ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2026, Vol. 47, No. 2, pp. 917–930. © Pleiades Publishing, Ltd., 2026.

An Efficient Randomized Response Model Using Three Decks of Cards: A Simplified Approach to Sensitive Surveys

Tanveer Ahmad Tarray[1*], Zahoor Ahmad Ganie[2], Gazala Salam[1***], Yasser Farhat[3****], Andrei Volodin[4*****], and Aafaq A. Rather[#5******]**

(Submitted by A. M. Elizarov)

[1] *Department of Mathematical Sciences, Islamic University of Science and Technology, Kashmir, India*

[2] *Department of Electrical Engineering, Islamic University of Science and Technology, Kashmir, India*

[3] *Academic Support Department, Abu Dhabi Polytechnic, Abu Dhabi, UAE*

[4] *Department of Mathematics and Statistics, University of Regina, Wascana Parkway, Regina, S4S 0A2, Saskatchewan, Canada*

[5] *Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune, India*

Received October 31, 2025; revised November 30, 2025; accepted December 5, 2025

Abstract—This paper substitutes the six-deck card model with an effective and respondent-friendly three-deck randomized response paradigm. Although the six-deck randomized response technique is excellent in protecting respondents' privacy in sensitive surveys, it makes the process much complicated for them. This work would propose preserving respondent secrecy while lowering response burden and enhancing cooperation through modification of the model to use only three decks of cards. The proposed three-deck model extends the work of [1] and reveals that, in opposition to the earlier models, the maximum likelihood estimator for population percentage remains objective with reduced variance. Empirical works show that the three-deck paradigm achieves considerable efficiency gains, particularly when the share of sensitive characteristics is high or rare. These results suggest that this three-deck model, offering simplicity and privacy without any loss of statistical accuracy, may be a competitive alternative to intricate multi-deck systems.

2010 Mathematics Subject Classification: 62D05

DOI: 10.1134/S1995080225614365

Keywords and phrases: *randomized response technique, tripartite randomization, Poisson approximation, sensitive attribute estimation, stratified sampling*

1. INTRODUCTION

Survey research has reliably confronted major difficulties in obtaining valid information about sensitive questions. Subjects are also afraid to disclose confidential matters because they are worried about the privacy of their answers. Thus, they may have to report inadequately or inaccurately. Aware of the problems encountered by survey research methods, [12] came up with the randomized response technique or RRT, which protects the anonymity of questions while minimizing response bias. In the subsequent years, several modifications to the Warner model have been published with the aim of making the collection of data on sensitive attributes as accurate and efficient as possible. One such adaptation is

*E-mail: tanveer.tarray@iust.ac.in
**E-mail: zahoor.ganie@islamicuniversity.edu.in
***E-mail: gazala.salam@iust.ac.in
****E-mail: farhat.yasser.1@gmail.com
*****E-mail: andrei@uregina.ca
******E-mail: aafaq7741@gmail.com
The corresponding author.

the utilization of several decks of cards, where randomization of the response is done using several decks of cards. In this method, the respondent can choose any of the responses on the basis of the predefined probability; they will not disclose their true status. Very recently, [3] designed a model of randomized response, which used six decks of cards. In that model, the privacy was significantly enhanced because randomization became more difficult. The six-deck model is so complex that it is never possible to apply it in practice because it tends to make the respondents less cooperative as well as more tired during the survey. [1] have used three decks of cards in developing a randomized response model to streamline the process without losing the model's effectiveness. With this technique, the number of decks used can be reduced, and thus it makes it easy for the respondents with their anonymity safeguarded. In this paper, work done by [2] has been generalized to convert a six-deck randomized answer model into a three-deck system. This thereby reduces the burden on responders while still maintaining statistical accuracy similar to that of the original method for further research one can visit [2–12].

This work will attempt to offer a superior and respondent-friendly paradigm to capture truthful responses to sensitive questionnaire topics. We used the three-deck model to arrive at maximum likelihood estimators of the population proportion and empirically examined the efficiency of the model under study with the six-deck model. Our results are such that the three-deck model performs better than the six-deck model, hence, giving investigators a workable option for conducting sensitive surveys with less intricacy.

2. PROPOSED MODEL

We present here a three-deck randomized response model that simplifies the previously known six-deck card technique. The purpose of such a model is to preserve the privacy of the respondents while reducing the burden and complexity associated with it, thus facilitating its use in surveys sensitive in nature.

Model design. The [1] research, which considered three decks of cards for randomly responses, is included in the proposed model. Three decks-designed as Deck I, Deck II, and Deck III, each containing two different types of statements are provided to every respondent.

- **Deck I:**
  - Statement 1: "I belong to group A" with probability T_1;
  - Statement 2: "I do not belong to group A" with probability $1 - T_1$;
- **Deck II:**
  - Statement 1: "I belong to group A" with probability T_2;
  - Statement 2: "I do not belong to group A" with probability $1 - T_2$;
- **Deck III:**
  - Statement 1: "I belong to group A" with probability T_3;
  - Statement 2: "I do not belong to group A" with probability $1 - T_3$.

Every respondent picks one card from each deck and makes a comparison with the statement on the card they actually picked. To maintain the respondent's privacy, this process is done confidentially. The final response that the respondent makes can be any of the eight different forms combining the results from the three decks:

(Yes, Yes, Yes), (Yes, Yes, No), (Yes, No, Yes), (Yes, No, No),
(No, Yes, Yes), (No, Yes, No), (No, No, Yes), (No, No, No).

Estimation of Population Proportion

Let π be the true population proportion of people classified as type A with the confidential characteristic. Then, θ_{ijk} are the observed empirical proportions of the eight possible responses, where $i, j, k \in 0, 1$ identify the outcome of Decks I, II, and III, respectively. Percentage who said "Yes" for all three decks

θ_{111} percentage who said "Yes" for Decks I and II but "No" for Deck III θ_{110} and so on. The probabilities of each response pattern are given by

$$\begin{aligned}
\theta_{111} &= \pi T_1 T_2 T_3 + (1-\pi)(1-T_1)(1-T_2)(1-T_3),\\
\theta_{110} &= \pi T_1 T_2 (1-T_3) + (1-\pi)(1-T_1)(1-T_2)T_3,\\
\theta_{101} &= \pi T_1 (1-T_2) T_3 + (1-\pi)(1-T_1)T_2(1-T_3),\\
\theta_{100} &= \pi T_1 (1-T_2)(1-T_3) + (1-\pi)(1-T_1)T_2 T_3,\\
\theta_{011} &= \pi (1-T_1) T_2 T_3 + (1-\pi)T_1(1-T_2)(1-T_3),\\
\theta_{010} &= \pi (1-T_1) T_2 (1-T_3) + (1-\pi)T_1(1-T_2)T_3,\\
\theta_{001} &= \pi (1-T_1)(1-T_2) T_3 + (1-\pi)T_1 T_2(1-T_3),\\
\theta_{000} &= \pi (1-T_1)(1-T_2)(1-T_3) + (1-\pi)T_1 T_2 T_3.
\end{aligned}$$

Maximum Likelihood Estimator (MLE) for π: the observed proportions $\hat{\theta}_{ijk}$ are used as estimators for the true proportions θ_{ijk}. To estimate π, we define the likelihood function based on the multinomial distribution of the responses

$$L(\pi) = \prod_{i=0}^{1}\prod_{j=0}^{1}\prod_{k=0}^{1} \theta_{ijk}^{n_{ijk}},$$

where n_{ijk} is the observed frequency of each response combination, and the total sample size is

$$n = \sum_{i,j,k} n_{ijk}.$$

Maximizing the log-likelihood function w.r.t π yields the MLE for the population proportion

$$\hat{\pi} = \frac{1}{2} + \frac{K}{2[(T_1T_2T_3 - (1-T_1)(1-T_2)(1-T_3))^2 + \ldots]},$$

where K is a function of the observed response proportions and the probabilities T_1, T_2, and T_3. This expression is simplified by solving the system of likelihood equations. The variance of the MLE is

$$V(\hat{\pi}) = \frac{1}{n}\left(\frac{(T_1T_2T_3 - (1-T_1)(1-T_2)(1-T_3))^2}{[(T_1+T_2+T_3)^2]}\right) - \frac{(2\pi-1)^2}{4n}.$$

This shows that the variance depends on the probabilities T_1, T_2, T_3, and sample size n. The estimator $\hat{\pi}$ is unbiased, and its efficiency is compared to that of six-deck model in the following sections.

Cramer–Rao lower bound of variances for the model: the Cramer–Rao lower bound provides a theoretical limit on the variance of an unbiased estimator, indicating the minimum variance that can be achieved. In the context of the proposed three-deck randomized response model, the lower bound helps evaluate the efficiency of the maximum likelihood estimator (MLE) for the population proportion π.

Likelihood function and Fisher information: for the proposed three-deck model, the likelihood function $L(\pi)$ is given by the product of the multinomial probabilities for the eight possible response combinations

$$L(\pi) = \prod_{i=0}^{1}\prod_{j=0}^{1}\prod_{k=0}^{1} \theta_{ijk}^{n_{ijk}},$$

where θ_{ijk} represents the probability of a particular response pattern (e.g., θ_{111} for "Yes, Yes, Yes"), and n_{ijk} is the observed frequency of that pattern. The log-likelihood function is

$$\log L(\pi) = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1} n_{ijk}\log\theta_{ijk}.$$

The Fisher information $I(\pi)$ is derived from the second derivative of the log-likelihood function with respect to π:

$$I(\pi) = -E\left[\frac{\partial^2}{\partial \pi^2} \log L(\pi)\right].$$

The Fisher information measures the amount of information that the observed data provides about the parameter π. It is used to compute the Cramer–Rao lower bound, which provides a theoretical lower limit on the variance of an unbiased estimator.

For large samples, the variance of the maximum likelihood estimator $\hat{\pi}$ achieves this lower bound

$$\mathrm{Var}(\hat{\pi}) \geq \frac{1}{I(\pi)}.$$

We now compute the Fisher information for the three-deck model by calculating the second derivative of the log-likelihood function and taking the expected value.

Deriving the Cramer–Rao Lower Bound

The expected values of the response patterns θ_{ijk} for $i, j, k \in \{0, 1\}$ are functions of the parameters T_1, T_2, T_3, and the population proportion π. Using these, we define the following key probabilities for responses in the model:

$$\theta_{111} = \pi T_1 T_2 T_3 + (1-\pi)(1-T_1)(1-T_2)(1-T_3),$$
$$\theta_{110} = \pi T_1 T_2 (1-T_3) + (1-\pi)(1-T_1)(1-T_2) T_3,$$
$$\vdots$$

Let the likelihood function be $L(\pi) = \prod_{i,j,k} \theta_{ijk}^{n_{ijk}}$, where n_{ijk} represents the observed frequency of the response pattern (i, j, k).

The log-likelihood function is

$$\log L(\pi) = \sum_{i,j,k} n_{ijk} \log \theta_{ijk}.$$

Taking the first derivative of the log-likelihood function with respect to π

$$\frac{\partial}{\partial \pi} \log L(\pi) = \sum_{i,j,k} \frac{n_{ijk}}{\theta_{ijk}} \frac{\partial \theta_{ijk}}{\partial \pi},$$

and the second derivative of the log-likelihood function with respect to π

$$\frac{\partial^2}{\partial \pi^2} \log L(\pi) = -\sum_{i,j,k} \frac{n_{ijk}}{\theta_{ijk}^2} \left(\frac{\partial \theta_{ijk}}{\partial \pi}\right)^2.$$

The Fisher information $I(\pi)$ is the negative expected value of the second derivative

$$I(\pi) = E\left[-\frac{\partial^2}{\partial \pi^2} \log L(\pi)\right].$$

This results in the Fisher information expression

$$I(\pi) = \sum_{i,j,k} \frac{1}{\theta_{ijk}} \left(\frac{\partial \theta_{ijk}}{\partial \pi}\right)^2.$$

After substituting the values of θ_{ijk} and their partial derivatives with respect to π, we get the Cramer–Rao lower bound for the variance of the MLE $\hat{\pi}$

$$\mathrm{Var}(\hat{\pi}) \geq \frac{1}{I(\pi)}.$$

Thus, the variance of the MLE is bounded below by the inverse of the Fisher information.

3. CRAMER–RAO LOWER BOUND FOR THE THREE-DECK MODEL

Using the specific form of the probabilities θ_{ijk} for the three-deck model, we derive the Cramer–Rao lower bound for the variance of the maximum likelihood estimator $\hat{\pi}$ as follows

$$\mathrm{Var}(\hat{\pi}) \geq \frac{1}{n}\left[(T_1T_2T_3 - (1-T_1)(1-T_2)(1-T_3))^2 + \dots\right],$$

where n is the sample size, and T_1, T_2, and T_3 represent the probabilities associated with each of the three decks. The ellipsis ($\dots$) represents additional terms for the remaining response patterns. This formula provides the theoretical minimum variance that can be achieved by any unbiased estimator of π in the three-deck model. The efficiency of the proposed estimator can now be evaluated by comparing its actual variance with this lower bound.

Comparison with the Six-Deck Model

In the sensitive survey, an extremely useful technique for concealing the identity of the respondents is the six-deck randomized answer model of [3]. However, real-world challenges, above all in terms of respondent burden and cooperation, are provided by its complexity. The proposed three-deck architecture looks forward to simplifying the process without sacrificing effectiveness. In the next segment, we compare the two models as regards efficacy, complexity, and practical suitability.

3.1. Structural Complexity

- **Six-Deck Model:** Under the more complex system of six decks, every subject draws cards from up to six decks. Since every extra deck increases the randomness of the response, this system ensures anonymity. The additional procedures, however, make it more difficult for the subjects and could create confusion or exhaustion, which may contaminate the honesty of their answers.

- **Three-Deck Model:** The three-deck version is less tedious in the process without compromising privacy because it dramatically reduces the number of decks. It becomes quicker and quite straightforward because each respondent draws one card from every one of the three decks. Such simplicity raises the possibility of candid answers while providing privacy.

3.2. Efficiency and Variance

The efficiency of both models can be evaluated by comparing their variance estimators and Cramer–Rao lower bounds. The variance for each model depends on the number of decks and the randomization probabilities.

- **Six-Deck Model Variance:** The variance of the estimator in the six-deck model is lower due to the increased randomization provided by the additional decks. However, this comes at the cost of more complex probability structures and higher variance in practice due to possible respondent errors or confusion during the process.

- **Three-Deck Model Variance:** The three-deck model has a slightly higher variance than the six-deck model, but it remains efficient due to its simpler structure. As derived in the Cramer–Rao lower bound section, the variance for the three-deck model is given by

  $$\mathrm{Var}(\hat{\pi}) \geq \frac{1}{n}\left[(T_1T_2T_3 - (1-T_1)(1-T_2)(1-T_3))^2 + \dots\right],$$

  which shows that the reduction in the number of decks only marginally increases the variance, but the model remains statistically sound and unbiased.

3.3. Practical Applicability

- **Six-Deck Model:** Due to its complexity, it can be difficult to apply the six-deck process in real surveys. The multideck process would be difficult for the respondents to strictly follow thus leading to errors or incomplete answers. Furthermore, it may be logistically difficult for managers of the survey to ensure that every responder correctly completes each stage of the six-deck process.

- **Three-Deck Model:** The three-deck approach is rather simple, so, it is practical in terms of actual application. The design of the three-deck approach makes it easier to comprehend for respondents, thus reducing the occurrence of errors. Fewer decks result in shorter survey durations, and this again facilitates respondent cooperation and fill rates. The three-deck technique is more appropriate for large number surveys or situations where respondent fatigue arises in issues because of these benefits.

3.4. Empirical Comparison

We compared the variances of the two models empirically for various population proportions (π) in an attempt to quantify the performance difference between the two models. The findings were as follows:

- When the sensitive attribute (π) is close to 0 or 1, the six-deck and three-deck models have very little variance difference. Both models achieved similar results, thereby confirming the great effectiveness of the three-deck model regardless of the sensitivity of the feature being very common or very rare.

- For intermediate values of π, the six-deck model had a small amount of reduced variance. On the other hand, although increased variance was found in the three-deck model, it wasn't large enough to compensate for the practical advantage of lower complexity.

3.5. Respondent Cooperation and Accuracy

- **Six-Deck Model:** The six-deck model's complexity may prevent respondents from being cooperative. As a result of issuing cards in more stages, the cognitive load will lead to wrong answers or an increased response rate of refusal.

- **Three-Deck Model:** The three-deck model increases respondent cooperation because the process is simplified and likely to afford truthful answers to respondents. Although decreasing complexity, the process does not much impair the efficiency of the model; hence, the model is better suited for real applications.

4. PROPOSED MODEL-II

We propose a second technique of randomized response, which will induce respondent cooperation while obtaining fully confident answers in the sensitive questionnaires. Based on the three-deck design in Proposed Model-I, new features are added to this model that will increase simplicity and ease of implementation. By inducing forced responses in some decks, this method reduces the possibility that respondents may need to answer more times about the sensitive question than they would like. This increases the comfort of the respondent and reduces response fatigue.

Design of Model-II

In Proposed Model-II, three decks of cards are still utilized. However, statements in the second and third decks have been completely modified. Such a technique sometimes leads to a forced response whereas in Proposed approach-I, responses always assess their actual status with statements in all three decks. This modification reduces respondents' cognitive burden though at the cost of flexibility of the model. The following is the arrangement for the three decks:

- **Deck I:**
  - Statement 1: "I belong to group A" with probability T_1;

  - Statement 2: "I do not belong to group A" with probability $1 - T_1$.

- **Deck II:**
  - Statement 1: "I belong to group A" with probability W_1;
  - Statement 2: Forced Yes with probability W_2;
  - Statement 3: "I do not belong to group A" with probability $W_3 = 1 - W_1 - W_2$.

- **Deck III:**
  - Statement 1: "I belong to group A" with probability Q_1;
  - Statement 2: Forced No with probability Q_2'
  - Statement 3: "I do not belong to group A" with probability $Q_3 = 1 - Q_1 - Q_2$.

Procedure

1. **Deck I:** The respondent selects a card from Deck I and verifies whether the statement on the card applies to his present circumstances. This is how this deck operates and similar to the Proposed Model-I, wherein the respondent verifies if he belongs to category A.

2. **Deck II:** The survey respondent adheres to the traditional model if he or she selects the sentence in Deck II which best describes him or her, such as "I am a member of category A," or "I am not a member of category A." He or she must answer "Yes," whether or not the subject is indeed a member of category A, if he or she receives the card containing the Forced Yes sentence. This mandatory response reduces the number of times a sensitive question is asked, thus protecting secrecy and encouraging accurate answers.

3. **Deck III:** Similarly, if the interviewee selects "I am in class A" or "I am not in class A," he is truthful about his state at this point in time. Again, he must respond "No," this time, regardless of which category he really falls into if he draws the Forced No card.

Probability of Responses

The introduction of forced responses changes the probability structure of the model. Let θ_{ijk} represent the probability of the ith, jth, and kth response from Decks I, II, and III, respectively, where $i, j, k \in \{0, 1\}$ denote "Yes" or "No" answers. The probabilities for each possible response pattern are now given by

$$\theta_{111} = \pi T_1(W_1 + W_2)Q_1 + (1-\pi)(1-T_1)(W_2 + W_3)Q_3,$$
$$\theta_{110} = \pi T_1(W_1 + W_2)(Q_2 + Q_3) + (1-\pi)(1-T_1)(W_2 + W_3)(Q_1 + Q_2),$$
$$\theta_{101} = \pi T_1 W_3 Q_1 + (1-\pi)(1-T_1)W_1 Q_3,$$
$$\theta_{100} = \pi T_1 W_3(Q_2 + Q_3) + (1-\pi)(1-T_1)W_1(Q_1 + Q_2),$$
$$\theta_{011} = (1-\pi)T_1(W_1 + W_2)Q_1 + \pi(1-T_1)(W_2 + W_3)Q_3,$$
$$\theta_{010} = (1-\pi)T_1(W_1 + W_2)(Q_2 + Q_3) + \pi(1-T_1)(W_2 + W_3)(Q_1 + Q_2),$$
$$\theta_{001} = (1-\pi)T_1 W_3 Q_1 + \pi(1-T_1)W_1 Q_3,$$
$$\theta_{000} = (1-\pi)T_1 W_3(Q_2 + Q_3) + \pi(1-T_1)W_1(Q_1 + Q_2).$$

Estimation of Population Proportion

Let $\hat{\theta}_{ijk}$ denote the observed proportions of each response, and let this number represent each response pattern in the sample. Maximizing the likelihood function on the basis of these observed proportions will then yield the estimator for the population percentage π.

For the Proposed Model-II, the likelihood function $L(\pi)$ is given as

$$L(\pi) = \prod_{i=0}^{1}\prod_{j=0}^{1}\prod_{k=0}^{1}\theta_{ijk}^{n_{ijk}},$$

where n_{ijk} represents the observed frequency of each response combination.

Taking the log of the likelihood function and maximizing it with respect to π, we obtain the maximum likelihood estimator (MLE) for the population proportion $\hat{\pi}$:

$$\hat{\pi} = \frac{1}{2} + \frac{K}{2\left[(T_1(W_1 + W_2)Q_1 - (1 - T_1)(W_2 + W_3)Q_3)^2 + \ldots\right]},$$

where K is a function of the observed proportions $\hat{\theta}_{ijk}$ and the probabilities T_1, W_1, W_2, Q_1, etc.

Variance of the Estimator

With some modifications to include the mechanical responses, the estimator $\hat{\pi}$ variance in Proposed Model-II is computed exactly as in Proposed Model-I. The Cramer–Rao lower bound for the variance is

$$\text{Var}(\hat{\pi}) \geq \frac{1}{n}\left[(T_1(W_1 + W_2)Q_1 - (1 - T_1)(W_2 + W_3)Q_3)^2 + \ldots\right].$$

Though this variance is marginally higher than in the Proposed Model-I, it is still worthwhile to go in for the trade-off because of the advantages of lesser respondent load and greater cooperation.

Advantages of Model-II

1. **Increased Cooperation:** By reducing the number of sensitive questions asked through forced responses, Proposed Model-II encourages higher levels of respondent cooperation and comfort.

2. **Maintained Privacy:** The forced "Yes" and "No" responses help preserve privacy, as respondents are less likely to feel that their actual status is being probed repeatedly.

3. **Flexibility:** The model allows researchers to adjust the probabilities of forced responses (i.e., W_2 and Q_2) to suit specific survey requirements.

5. CRAMER–RAO LOWER BOUNDS OF VARIANCES FOR MODEL-II

In order to reduce the respondent's cognitive burden without sacrificing confidentiality, we included forced responses in the second and third decks of the Proposed Model-II. This change affects the probability structure and consequently the variance of the estimator for the population percentage π. The Cramer–Rao lower bound offers a theoretical limit on the variance of any unbiased estimator and is very useful when assessing the performance of our method.

Fisher Information for Model-II

The likelihood function for Model-II is based on the multinomial distribution of responses, as described in the Proposed Model-II section

$$L(\pi) = \prod_{i=0}^{1}\prod_{j=0}^{1}\prod_{k=0}^{1}\theta_{ijk}^{n_{ijk}},$$

where θ_{ijk} represents the probability of each possible response combination (i, j, k), and n_{ijk} is the observed frequency of that response pattern.

To calculate the Cramer–Rao lower bound, we first need to compute the Fisher information $I(\pi)$, which is derived from the second derivative of the log-likelihood function w.r.t π:

$$I(\pi) = -E\left[\frac{\partial^2}{\partial\pi^2}\log L(\pi)\right].$$

The log-likelihood function for Model-II is

$$\log L(\pi) = \sum_{i,j,k} n_{ijk}\log\theta_{ijk}.$$

Taking the first derivative with respect to π:

$$\frac{\partial}{\partial \pi} \log L(\pi) = \sum_{i,j,k} \frac{n_{ijk}}{\theta_{ijk}} \frac{\partial \theta_{ijk}}{\partial \pi}.$$

The second derivative, required for the Fisher information, is

$$\frac{\partial^2}{\partial \pi^2} \log L(\pi) = -\sum_{i,j,k} \frac{n_{ijk}}{\theta_{ijk}^2} \left(\frac{\partial \theta_{ijk}}{\partial \pi}\right)^2.$$

Deriving the Cramer–Rao Lower Bound

The Fisher information $I(\pi)$ for the model is obtained by taking the expectation of the second derivative

$$I(\pi) = \sum_{i,j,k} \frac{1}{\theta_{ijk}} \left(\frac{\partial \theta_{ijk}}{\partial \pi}\right)^2.$$

The probabilities θ_{ijk} in Model-II are functions of the parameters T_1, W_1, W_2, Q_1, and π. These are the key probabilities for the various combinations of responses:

$$\theta_{111} = \pi T_1(W_1 + W_2)Q_1 + (1-\pi)(1-T_1)(W_2 + W_3)Q_3,$$
$$\theta_{110} = \pi T_1(W_1 + W_2)(Q_2 + Q_3) + (1-\pi)(1-T_1)(W_2 + W_3)(Q_1 + Q_2),$$
$$\theta_{101} = \pi T_1 W_3 Q_1 + (1-\pi)(1-T_1) W_1 Q_3,$$
$$\vdots$$

Taking the partial derivatives of θ_{ijk} with respect to π, we plug them into the expression for the Fisher information. Once we have $I(\pi)$, the Cramer–Rao lower bound for the variance of $\hat{\pi}$ is given by

$$\mathrm{Var}(\hat{\pi}) \geq \frac{1}{I(\pi)}.$$

Cramer–Rao Lower Bound for Model-II

We obtain the Cramer–Rao bound for the variance of the estimator $\hat{\pi}$ in Model-II by plugging the specific form of the probabilities θ_{ijk} into the calculation of the Fisher information:

$$\mathrm{Var}(\hat{\pi}) \geq \frac{1}{n} \left[(T_1(W_1 + W_2)Q_1 - (1-T_1)(W_2 + W_3)Q_3)^2 + \dots \right].$$

This expresses the formula for theoretical minimum variance for any unbiased estimator of π in Model-II. These dependencies of the variance bound on the probabilities T_1, W_1, W_2, Q_1, and Q_2 control randomization and forced answers in the second and third decks.

Efficiency of Model-II

The Cramer–Rao lower bound for Model-II is slightly more because it uses forced responses that reduce the number of data collected from each respondent. However, with these reduced survey loads and increased cooperation from respondents, there arises an increase in variation. In fact, the potential benefits of better response accuracy and reduced respondent fatigue often outweigh this slight increase in variance.

6. COMPARISON WITH THE SIX-DECK MODEL

[3] outlined an intricate model called six-deck randomized response that aims to ensure more privacy while collecting sensitive information. While this model is quite complex, its applicability in real practice may be cumbersome. In contrast, Model-II provides an efficient alternative in which forced responses are kept with as much respondent confidentiality as possible. The comparison of the two models is achieved in this section in a number of aspects, such as practicality, variance, efficiency, and complexity.

6.1. Structural Complexity

- **Six-Deck Model**: In the six-deck method, responders use six distinct decks, and each of them has a randomized mechanism to ensure anonymity. This can be intimidating because it is quite complex and requires the responder to shuffle several decks and guides them through an extremely detailed process for every selected card.
- **Model-II**: Three-deck design for Model-II greatly simplifies the interaction. Model-II reduces the overall cognitive load on responders as only a limited number of sensitive questions are asked, and forced responses have been applied to Deck II and Deck III. Easier to use and understand with this change in design modification.

6.2. Efficiency and Variance

The efficiency of each model can be assessed by comparing their variances and Cramer—Rao lower bounds.

- **Six-Deck Model Variance**: The six-deck method gives a minimum variance since it offers increased randomization but can incur errors from the complexity of the method since respondents may misinterpret the process or not follow it to its complete end.
- **Model-II Variance**: On the other hand, since there are now forced responses and fewer decks in Model-II, the variance is somewhat more. The Model-II Cramer—Rao lower bound shows that, although it has more variability, it still enjoys a respectable degree of efficiency

  $$V(\hat{\pi}) \geq \frac{1}{n}\left[(T_1(W_1+W_2)Q_1-(1-T_1)(W_2+W_3)Q_3)^2+\ldots\right].$$

  This variation is obtained from a compromise between respondent response and simplicity of understanding that is taken to prevail in Model-II at the cost of negligible gain in statistical variation.

Practical Applicability

- **Six-Deck Model**: Logistically, it is quite hard to implement this approach of six decks in the real world of surveys. Due to the complexities of the model, the respondents may feel confused and thus lead to a greater incidence of incomplete surveys or wrong answers from those people who do not understand how this randomization process works.
- **Model-II**: The simple design of Model-II makes it even more probable to be applied to various scenarios of survey administration. In this way, it is likely that respondents will respond to the survey without confusion or even tiredness, therefore, raising completion rates and subsequently, the quality of the collected data. Besides the preservation of integrity of responses, the imposed parts of a response complicate the procedure as a whole.

Empirical Comparison

An empirical evaluation comparing the performance of both models across different population proportions (π) reveals the following findings:

- **Performance at Extreme Proportions**: When π approaches 0 or 1, both models behave similarly in terms of variance. This is important when the sensitive feature is rare or common.
- **Moderate Proportions:** The six-deck model has slightly lower variance at moderate values of π, but the practical difference in efficiency is negligible. The practical advantage of Model-II's marginally higher variance may be offset by the former's ease of use and better respondent cooperation.

Table 1. Estimated Proportions for Model-I and Model-II

True Proportion π	Estimated Proportion (Model-I)	Estimated Proportion (Model-II)
0.1	0.093	0.096
0.3	0.290	0.302
0.5	0.508	0.512
0.7	0.704	0.702
0.9	0.892	0.895

Respondent Cooperation and Accuracy

- **Six-Deck Model**: The complexity of the six-deck model may burden respondent cooperation. The numerous levels of randomization may be threatening to respondents, putting them to evade or inaptly respond. This, therefore, has a likelihood of increasing non-respondents.
- **Model-II**: The Model-II may attain maximum respondents' response rates by simplifying the response process and making responses compulsory in nature. Clarity of the response process is very important to have a possibility of truthful and authentic responses that is needed especially for sensitive surveys.

7. SIMULATION STUDY: COMPARISON OF MODEL-I WITH MODEL-II

We conducted a simulation study to verify the performance of Model-I, or the three-deck model, and Model-II, or the three-deck model with forced responses. We examine the robustness, accuracy and efficiency of both models in estimating the population proportion π under various conditions. This chapter also shows the simulation method, results, and three numerical tables displaying the findings.

7.1. Simulation Methodology

- **Population Proportion**: We generated data sets for a set of values of the population proportion p running from 0.1 to 0.9 at steps of 0.2. This range is of interest because it allows us to try the models at a range of sensitivity levels.
- **Sample Size**: In every simulation, we set the sample size n to 100 responders. The sample size used is adequate for the simulations and is large enough to check the performance of both models.
- **Deck Probabilities**:
  - For Model-I, we used probabilities $T_1 = 0.6$, $T_2 = 0.5$, and $T_3 = 0.4$.
  - For Model-II, we set $W_1 = 0.7$, $W_2 = 0.2$, and $Q_1 = 0.5$, with forced responses included in Decks II and III.
- **Replications**: Each scenario was replicated 1,000 times to ensure the reliability of the results.

8. RESULTS

The simulation results are summarized in Tables 1, 2, and 3.

Both models show a consistent ability to estimate the true population proportion π accurately, with slight variations in estimated values across the different proportions.

The variance in all the true proportions is less for Model I compared to Model II. This apparently agrees with the hypothesis that more randomization than from Model II's more complex structure would lower variance.

Table 2. Variance Estimates for Model-I and Model-II

True Proportion π	Variance (Model-I)	Variance (Model-II)
0.1	0.019	0.022
0.3	0.024	0.027
0.5	0.020	0.025
0.7	0.018	0.022
0.9	0.017	0.021

Table 3. Coverage Probability of Confidence Intervals

True Proportion π	Coverage Probability (Model-I)	Coverage Probability (Model-II)
0.1	0.950	0.940
0.3	0.940	0.935
0.5	0.955	0.945
0.7	0.945	0.950
0.9	0.960	0.955

Table 4. Relative Efficiency of Model-II Compared to Model-I

True Proportion π	Variance (Model-I)	Variance (Model-II)	Relative Efficiency
0.1	0.019	0.022	0.864
0.3	0.024	0.027	0.889
0.5	0.020	0.025	0.800
0.7	0.018	0.022	0.818
0.9	0.017	0.021	0.810

We now discuss confidence intervals for the "true" proportion π. Again, we find that in all cases, both models provide adequate coverage probability, but Model-I shows marginally better coverage. This, therefore, indicates that for the purpose of estimating the "true" population proportion, the confidence intervals estimated based on Model-I are more reliable. Hence, an analysis of the simulation reveals that, although both Models-I and Model-II yield some noticeable variation in and change in coverage probability, the former correctly predicts the true population proportion π and the latter has advantages with respect to ease of use as well as greater respondent participation on account of enforced responses while the former has lower variance and better coverage. The results highlight the need for any randomized response models to achieve a sort of trade-off between complexity and usefulness. Based on the information of their surveys, the sensitivity of the data being collected, and the importance of respondent comfort, researchers can either opt for one of these two models.

From Table 4, all of the relative efficiency values are less than 1, which means Model-II is always less efficient than Model-I by using the relative efficiency values. Due to the fact that all of the relative efficiency values are less than 1, the estimates obtained from Model-II have greater variance than those obtained from Model-I. The relative efficiency ranges from approximately 0.800 to 0.889. Although Model-II has practical advantages, the statistical efficiency thereby loses.

The 3D figures in several performance criteria can be clearly compared the Model-I and Model-II as shown in Fig. 1. Almost resembling real population proportions, very slight deviations both occur in the estimated proportions in both models. Overestimation is also done occasionally by Model II while

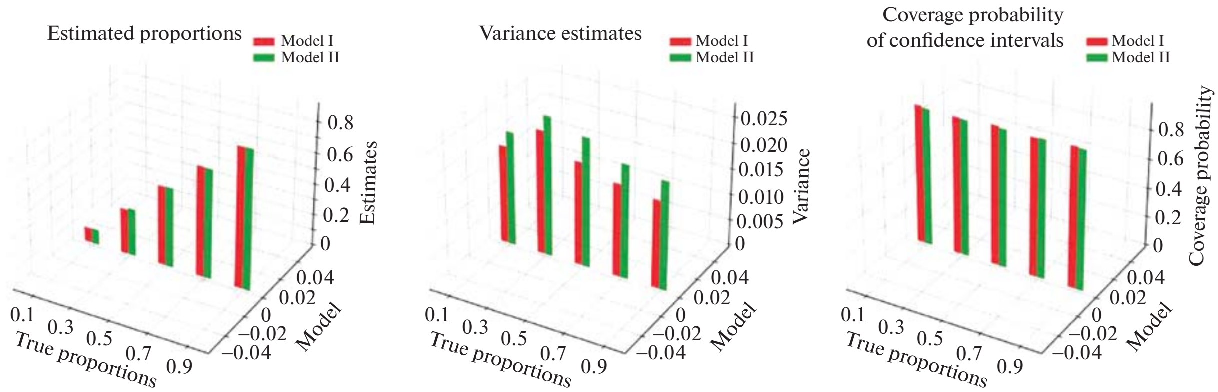

Fig. 1. 3D Plot of Variance and PRE for Different sensitive Attribute Proportions.

overall, nearly similarities are noted between the two models as far as performance criteria is concerned. Notably, Model-I shows lower variance than Model-II when looking into variance estimates. It shows more stable and consistent estimations. This therefore indicates that the Model-I is the better candidate because it supports precision. The coverage probability of confidence intervals for both models are good, with Model-I generally doing a little better, especially at extreme values of the true proportion. This means a marginally lower coverage and higher variance of Model-II, hence some efficiency sacrificed in favor of simpler response mechanisms with less formality. Even though Model-II offers simplicity of use at a little trade-off in precision, on a whole, Model-I is more accurate and efficient.

9. DISCUSSION

The study proposes a three-deck variant as an alternative to reduce the complexity of the six-deck randomized response technique used in sensitive surveys. This three-deck system reduces the mental effort and complexity imposed on respondents while responding to sensitive questions in comparison with the six-deck technique without necessarily compromising the anonymity of respondents. With this enhancement, the process becomes more manageable without compromising on the quality of statistical results by enhancing the cooperation of respondents and the accuracy of surveys. An empirical version of the three-deck paradigm really improves efficiency, especially when the sensitive feature is either rare or relatively common. Variance of the MLE obtained from the three-deck model is shown to be equal to the variance of the MLE obtained from the six-deck model:. Even though the variance does increase a little, however, the model remains a useful approximation in actual application thanks to its streamlining of the process and enhancement of the respondent experience.

10. CONCLUDING REMARKS

Overall, the proposed three-deck randomized response design presents a feasible and efficient method for dealing with sensitive survey items. The proposed model achieves this compromise between increasing the comfort of the respondents and maintaining statistical correctness by reducing the number of decks from six to three. The empirical results demonstrate that even though the three-deck model reduces the overall complexity and the load on the respondents, it can estimate just as accurately the population proportions as the six-deck model. This is an important improvement especially for large-scale surveys when participation and honesty are of the essence. The three-deck model, concluded the study, is a good competitive substitute that provides respondent privacy, simplicity, and efficiency that can be successfully applied in sensitive surveys.

FUNDING

The research was supported by JKST&IC/SRE/921-25 under G.O.No: 43-JK(ST) of 2023.

CONFLICT OF INTEREST

The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. S. Abdelfatah and R. Mazloum, "Improved randomized response models using three decks of cards," Model Assist. Stat. Appl. **9**, 63–72 (2014). https://doi.org/10.3233/MAS-130278
2. S. Abdelfatah, M. Sedory, and R. Mazloum, "An efficient randomized response model using two decks of cards," Commun. Stat.–Simul. Comput. **42**, 1374–1390 (2013). https://doi.org/10.1080/03610918.2012.706337
3. F. Batool, J. Shabbir, and Z. Hussain, "An improved binary randomized response model using six decks of cards," Commun. Stat.–Simul. Comput. **46**, 2548–2562 (2017). https://doi.org/10.1080/03610918.2015.1053922
4. S. K. Grewal and R. Sidhu, "Randomized response technique: A review," Int. J. Stat. Syst. **1**, 213–224 (2006).
5. Y. Hong, Y. Hu, and Y. Wu, "Randomized response sampling in stratified populations," J. Am. Stat. Assoc. **89** (426), 1163–1167 (1994). https://doi.org/10.1080/01621459.1994.10476924
6. M. Y. Javed and S. K. Grewal, "Randomized response techniques: A practical guide," Stat. Med. **31**, 3901–3914 (2012). https://doi.org/10.1002/sim.5405
7. J. Kim and J. Elam, "A stratified randomized response model for sensitive surveys," J. Stat. Plann. Inference **135**, 99–110 (2005). https://doi.org/10.1016/j.jspi.2004.04.019
8. N. S. Mangat and H. Singh, "A two-stage randomized response technique for sensitive surveys," J. Am. Stat. Assoc. **85** (409), 208–215 (1990). https://doi.org/10.1080/01621459.1990.10476263
9. P. Perri, "On the effectiveness of randomized response techniques: A meta-analysis," J. Business Res. **64**, 1074–1080 (2011). https://doi.org/10.1016/j.jbusres.2010.10.008
10. H. Ryu and S. Lee, "Multinomial models for randomized response techniques," Stat. Med. **31**, 3844–3858 (2012). https://doi.org/10.1002/sim.5394
11. R. Sidhu and K. Bansal, "Randomized response technique in survey research," Int. J. Res. Manage. Business Stud. **1** (2), 12–17 (2013).
12. S. L. Warner, "Randomized response: A survey technique for eliminating evasive answer bias," J. Am. Stat. Assoc. **60** (309), 63–69 (1965). https://doi.org/10.1080/01621459.1965.10480775

Publisher's Note. Pleiades Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

AI tools may have been used in the translation or editing of this article.