student or <br> stole during e-examinations before? | 200 | 40.0 | 292 | 58.4 | 8 | 1.6 | 500 |  |  |  |  |  |  |  |
| Have you cheated during an <br> examination in FUNAAB before? | 268 | 53.6 | 225 | 45.0 | 7 | 1.4 | 500 |  |  |  |  |  |  |  |
| Have you bribed a lecturer for <br> marks before? | 214 | 42.8 | 284 | 56.8 | 2 | 0.4 | 500 |  |  |  |  |  |  |  |

### 3.1 Existing Procedure

When the 300 sample was selected, to know the particular wrong vices that is been carried out by the students of FUNAAB, cheating in an examination was the question with highest percentage across the varied sample sizes.
Using equation 1 , when $\alpha=20, \beta=11$ and $\delta=2$, an estimate and the mean square error(MSE) of true proportion of students who engaged in some of the wrong vices using existing procedure are as shown in table 2 . The table shows that for 300 sample size, question 7 which talks about examination malpractice has the highest proportion among all other vices in which FUNAAB students do engage in.

Table 2: Proportion of Students who engaged in Wrong Vices and its Varianceusing existing procedure.

| Sample size | Question | Proportion of students with the stigmatized or sensitive behavior $(\pi)$ | Variance of $(\pi)$ |
| :--- | :--- | :--- | :--- |
| 300 | 1 | 0.41073 | 0.000819 |
|  | 2 | 0.390893 | 0.000806 |
|  | 3 | 0.380974 | 0.000799 |
|  | 4 | 0.376014 | 0.000795 |
|  | 5 | 0.366096 | 0.000786 |
|  | 6 | 0.371055 | 0.000791 |
| 7 | 0.891794 | 0.000334 |  |
|  | 8 | 0.400812 | 0.000813 |

### 3.2 Proposed Procedure

A simple random sampling scheme was used to select 500 students which comprised of 332 males and 168 females among students of Federal University of Agriculture, Abeokuta, Ogun State. These students were stratified into two strata using gender as stratifying factor and proportional allocation procedure was used to select 199 male students and 101 female students randomly. For $\alpha=20, \beta=11$ and $\delta=2$, an estimate and the mean square error(MSE) of true proportion of male students and female students who engaged in some of the wrong vices are as shown in table 6 and 7 respectively. From the tables it can be seen that question 7 which talks about examination malpractice has the highest proportion among all other vices in which FUNAAB students engage in. For male, for the sample size of 300 , question 7 is the highest proportion among all vices, while all other questions have the same proportion of male students who engage in wrong vices. An estimate and the mean square error(MSE) of true proportion of female students who engaged in some of the wrong vices are as shown in table 7 . From the table it can be seen that question 7 has the highest proportion among all vices, followed by question 8 , while all other questions has the same proportion of male students who engage in wrong vices.
Table 3: Frequency of Male Respondents

| Question | Sample <br> size | Yes | No |  | Undecided |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  | Frequency | Percent | Frequency | Percent | Frequency | Percent |  |
| Have you contracted sexually <br> transmitted disease before? | 199 | 79 | 39.7 | 117 | 58.8 | 3 | 1.5 |
| Have you raped or being raped <br> before? | 199 | 77 | 38.7 | 120 | 60.3 | 2 | 1.0 |
| Have you aborted or aborted for <br> someone before1? | 199 | 76 | 38.2 | 118 | 59.3 | 5 | 2.6 |
| Are you a member of a secret <br> cult? | 199 | 79 | 39.7 | 120 | 60.3 | 0 | 0 |
| Have you killed before? | 199 | 75 | 37.7 | 123 | 61.8 | 1 | 0.5 |
| Have you robbed a fellow <br> student or stole during <br> examinations before? | 199 | 77 | 38.7 | 118 | 59.3 | 4 | 2.0 |
| Have you cheated during an <br> examination in <br> before? | 199 | 100 | 50.3 | 95 | 47.7 | 4 | 2.0 |
| Have you bribed a lecturer for <br> marks before? | 199 | 77 | 38.7 | 120 | 60.3 | 2 | 0.5 |

Table 4: Frequency of Female Respondents

| Question | Sample size | Yes |  | No |  | Undecided |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Frequency | Percent | Frequency | Percent | Frequency | Percent |
| Have you contracted sexually transmitted disease before? | 101 | 40 | 39.6 | 60 | 59.4 | 1 | 1.0 |
| Have you raped or being raped before? | 101 | 41 | 40.6 | 60 | 59.4 | 0 | 0 |
| Have you aborted or aborted for someone before1? | 101 | 42 | 41.6 | 59 | 58.4 | 0 | 0 |
| Are you a member of a secret cult? | $101$ | 38 | 37.6 | 63 | 62.4 | 0 | 0 |
| Have you killed before? | 101 | 38 | 37.6 | 63 | 62.4 | 0 | 0 |
| Have you robbed a fellow student or stole during eexaminations before? | 101 | 41 | 40.6 | 58 | 57.4 | 2 | 2.0 |
| Have you cheated during an examination in FUNAAB before? | $101$ | 54 | 53.5 | 46 | 45.5 | 1 | 1.0 |
| Have you bribed a lecturer for marks before? | $101$ | 45 | 44.6 | 56 | 55.4 | 0 | 0 |

Table 5: Proportion of Male Students who engage in Wrong Vices and its Variance

| Sample size | Question | proportion of students with the stigmatized or sensitive behavior $(\pi)$ | Variance of $(\pi)$ |
| :--- | :--- | :--- | :--- |
| 300 | 1 | 0.346731 | 0.001157 |
|  | 2 | 0.331778 | 0.001133 |
|  | 0.324302 | 0.00112 |  |
|  | 0.346731 | 0.001157 |  |
|  | 0.316825 | 0.001107 |  |
|  | 0.331778 | 0.001133 |  |
|  | 0.944853 | 0.000281 |  |
| 7 | 0.331778 | 0.001133 |  |

Table 6: Proportion of Female Students who engaged in Wrong Vices and its Variance

| Sample size | Question | proportion of students with the stigmatized or sensitive behavior $(\pi)$ | Variance of $(\pi)$ |
| :--- | :--- | :--- | :--- |
| 300 | 1 | 0.345325 | 0.002276 |
|  | 2 | 0.360056 | 0.002319 |
| 3 | 0.374787 | 0.002358 |  |
|  | 0.315863 | 0.002177 |  |
|  |  | 0.315863 | 0.002177 |
|  | 0.360056 | 0.002319 |  |
|  | 6 | 0.964025 | 0.000381 |
| 7 | 0.41898 | 0.002448 |  |

When $\alpha=20, \beta=11$ and $\delta=2$, the combined estimate and the mean square error(MSE) of true proportion of students who engaged in some of the wrong vices are as shown in table 8 . From table 9 , it can be seen that across the entire selected sensitive question, the proposed procedure with stratification is more efficient than the existing without stratification. This is represented in the line graph below. It should be noted that the proportion of students who cheats in an examination is the highest among all other vices. This suggests that the higher the proportion of people with stigmatized behavior the lower the variances.

Table 7: Proportion of Students who engaged in Wrong Vices and its Variance Using the Proposed Procedure

| Sample size | Question | proportion of students with the stigmatized or sensitive behavior $(\pi)$ | Variance of $(\pi)$ |
| :--- | :--- | :--- | :--- |
| 300 | 1 | 0.346259 | 0.000767 |
|  | 2 | 0.34128 | 0.000761 |
|  | 3 | 0.341265 | 0.00076 |
|  | 4 | 0.33636 | 0.000756 |
|  | 5 | 0.316502 | 0.000734 |
|  | 6 | 0.34128 | 0.000761 |
| 7 | 0.951295 | 0.000166 |  |
|  | 0.361078 | 0.000776 |  |

Table 8: Efficiency of the Proposed Procedure over the Existing Procedure

| Sample size | Question | Variance of $(\pi)$ <br> Proposed | Variance of $(\pi)$ <br> Existing | Efficiency |
| :--- | :--- | :--- | :--- | :--- |
| 300 | 1 | 0.000767 | 0.000819 | 1.067797 |
|  | 2 | 0.000761 | 0.000806 | 1.059133 |
|  | 3 | 0.00076 | 0.000799 | 1.051316 |
|  | 4 | 0.000756 | 0.000795 | 1.051587 |
|  | 5 | 0.000734 | 0.000786 | 1.070845 |
|  | 6 | 0.000761 | 0.000791 | 1.039422 |
|  | 7 | 0.000166 | 0.000334 | 2.012048 |
|  | 8 | 0.000776 | 0.000813 | 1.04768 |



Figure 1: $\quad$ Efficiency of the Proposed Procedure over the Existing Procedure

### 4.0 Conclusion

The proportion of male students who had contracted sexually transmitted disease, involved in rape, abortion, cultism, murder, stealing and examination malpractice is higher than that for the females. The proposed estimator showed that the variance of proportion for both male and female students involved in the sensitive and stigmatized characteristics is lower than that for the existing estimator. Hence, the proposed estimator is efficient.

### 5.0 References

[1] Warner, S.L. 1965. Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias, Journal of the American Statistical Association 60: 63-69.
[2] Hussain, Z. and Shabbir, J. 2012. Bayesian Estimation of Population Proportion in Kim and Warde Mixed Randomized Response Technique. Electronic Journal of Applied Statistical Analysis, Vol. 5, Issue 2: 213 - 225.
[3] Adebola, F.B. and Aydepetun, A.O. 2011. A New Tripartite Randomized Response Technique. Journal of the Nigerian Association of Mathematical Physics, Volume 19: 119-122.
[4] Greenberg, B.G., Abul-Ela, A.A., Simmons, W.R. and Horvitz, D.G. 1969. The Unrelated Question Randomized Response: Theoretical Framework, Journal of the American Statistical Association 64: 520-539.
[5] Mangat, N.S. and Singh, R. 1990. An Alternative Randomized Response Procedure. Biometrika 77: 439-442.
[6] Mangat, N.S. 1994. An Improved Randomized Response Strategy. J. Roy. Statist. Soc. Ser. B, 56(1): 93-95.
[7] Singh, S., Horn, S. and Chowdhuri, S. 1998. Estimation of Stigmatized Characteristics of a Hidden Gang in Finite Population.Austral.\& New Zealand J. Statist, 40(3): 291-297.
[8] Kim, J.M., Tebbs, J.M. and An, S.W. 2006. Extension of Mangat's Randomized Response Model. Journal of Statistical Planning and Inference, 36: 1554-1567.
[9] Adebola, F.B. and Adepetun, A.O. 2012. On a Numerical Comparison of the Proposed Randomized Response Technique with Hussain and Shabbir. Journal of the Nigerian Association of Mathematical Physics, Volume 20: 379-384.
[10] Adepetun, A.O. and Adebola, F.B. 2014. On the Relative Efficiency of the Proposed Reparametized Randomized Response Model. International Journal of Mathematical Theory and Modeling, Volume 4: 58-67.
[11] Winkler, R. and Franklin, L. 1979. Warner's Randomized Response Model: A Bayesian Approach. Journal of the American Statistical Association, 74: 207-214
[12] Pitz, G.F. 1980. Bayesian Analysis of Random Response Models. Psychological Bulletin, 87: 209-212.
[13] Spurrier, J. and Padgett, W. 1980. The Application of Bayesian Techniques in Randomized Response. Sociological Methodology, 11: 533-544.
[14] O’Hagan, A. 1987. Bayes Linear Estimators for Randomized Response Models. Journal of the American Statistical Association, 82: 580-585.
[15] Migon, H. and Tachibana, V. 1997. Bayesian Approximations in Randomized Response Models. Computational Statistics and Data Analysis, 24: 401-409.
[16] Unnikrishnan, N. and Kunte, S. 1999. Bayesian Analysis for Randomized Response Models. Sankhya, B, 61, 422432.
[17] Bar-Lev, S., Bobovich, E. and Boukai, B.A. 2003. Common Conjugate Prior Structure for Several Randomized Response Models. TEST, 2003; 12: 101-113.
[18] Barabesi, L. and Marcheselli, M. 2006. A Practical Implementation and Bayesian Estimation in Franklin's Randomized Response Procedure. Communication in Statistics- Simulation and Computation, 35: 365-573.
[19] Hussain, Z., and Shabbir, J. 2009a. Bayesian Estimation of Population Proportion of a Sensitive Characteristic using Simple Beta Prior. Pakistan Journal of Statistics, 25(1): 27-35.
[20] Hussain, Z. and Shabbir, J. 2009b. Bayesian Estimation of population proportion in Kim and Warde (2005) Mixed Randomized Response using Mixed Prior Distribution. Journal of probability and Statistical Sciences, 7(1): 71-80.
[21] Adepetun, A.O. and Adewara, A.A. 2014. Bayesian Analysis of Kim and Warde Randomized Response Technique using Alternative Priors. American Journal of computation and Applied Mathematics, 4(4): 130-140

