Journal of the Nigerian Association of Mathematical Physics
Volume **20 (March, 2012), pp 379 – 384**
© *J. of NAMP*

On a Numerical Comparison of the Proposed Randomized Response Technique with Hussain and Shabbir

[1]Adebola, F.B and [2]Adepetun, A.O

Department of Mathematical Sciences,
Federal University of Technology, PMB 704,Akure,
Ondo State, Nigeria.

Abstract

In this paper, we compared our newly developed Randomized Response Technique (RRT) with that of Hussain-Shabbir's dichotomous Randomized Response Technique (RRT) when data are obtained through the randomized response technique (RRT) proposed by Hussain and Shabbir (2007). It was established that the variance of the proposed technique is less than that of the conventional technique for various orders of probabilities of each answer option. Hence, the proposed technique is more efficient than the conventional technique.

Keywords: Probability of Selection, Privacy Protection, Sensitive Characters, Randomization Device

1.0 Introduction

The problem of estimation of the total population of a sensitive quantitative variable is well known in survey sampling. Warner [8] was the first to suggest an ingenuous method to estimate the proportion of sensitive characters like induced abortions, drug used etc., through a randomization device like a deck of cards, spinners etc. such that the respondents' privacy should be protected. To date, a large number of developments and variants of Warner's Randomized Response Technique (RRT) have been put forward by several researchers. Greenberg et al.[2], Mangat and Singh [5], Mangat [6], Singh et al.[7], Christofides [1], Kim and Warde [4] are some of the many to be referenced. In sections to follow, we present Hussain and Shabbir (2007), Proposed Randomized Response Technique and subsequently its efficiency over the conventional one.

2.0 Hussain and Shabbir Technique

Hussain and Shabbir (2007) proposed a Randomized Response Technique (RRT) based on the random use of one of the two randomization devices R_1 and R_2. In design, the two randomization devices R_1 and R_2 are the same as that of Warner's (1965) device but with different probabilities of selecting the sensitive question. The idea behind this suggestion is to decrease the suspicion among the respondents by providing them choice to randomly choose the randomization device itself. As a result, respondents may divulge their true status. A simple random sample with replacement (SRSWR) sampling is assumed to select a sample of size n. Let $\propto$ and β be any two positive real numbers chosen such that $q = \frac{\alpha}{\propto+\beta}, (\alpha \neq \beta)$ is the probability of using R_1, where R_1 consists of the two statements of Warner's device but with preset probabilities P_1 and $1 - P_1$ and $1 - q = \frac{\beta}{\propto+\beta}$ is the probability of using R_2 ,where R_2 consists of the two statements of Warner's device also with preset probabilities P_2 and $1 - P_2$ respectively. For the i^{th} respondent, the probability of a "yes" response is given by

$$P(\text{yes}) = \emptyset = \frac{\alpha}{\alpha+\beta}[P_1\pi + (1 - P_1)(1-\pi)] + \frac{\beta}{\alpha+\beta}[P_2\pi + (1 - P_2)(1-\pi)] \tag{2.1}$$

To provide the equal privacy protection in both the randomization devices R_1 and R_2, we put $P_1 = 1 - P_2$ into equation (2.1),to obtain

$$\emptyset = \frac{\pi[(\alpha-\beta)(2P_1-1)] + P_1\beta + P_2\alpha}{\alpha+\beta} \tag{2.2}$$

[1]Corresponding author: ***Adebola, F.B,*** E-mail: akinolaoladiran@yahoo.com, Tel.: +2348038380448

Hence,

$$\pi = \frac{\emptyset(\alpha+\beta) - P_1\beta - P_2\alpha}{(2P_1 - 1)(\alpha-\beta)}, P_1 \neq 1/2, \alpha \neq \beta \qquad (2.3)$$

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

$$\hat{\pi} = \frac{\hat{\emptyset}(\alpha+\beta) - P_1\beta - P_2\alpha}{(2P_1 - 1)(\alpha-\beta)} \qquad (2.4)$$

Where $\hat{\emptyset} = \frac{y}{n}$ and y is the number of respondents reporting a "yes" answer when $P_1 = 1 - P_2$. The variance of the estimator was given then by

$$V(\hat{\pi})_{conv} = \frac{\pi(1-\pi)}{n} + \frac{(P_2\alpha + P_1\beta)(P_1\alpha + P_2\beta)}{n(2P_1-1)^2(\alpha-\beta)^2(\alpha+\beta)^2} \qquad (2.5)$$

3.0 The Proposed Technique

It has been discovered that despite the successful attempts by several authors in developing an efficient Randomized Response Techniques (RRTs), the developed techniques only considered a two-option of "yes" and "no" response. As a result of which we propose a new Randomized Response Technique (RRT) that will be based on the random use of one of the three randomization devices, R_1, R_2 and R_3. In design, the three randomization devices R_1, R_2 and R_3 are similar to that of Warner's device but with different probabilities of selection. In addition to α and β proposed earlier by Hussain and Shabbir, we introduce δ, a positive real number such that $q = \frac{\alpha}{\alpha+\beta+\delta}, \alpha \neq \beta \neq \delta$ is the probability of using R_1, where R_1 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 adopting Hussain and Shabbir's probability of a "yes" response for the i[th] respondent, the probability of a "yes" response when the third option "undecided" is included is given by

$$Q(yes) = \varphi = \frac{\alpha}{\alpha+\beta+\delta}[P_1\pi + (1-P_1)(1-\pi)] + \frac{\beta}{\alpha+\beta+\delta}[P_2\pi + (1-P_2)(1-\pi)] + \frac{\delta}{\alpha+\beta+\delta}[P_3\pi + (1-P_3)(1-\pi)] \qquad (3.1)$$

In order to provide the equal privacy protection in the three randomization devices R_1, R_2, and R_3, we put $P_1 = 1 - P_2 - P_3$ into equation (3.1), to obtain

$$\pi = \frac{\varphi(\alpha+\beta+\delta) - [(\alpha+\beta+\delta) - P_1\alpha - P_2\beta - P_3\delta]}{2P_1\alpha + 2P_2\beta + 2P_3\delta - \alpha - \beta - \delta} \qquad (3.2)$$

Hence, the unbiased sample estimate of π is given as

$$\hat{\pi} = \frac{\hat{\varphi}(\alpha+\beta+\delta) - [(\alpha+\beta+\delta) - P_1\alpha - P_2\beta - P_3\delta]}{2P_1\alpha + 2P_2\beta + 2P_3\delta - \alpha - \beta - \delta} \qquad (3.3)$$

Where $\hat{\varphi} = \frac{x}{n}$ and x is the number of respondents reporting a "yes" answer when $P_1 = 1 - P_2 - P_3$. The variance of the estimator is given then by

$$V(\hat{\pi}) = \frac{(\alpha+\beta+\delta)^2 \begin{bmatrix} \pi(4\alpha^2P_1^2 - 4\alpha^2P_1 - 4\alpha\beta P_1 + 8\alpha\beta P_1P_2 - 4\alpha\delta P_1 \\ +8\alpha\delta P_1P_3 + \alpha^2 + 2\alpha\beta - 4\alpha\beta P_2 + 2\alpha\delta \\ -4\alpha\delta P_3 - 4\beta^2P_2 + 4\beta^2P_2^2 - 4\beta\delta P_2 + 8\beta\delta P_2P_3 \\ +\beta^2 + 2\beta\delta - 4\beta\delta P_3 - 4\delta^2P_3 + 4\delta^2P_3^2 + \delta^2) \\ -\pi^2(4\alpha^2P_1^2 - 4\alpha^2P_1 - 4\alpha\beta P_1 + 8\alpha\beta P_1P_2 \\ -4\alpha\delta P_1 + 8\alpha\delta P_1P_3 + \alpha^2 + 2\alpha\beta - 4\alpha\beta P_2 \\ +2\alpha\delta - 4\alpha\delta P_3 - 4\beta^2P_2 + 4\beta^2P_2^2 - 4\beta\delta P_2 \\ +8\beta\delta P_2P_3 + \beta^2 + 2\beta\delta - 4\beta\delta P_3 - 4\delta^2P_3 \\ +4\delta^2P_3^2 + \delta^2) \end{bmatrix} \begin{matrix} +[\alpha^2P_1 + \alpha\beta P_2 + \alpha\delta P_3 - \alpha^2P_1^2 + \alpha\beta P_1 - 2\alpha\beta P_1P_2 \\ +\alpha\delta P_1 - 2\alpha\delta P_1P_3 + \beta^2P_2 + \beta\delta P_3 - \beta^2P_2^2 + \\ \beta\delta P_2 - 2\beta\delta P_2P_3 + \delta^2P_3 - \delta^2P_3^2] \end{matrix}}{n[2P_1(\alpha-\delta) + 2P_2(\beta-\delta) - (\alpha+\beta-\delta)]^2(\alpha+\beta+\delta)^2} \qquad (3.4)$$

Hence, we have

$$V(\hat{\pi}) = \frac{\pi(1-\pi)}{n} + \frac{(P_1\alpha + P_2\beta + P_3\delta)(P_3\alpha + P_2\beta + P_1\delta)}{n[2P_1(\alpha-\delta) + 2P_2(\beta-\delta) - (\alpha+\beta-\delta)]^2(\alpha+\beta+\delta)^2} \qquad (3.5)$$

4.0 Efficiency Comparison

Here, we show that the new RRT is better than the existing ones by comparing its variance with the variance of Hussain-Shabbir technique under consideration. We adopt the data used by Hussain and Shabbir (2007) in comparing their results with others. In what follows, the proposed tripartite Randomized Response Technique (RRT) is more efficient than Hussain and Shabbir (2007) dichotomous Randomized Response Technique (RRT) if we have

$$V(\hat{\pi})_{conv} - V(\hat{\pi})_{prop} > 0 \qquad (4.1)$$

Or if

$$\frac{\pi(1-\pi)}{n}+\frac{(P_2\alpha+P_1\beta)(P_1\alpha+P_2\beta)}{n(2P_1-1)^2(\alpha-\beta)^2(\alpha+\beta)^2}-\frac{\pi(1-\pi)}{n}-\frac{(P_1\alpha+P_2\beta+P_3\delta)(P_3\alpha+P_2\beta+P_1\delta)}{n[2P_1(\alpha-\delta)+2P_2(\beta-\delta)-(\alpha+\beta-\delta)]^2(\alpha+\beta+\delta)^2}>0 \quad (4.2)$$

Or if

$$\frac{(P_2\alpha+P_1\beta)(P_1\alpha+P_2\beta)}{n(2P_1-1)^2(\alpha-\beta)^2(\alpha+\beta)^2}-\frac{(P_1\alpha+P_2\beta+P_3\delta)(P_3\alpha+P_2\beta+P_1\delta)}{n[2P_1(\alpha-\delta)+2P_2(\beta-\delta)-(\alpha+\beta-\delta)]^2(\alpha+\beta+\delta)^2}>0 \quad (4.3)$$

Table 4.1:Comparison between conventional RRT and proposed RRT when $P_1 = 0.6, P_2 = 0.3, P_3 = 0.1, \pi = 0.5, \alpha = 20, \beta = 11, \delta = 2$, for varying sample sizes (n)

n	P_1	P_2	P_3	π	α	β	δ	Conventional Variance (eqn. 2.5)	Proposed Variance (eqn.3.5)
50	0.6	0.3	0.1	0.5	20	11	2	0.00624	0.00546
100	0.6	0.3	0.1	0.5	20	11	2	0.00312	0.00273
150	0.6	0.3	0.1	0.5	20	11	2	0.00208	0.00182
200	0.6	0.3	0.1	0.5	20	11	2	0.00156	0.00136
500	0.6	0.3	0.1	0.5	20	11	2	0.000624	0.000546

Table 4.2 :Comparison between conventional RRT and proposed RRT when $P_1 = 0.4, P_2 = 0.4, P_3 = 0.2$, $\pi = 0.4$, $\alpha = 20, \beta = 11, \delta = 2$, for varying sample sizes (n)

n	P_1	P_2	P_3	π	α	β	δ	Conventional Variance (eqn.2.5)	proposed variance (eqn.3.5)
50	0.4	0.4	0.2	0.4	20	11	2	0.00579	0.00484
100	0.4	0.4	0.2	0.4	20	11	2	0.00289	0.00242
150	0.4	0.4	0.2	0.4	20	11	2	0.00193	0.00161
200	0.4	0.4	0.2	0.4	20	11	2	0.00145	0.00121
500	0.4	0.4	0.2	0.4	20	11	2	0.000579	0.000484

Table 4.3: Comparison between conventional RRT and proposed RRT when $P_1 = 0.2, P_2 = 0.5, P_3 = 0.3$, $\pi = 0.7$, $\alpha = 20, \beta = 11, \delta = 2$, for varying sample sizes (n)

n	P_1	P_2	P_3	π	α	β	δ	conventional variance (eqn.2.5)	Proposed Variance (eqn.3.5)
50	0.2	0.5	0.3	0.7	20	11	2	0.00428	0.00421
100	0.2	0.5	0.3	0.7	20	11	2	0.00214	0.00211
150	0.2	0.5	0.3	0.7	20	11	2	0.00143	0.00140
200	0.2	0.5	0.3	0.7	20	11	2	0.00107	0.00105
500	0.2	0.5	0.3	0.7	20	11	2	0.000428	0.000421

Table 4.4: Comparison between conventional RRT and proposed RRT when $P_1 = 0.15, P_2 = 0.6, P_3 = 0.25$, $\pi = 0.8$, $\alpha = 20, \beta = 11, \delta = 2$, for varying sample sizes (n)

n	P_1	P_2	P_3	π	α	β	δ	conventional variance (eqn.2.5)	Proposed Variance (eqn.3.5)
50	0.15	0.6	0.25	0.8	20	11	2	0.00327	0.00321
100	0.15	0.6	0.25	0.8	20	11	2	0.00163	0.00161
150	0.15	0.6	0.25	0.8	20	11	2	0.00109	0.00107
200	0.15	0.6	0.25	0.8	20	11	2	0.00082	0.000803
500	0.15	0.6	0.25	0.8	20	11	2	0.000327	0.000321

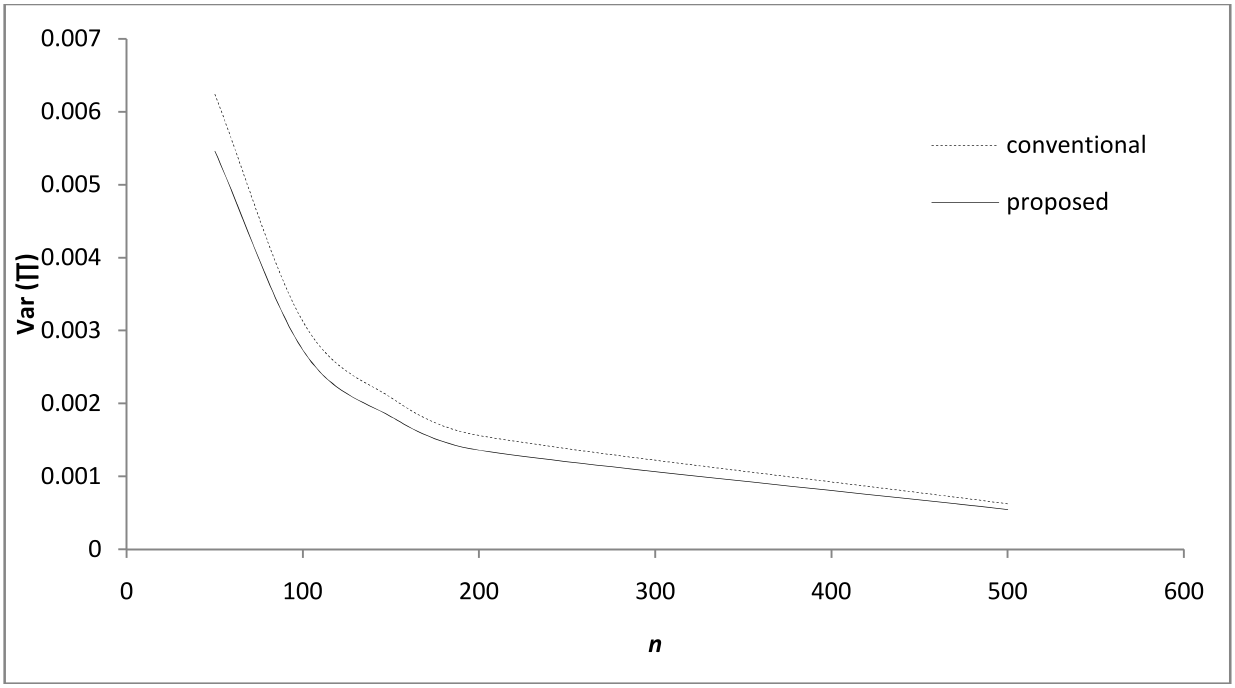

Figure 4.1:Comparison between conventional RRT and proposed RRT when $P_1 = 0.6, P_2 = 0.3, P_3 = 0.1$, $\pi = 0.5, \alpha = 20, \beta = 11, \delta = 2$, for varying sample sizes (n)

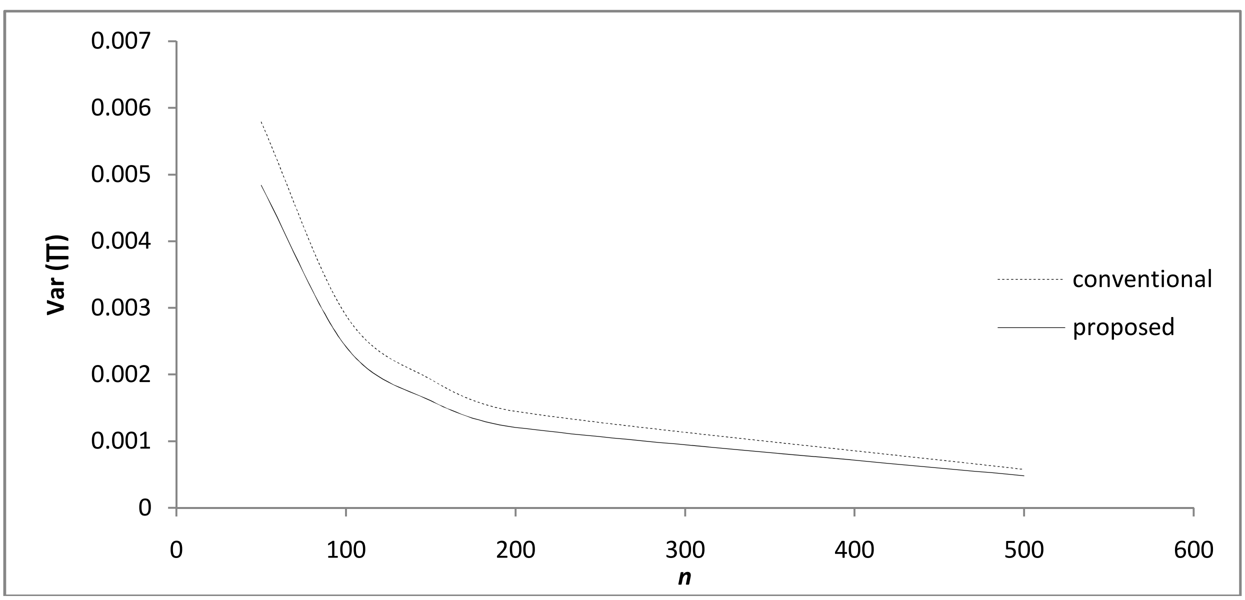

Figure 4.2:Comparison between conventional RRT and proposed RRT when $P_1 = 0.4, P_2 = 0.4, P_3 = 0.2$, $\pi = 0.4$, $\alpha = 20, \beta = 11, \delta = 2$, for varying sample sizes (n)

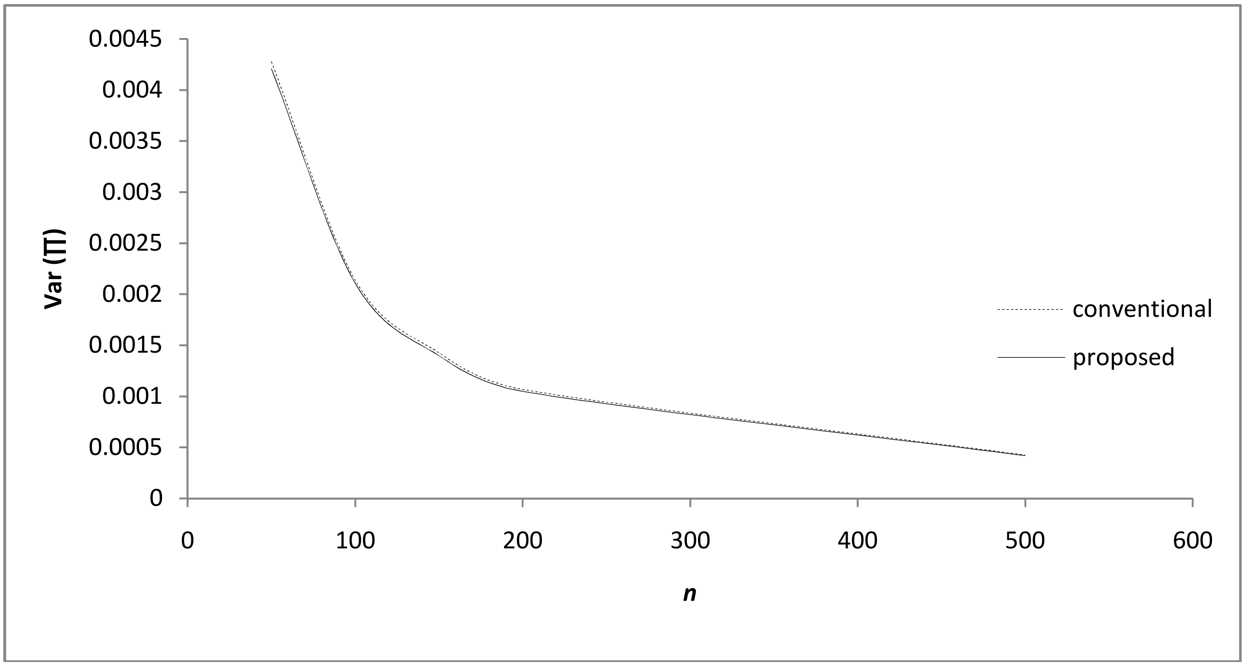

Figure 4.3:Comparison between conventional RRT and proposed RRT when $P_1 = 0.2, P_2 = 0.5, P_3 = 0.3,\quad \pi = 0.7,$ $\alpha = 20, \beta = 11, \delta = 2$, for varying sample sizes (n)

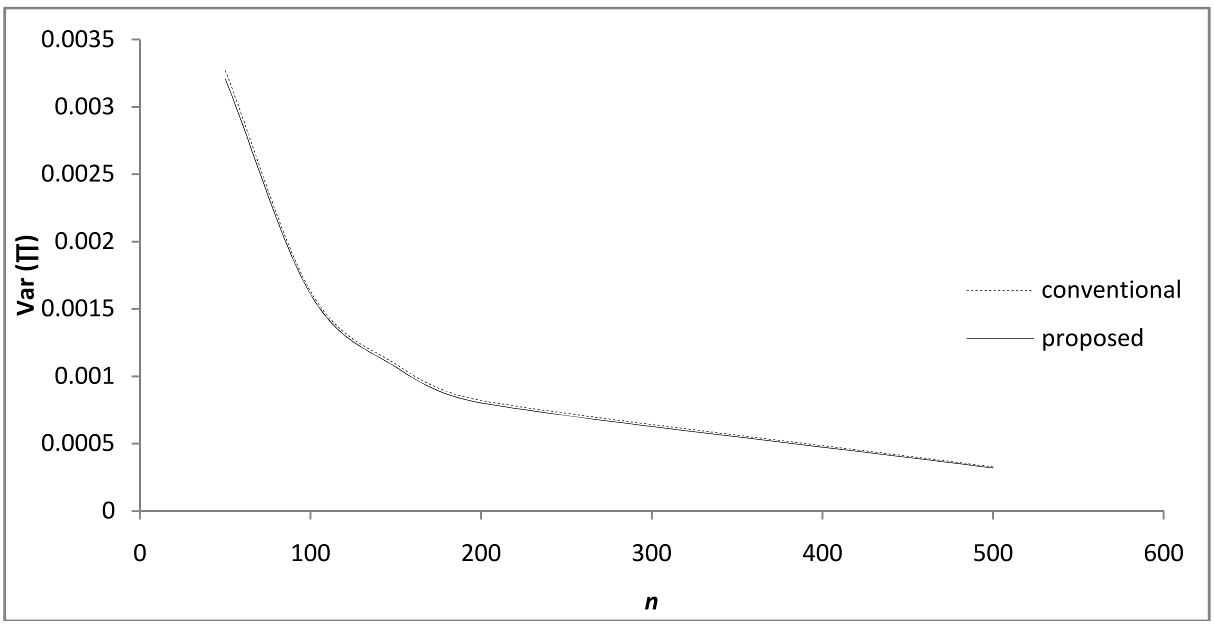

Figure 4.4:Comparison between conventional RRT and proposed RRT when $P_1 = 0.15, P_2 = 0.6, P_3 = 0.25,\quad \pi = 0.8$, $\alpha = 20, \beta = 11, \delta = 2$, for varying sample sizes (n)

Note:

Var(π) in the figures above represents variance for both conventional and proposed techniques as obtained in equations 2.5 and 3.5 respectively.

…….conventional variance (equation 2.5)

_____proposed variance (equation 3.5)

5.0 Conclusion

In this study, the review of the work of Hussain and Shabbir (2007) was presented.

The efficiency of our proposed Randomized Response Technique over that of the conventional one with respect to their variances was also verified by adopting the data used by Hussain and Shabbir (2007). It was evident in the results on Tables and Figures that the proposed technique is indeed more efficient than the conventional one.

References

[1] Christofides, T.C. (2003): *A generalized randomized response technique*. Metrika, 57, 195-200.

[2] Greenberg, B., Abul-Ela, A., Simmons, W. and Horvitz, D. (1969): *The unrelated question randomized response: theoretical framework*. J. Amer. Statist. Assoc, 64, 529-539. Zawar Hussain and Javid Shabbir 35

[3] Hussain, Z. and Shabbir, J. (2007): *Randomized use of Warner's Randomized Response Model*. InterStat: April # 7. http://interstat.statjournals.net/INDEX/Apr07.html

[4] Kim, J.M. and Warde, D.W. (2004): *A stratified Warner's Randomized Response Model*. J. Statist. Plann. Inference, 120(1-2), 155-165.

[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] Warner, S.L. (1965): *Randomized Response: a survey technique for eliminating evasive answer bias*. J. Amer. Statist. Assoc., 60, 63-69.