SCOPUA Journal of Applied Statistical Research (JASR)

Vol.2, Issue 2, June 2026
DOI: https://doi.org/10.64060/JASR2V2i3

On New Quantitative Randomized Response Techniques and a Weighted Privacy–Efficiency Measure

Shoaib Iqbal [1], Zawar Hussain [1*]

[1]Department of Statistics, Faculty of Computing, The Islamia University of Bahawalpur 63100, Pakistan

* Corresponding Email: zhlangah@yahoo.com

Received: 03 March 2026 / Revised: 01 May 2026 / Accepted: 05 May 2026 / Published online: 15 May 2026

ABSTRACT

Randomized response techniques are widely used for collecting information on sensitive variables while preserving respondent privacy. In the context of quantitative data, existing models are typically based on either additive or multiplicative scrambling mechanisms, which limit their flexibility in balancing estimation efficiency and privacy protection. This paper proposes a generalized class of quantitative randomized response models that integrates both additive and multiplicative components within a unified framework. From this formulation, three specific models (M_1 - M_3) are developed, each representing a distinct configuration of the scrambling process and corresponding to different efficiency–privacy trade-offs. In addition, a weighted measure combining efficiency and privacy is introduced to facilitate a unified comparison of competing models. The proposed measure provides a flexible and interpretable criterion for model evaluation and supports consistent ranking under varying preference weights. Theoretical properties of the estimators are derived, and comparative analysis demonstrates the improved performance of the proposed models over existing techniques under a range of parameter settings. The proposed framework enhances the adaptability of randomized response methods and offers a practical tool for selecting appropriate models in sensitive survey applications.

Keywords*:*Randomized response models; Scrambling models; Estimation of mean; Relative efficiency; Privacy protection

1. Introduction

Amongst the indirect questioning techniques including item count technique by Miller [1], Nominative technique by Sirken [2], List experiment by Raghavarao and Federer [3], the randomized response technique by Warner [4] have long been recognized as an effective tool for collecting reliable data on sensitive characteristics while preserving respondent confidentiality. Since the seminal contribution of Warner [4], a wide range of models have been developed to address both qualitative and quantitative survey settings. Some of the important developments may include Warner [5], Greenberg et al. [6], Horvitz et al. [7], Kuk [8], Mangat [9], Christofides [10], Kim and Warde [11], Arnab and Singh [12], Gupta et al. [13], Eichhorn and Hayre [14]. A detailed review on qualitative and quantitative randomized response models is available in Lee et al [15]. The interested readers may refer to Lee at al. [15] and the references cited therein.

In the context of quantitative variables, these techniques aim to estimate population parameters, particularly the mean, without requiring respondents to disclose their true values directly. Most existing quantitative RRT models are built upon either additive or multiplicative scrambling mechanisms. Additive approaches introduce random noise independent of the sensitive variable, whereas multiplicative schemes distort the reported value through scaling. While both strategies offer a degree of privacy protection, they often involve a delicate trade-off between estimation efficiency and confidentiality, and their structural forms remain somewhat restrictive. In particular, the literature lacks a comprehensive framework that systematically integrates these two mechanisms.

In addition to model development, considerable attention has been given to the evaluation of randomized response techniques in terms of efficiency and privacy. Traditional comparisons are often based on variance as a measure of efficiency, while privacy protection is assessed through various ad hoc or model-specific indices. More recently, Yan et al.

[16], Gupta et al. [17] and Azeem [18] measures have been proposed, which allow the incorporation of user-defined preferences. Despite their usefulness, these measures exhibit certain limitations. In particular, they often lack analytical transparency, making it difficult to characterize conditions under which one model dominates another. Moreover, the resulting rankings may not be stable across different weight choices, which complicates interpretation in practical applications. These limitations highlight the need for a more robust and interpretable combined measure that can support consistent decision-making.

The present study addresses these limitations by proposing a generalized class of quantitative randomized response models that combines additive and multiplicative scrambling within a unified formulation. This perspective allows the construction of multiple model configurations under a common structure. In particular, three specific models ($M_1 - M_3$) are developed, each representing a distinct balance between privacy protection and estimation efficiency. Rather than being arbitrary variations, these models correspond to different operational regimes within the proposed framework.

In addition, this paper introduces a weighted measure of privacy and efficiency, denoted by ψ, which provides a unified basis for comparing competing models. Unlike existing measures, the proposed formulation facilitates a more transparent interpretation of trade-offs and supports consistent ranking across different preference settings. The analytical properties of this measure are also examined to highlight its practical relevance.

Overall, this study contributes to the literature by developing a flexible modeling framework and a coherent evaluation criterion for quantitative randomized response techniques. The proposed approach not only clarifies the relationships among different scrambling strategies but also enhances the decision-making process in selecting an appropriate model for sensitive surveys.

The remainder of the paper is organized as follows. Section 2 introduces the generalized scrambling framework and presents the proposed models. Section 3 derives the estimators and their properties. Section 4 discusses the weighted measure and its characteristics. Section 5 provides comparative analysis, and Section 6 concludes the paper.

02. Generalized Scrambling Framework

In this section, we introduce a generalized framework for constructing quantitative randomized response models that integrates both additive and multiplicative scrambling mechanisms.

Let Y_i denote the true value of a sensitive quantitative variable for the i^{th} respondent, with mean $\mu_y = E(Y_i)$ and variance $\sigma_y^2 = Var(Y_i)$. To protect respondent privacy, the observed response is generated through a randomized mechanism involving independent scrambling variable. Let S_i be a scrambling variable such that $E(S_i) = \mu_s = 0$, $Var(S_i) = \sigma_s^2$. Assume that S_i and Y_i are mutually independent. We define the generalized scrambling mechanism as:

$$R_i = Y_i + C_i Y_i S_i, \tag{1}$$

where C_i is a random variable representing random coefficient mechanism. Using the assumptions above, and applying expectation on (1), we obtain the expected response as:

$$E(R_i) = E(Y_i + C_i Y_i S_i) = \mu_y + E(C_i)\mu_y \mu_s$$

$$E(R_i) = \mu_y + E(C_i)\mu_y(0) = \mu_y. \tag{2}$$

Using (2), the unbiased moment estimator of μ_y may be defined as

$$\hat{\mu}_y = \bar{R} = \frac{1}{n}\sum_{i=1}^{n} R_i \,. \tag{3}$$

The variance of the moment estimator $\hat{\mu}_y$, in (3), may be derived as follows. From (1), we have

$$Var(R_i) = Var(Y_i + C_i Y_i S_i) = Var(Y_i) + Var(C_i Y_i S_i) + 2Cov(Y_i, C_i Y_i S_i).$$

Since $E(S_i) = 0$, we have $Cov(Y_i, Y_i S_i) = 0$. This implies $Cov(Y_i, C_i Y_i S_i) = 0$. Thus, we get

$$Var(R_i) = \sigma_y^2 + Var(C_i Y_i S_i). \tag{4}$$

Now

$$Var(C_i Y_i S_i) = E(C_i^2)Var(Y_i S_i) = E(C_i^2)E(Y_i^2 S_i^2) = E(C_i^2)E(Y_i^2)E(S_i^2).$$

$$Var(C_i Y_i S_i) = E(C_i^2)(\mu_y^2 + \sigma_y^2)\sigma_s^2.$$

Substituting the above result in (4), we get

$$Var(R_i) = \sigma_y^2 + E(C_i^2)(\mu_y^2 + \sigma_y^2)\sigma_s^2 . \quad (5)$$

Finally,

$$Var(\hat{\mu}_y) = \frac{1}{n} Var(R_i) = \frac{1}{n}\left\{\sigma_y^2 + E(C_i^2)(\mu_y^2 + \sigma_y^2)\sigma_s^2\right\} . \quad (6)$$

Also note that $Var(\hat{\mu}_y) = \frac{1}{n}\left\{\sigma_y^2 + E(C_i^2)(\mu_y^2 + \sigma_y^2)\sigma_s^2\right\} = \frac{1}{n}\left\{\sigma_y^2 + KE(C_i^2)\right\}$, where $K = (\mu_y^2 + \sigma_y^2)\sigma_s^2$. Thus, the efficiency of the estimator depends entirely on $E(C_i^2)$. The estimator $\bar{R}$ is more efficient when $E(C_i^2)$ is minimized. To quantify the privacy protection, we define Yan et al. [16] privacy measure as

$$\nabla = E(R_i - Y_i)^2 = E(C_i Y_i S_i) = E(C_i^2)(\mu_y^2 + \sigma_y^2)\sigma_s^2 . \quad (7)$$

So, from (5) and (7), we can write

$$Var(R_i) = \sigma_y^2 + \nabla . \quad (8)$$

The expression in (8) shows that an increase in privacy (higher ∇) leads to a proportional increase in variance. Hence, there exists an inherent efficiency–privacy trade-off governed by $E(C_i^2)$. Three special cases of this generalized framework will be discussed in section 5.

3. Some Existing Comparison Measures

This section discusses the available measures of privacy and efficiency commonly used for comparing quantitative randomized response models.

3.1 The Relative Efficiency Measure

Relative efficiency is one of the most widely used methods for performance comparison and has been applied by nearly all researchers in evaluating randomized response models. Let Model *S* be the model of interest and Model *O* is the reference model used for comparison. Denote the mean squared error (*MSE*) of the mean estimator under Model *S* and Model *O* as MSE_S and MSE_O, respectively. The relative efficiency is then calculated as:

$$E = \frac{MSE_O}{MSE_s} .$$

Since the MSE is always positive, $E \geq 0$. If $E \geq 1$, it indicates that Model *S* is more efficient than Model *O*.

3.2 The Yan et al. [16] Measure of Privacy

Yan et al. [16] proposed a metric to assess the privacy level of a given quantitative model, denoted by ∇ and defined as:

$$\nabla = (Z - Y)^2$$

For any randomized response model, $\nabla \geq 0$, with higher values indicating greater privacy protection. The metric developed by Yan et al. [16] emphasizes privacy protection exclusively.

3.3 The Gupta et al. [17] Unified Measure of privacy and efficiency

Separate measures for quantifying privacy level and efficiency have been used by different authors prior to the publication of the work by Gupta et al. [17]. To make simultaneous use of privacy and efficiency, Gupta et al. [17] proposed a combined metric that integrates both privacy and efficiency, as outlined below:

$$\delta = \frac{MSE}{\nabla} .$$

The value of δ ranges from 0 to ∞. A smaller δ reflects lower mean squared error, a higher privacy level, or both. This measure provides an overall assessment of a randomized response model as a single δ value. However, it does not assign relative weights to privacy or efficiency, which can limit its usefulness.

3.4 The Azeem [18] Weighted Privacy Measure

Let ∇_S and ∇_O be the Yan et al. [16] measure of privacy protection for Model *S* and Model *O*, respectively. Then define

$$P = \frac{\nabla_S}{\nabla_O} .$$

Since ∇_S is in the numerator, so a higher value of P indicates higher privacy protection for Model *S* than Model *O*. If the two models under consideration offer equal privacy protection, then $P = 1$. Azeem [18] defined weighted privacy measure as

$$\phi = \left(\frac{w_1 E + w_2 P}{w_1 + w_2} \right). \tag{9}$$

The proposed measure, ϕ, given in (9), ranges from 0 to ∞. A ϕ value greater than 1 signifies that Model *S* outperforms Model *O*. Conversely, if $0 < \phi < 1$, it indicates that Model *S* is less effective than Model *O* overall.

4. The Proposed Weighted Measure of Privacy

In an effort to overcome the inadequacies of existing measures, we introduce the weighted measure, denoted by ψ, as:

$$\psi = P^{w_1} \times E^{w_2}, \tag{10}$$

where E and P are as defined above and w_1, w_2 are non-negative weights such that $w_1 + w_2 = 1$.

The weighted measure defined in (10), is an inclusive and unified way of evaluating the randomized response models based on whether they combine efficiency and privacy as a single factor. Unlike the existing measures, which either used only efficiency or only privacy, the proposed measure strikes a balance between the two based on its definition in terms of being a weighted geometric mean of the two. Another striking strength of the proposed measure is its ease of reading: the sign of $\log(\psi)$ clearly indicates whether the proposed model provides improvement, equal, or worse performance than the competing model. The ψ measure not only generalizes earlier approaches but also extends them, delivering a more powerful, flexible, and helpful practical tool for the comparative evaluation of randomized response models. Moreover, the ψ measure is a versatile measure. With the use of appropriate weights given to efficiency and privacy, researchers can choose to use the measure for purposes specific to their research. For a survey dealing with very sensitive information, for instance, more importance can be attributed to privacy but, for reliability or industrial purposes, efficiency takes precedence. This weighting flexibility guarantees that the ψ measure is highly flexible and applicable in a wide range of research environments, thus guaranteeing its utility in theoretical work as well as in practice.

The proposed ψ measure has several distinct advantages:

1. *Balanced consideration*: Unlike previous methods, which focused on either relative efficiency *E* in isolation or privacy *P*, the ψ measure addresses them at the same time, resulting in balanced evaluation.
2. *Flexibility*: The researchers can tailor the measurement by altering w_1 and w_2. For instance, in medical questionnaires, privacy can be weighted heavily, and in industrial reliability, efficiency.
3. *Geometric mean formula* – The use of a geometric mean avoids either dimension from winning out and instead captures a natural trade-off between efficiency and privacy.
4. *Interpretability*: The measurement is simple to interpret as $\log(\psi) > 0$ indicates Model *S* dominates Model *O*. This can be seen as $\log(\psi) = 0$ suggests that the competing models are equally good and $\log(\psi) < 0$ means that the Model *S* is inferior to the Model *O*.
5. *Generalization of earlier measurements*: By setting $w_1 = 1$, $w_2 = 0$, it reduces to efficiency-only measurement (*E*) and $w_1 = 0$, $w_2 = 1$, it reduces to privacy-only measurement as that of Yan et al. [16].

The Proposed weighted ψ measure thus not only integrates efficiency and privacy into a single criterion but also surpasses earlier methods through robustness, flexibility, and interpretability. It represents a clear advancement in the comparative evaluation of randomized response models and provides researchers with a powerful tool for balancing statistical accuracy with privacy protection.

4.1 Further Properties of ψ

(i) ψ (and $\log(\psi)$ is monotone in both the components P and E. This can be verified as $\frac{\partial \log(\psi)}{\partial P} = \frac{w_1}{P}$. Since $w_1, P > 0$, we have $\frac{\partial \log(\psi)}{\partial P} > 0$. This implies $\log(\psi)$ is strictly increasing in *P* .Similarly, $\frac{\partial \log(\psi)}{\partial E} = \frac{w_2}{E} > 0$ implies that $\log(\psi)$ is strictly increasing in *E*. Hence, we can conclude that for $w_1, w_2 > 0$, the ψ is strictly increasing in both *P* and *E*. This implies that increasing both *P* and *E* improve ψ and there is no paradoxical behavior of ψ. Consequently, ψ is a consistent decision measure.

(ii) ψ is scale and unit invariant. To show this, suppose MSE_o is rescaled as $cMSE_0$ and MSE_s is rescaled as $cMSE_s$ then $E=\frac{cMSE_o}{cMSE_s}=\frac{MSE_o}{MSE_s}$. Similarly, ∇_o is rescaled as $k\nabla_o$ and ∇_s is rescaled as $k\nabla_s$. So ∇_o is rescaled as $P=\frac{k\nabla_s}{k\nabla_o}=\frac{\nabla_s}{\nabla_o}$. Thus ψ is invariant under common scaling. Now as we know that changing units (e.g., variance scale, measurement units) multiplies numerator and denominator equally, ratios remain unchanged. This means ψ is a unit free measure.

(iii) ψ is log- scale translation invariant. This can be shown by letting P is translated to cP then $\log(\psi)$ translates to $\log(\psi)+w_1\log(c)$. We see there is an absolute value shift in $\log(\psi)$. It implies ranking between models remain unchanged. Hence, ordering is invariant under multiplicative scaling.

(iv) ψ is a symmetric measure. To check this, we reverse the comparison of Model *S* versus Model *O* as Model O versus Model *S* then $(\psi)^{-1}=\left(\frac{\nabla_o}{\nabla_s}\right)^{w_1}\times\left(\frac{MSE_s}{MSE_o}\right)^{w_2}$. Then $\log(\psi)^{-1}=-\log(\psi)$. Hence, we have a perfect symmetry between the competing Models *O* and *S*.

5. Some Existing Scrambling Models

We now present some of the important and recent scrambling models.

5.1 Warner [5] Model

Using Warner [5] model, the response observed from the i^{th} respondent can be written as:

$$Z_{1i}=Y_i+S_i.$$

Since $E(Z_{1i})=\mu_y$, an unbiased estimator of μ_y, under the Warner [5] model is written as:

$$\hat{\mu}_W=\frac{1}{n}\sum_{i=1}^{n}Z_{1i},$$

with variance given by

$$Var(\hat{\mu}_W)=\frac{1}{n}\left(\sigma_y^2+\sigma_s^2\right). \tag{11}$$

The privacy measure ∇, is calculated as

$$\nabla_W=\sigma_s^2. \tag{12}$$

5.2 Eichhorn and Hayre [19] Model

The observed i^{th} responses according to the Eichhorn and Hayre [19] model is given below:

$$Z_{2i}=Y_iS_i.$$

An unbiased estimator of sensitive mean μ_y based on the responses Z_{2i}, $i=(1,2,...,n)$ is given as:

$$\hat{\mu}_{EH}=\frac{1}{n}\sum_{i=1}^{n}Z_{2i}.$$

The variance of $\hat{\mu}_{EH}$ is given by:

$$Var(\hat{\mu}_{EH})=\frac{1}{n}\left[\sigma_y^2+\sigma_t^2\mu'_{2y}\right]. \tag{13}$$

The privacy measure ∇, is given as

$$\nabla_{EH}=\sigma_t^2\mu'_{2y}. \tag{14}$$

5.3 The Diana and Perri [20] Model

The i^{th} response revealed under the Diana and Perri [20] model is written as:

$$Z_{3i}=Y_iT_i+S_i.$$

An unbiased estimator of μ_y using Z_{3i} is given by:

$$\hat{\mu}_{DP}=\frac{1}{n}\sum_{i=1}^{n}Z_{3i}.$$

Its variance is given by:

$$Var(\hat{\mu}_{DP}) = \frac{1}{n}\left[\sigma_y^2 + \sigma_s^2 + \sigma_t^2 \mu'_{2y}\right]. \tag{15}$$

The privacy measure ∇, for this model is given by :

$$\nabla_{DP} = \sigma_y^2 + \sigma_s^2 + \sigma_t^2 \mu'_{2y}. \tag{16}$$

5.4 The Gjestvang ang Singh [21] Model

The distribution of i^{th} response under the Gjestvang and Singh [21] model can be written as:

$$Z_{4i} = \begin{cases} Y_i - \beta S_i & \text{with probability } p_1 = \dfrac{\alpha}{\alpha+\beta} \\ Y_i + \alpha S_i & \text{with probability } p_2 = \dfrac{\beta}{\alpha+\beta}, \end{cases}$$

where α, and β are constants determined by the interviewer. Gjestvang and Singh [21] suggested the mean estimator as

$$\hat{\mu}_{GS} = \frac{1}{n}\sum_{i=1}^{n} Z_{4i}.$$

The variance of $\hat{\mu}_{GS}$ is given by:

$$Var(\hat{\mu}_{GS}) = \frac{1}{n}\left[\sigma_y^{\ 2} + \alpha\beta\left(\sigma_s^{\ 2}\right)\right]. \tag{17}$$

The ∇ measure for this model is given by:

$$\nabla_{GS} = \alpha\beta\left(\sigma_s^{\ 2}\right). \tag{18}$$

5.5 Murtaza et al. [22] Model

The distribution of i^{th} response under the Murtaza et al. [22] model may be written as:

$$Z_{5i} = \begin{cases} Y_i, & \text{with probability } (1-w) \\ Y_i T_i + \alpha S_i, & \text{with probability } (w), \end{cases}$$

where $\boldsymbol{\alpha}$ is a constant and w is unknown proportion of the individuals in the population considering the study variable sensitive enough to hide. An unbiased estimator of μ_y through Murtaza et al. [22] model is given by:

$$\hat{\mu}_M = \frac{1}{n}\sum_{i=1}^{n} Z_{5i}.$$

Assuming independent variables, the variance of $\hat{\mu}_M$ is given by:

$$Var(\hat{\mu}_M) = \frac{1}{n}\left[\sigma_y^{\ 2} + W\left\{\alpha^2\sigma_s^{\ 2} + \sigma_t^{\ 2}\mu'_{2y}\right\}\right]. \tag{19}$$

The ∇ measure of respondent's privacy is given by :

$$\nabla_M = \left\{\sigma_t^{\ 2}\mu'_{2y} + \alpha^2\sigma_s^{\ 2}\right\}. \tag{20}$$

5.6 Narjis and Shabbir [23] scrambling Model

The distribution of i^{th} response based on the Narjis and Shabbir [23] Model is as follows:

$$Z_{6i} = \begin{cases} Y_i - \beta S_i & \text{with probability } P_1 = \dfrac{\alpha}{\alpha+\beta+\gamma} \\ Y_i + \alpha S_i & \text{with probability } P_2 = \dfrac{\beta}{\alpha+\beta+\gamma} \\ Y_i & \text{with probability } P_3 = \dfrac{\gamma}{\alpha+\beta+\gamma}, \end{cases}$$

where α, β and γ denote the constants determined by the interviewer.

The estimator of μ_y using the Narjis and Shabbir [23] model is defined as :

$$\hat{\mu}_{NS} = \frac{1}{n}\sum_{i=1}^{n} Z_{6i}.$$

The variance of $\hat{\mu}_{NS}$ is given by :

$$Var(\hat{\mu}_{NS}) = \frac{1}{n}\left[\sigma_y^{\,2} + \frac{\alpha\beta(\alpha+\beta)\sigma_s^{\,2}}{\alpha+\beta+\gamma}\right]. \tag{21}$$

The privacy measure ∇, for this model is given by:

$$\nabla_{NS} = \frac{\alpha\beta(\alpha+\beta)\sigma_s^{\,2}}{\alpha+\beta+\gamma}. \tag{22}$$

5.7 Gupta et al. [24] Model

Using the Gupta et al. [24] scrambling model, the distribution of i^{th} response can be written as:

$$Z_{7i} = \begin{cases} Y_i & \text{with probability} \quad 1\text{-}W \\ Y_i + S_i & \text{with probability} \quad AW \\ Y_i T_i + S_i & \text{with probability} \quad (1\text{-}A)W. \end{cases}$$

An unbiased estimator of μ_y through the Gupta et al. [24] model is given by:

$$\hat{\mu}_G = \frac{1}{n}\sum_{i=1}^{n} Z_{7i}.$$

The variance of $\hat{\mu}_G$ is given below.

$$Var(\hat{\mu}_G) = \frac{1}{n}\left[\sigma_y^2 + W\sigma_s^2 + W(1-A)\sigma_t^2\mu'_{2y}\right] \tag{23}$$

Finally, the privacy measure ∇, for Gupta et al. [24] model is given by :

$$\nabla_G = \frac{1}{n}\left[\sigma_s^{\,2} + (1-A)\sigma_t^{\,2}\mu'_{2y}\right]. \tag{24}$$

6. Proposed Models as Special Cases of Generalized Framework

In this section, we present three different proposed randomized response models which are special cases of the generalized framework.

6.1 Proposed Model 1 (M_1)

The i^{th} response under M_1 is written as

$$R_{M_{1i}} = D_{1i}Y_i(1+S_i) + (1-D_{1i})Y_i(1-S_i),$$

where $D_{1i} \sim Bernoulli(P_1)$. The distribution of i^{th} respondent response under the M_1 is given as:

$$R_{M_{1i}} = \begin{cases} Y_i(1+S_i) & \text{with probability} \quad P_1 \\ Y_i(1-S_i) & \text{with probability}\,(1-P_1) \end{cases}$$

An unbiased mean estimator μ_y of the sensitive variable is presented by M_1 can be expressed as:

$$\hat{\mu}_{M_1} = \frac{1}{n}\sum_{i=1}^{n} R_{M_{1i}}$$

The variance of $\hat{\mu}_{M_1}$ is given by:

$$Var(\hat{\mu}_{M_1}) = \frac{1}{n}\left[\sigma_y^{\,2} + \sigma_s^{\,2}\mu'_{2y}\right] \tag{25}$$

For the M_1, respondent privacy can be written as:

$$\nabla_{M_1} = \sigma_s^{\,2}\mu'_{2y} \tag{26}$$

6.2 Proposed Model 2 (M_2)

The i^{th} response under M_2 is written as

$$R_{M_{2i}} = D_{2i}Y_i(1+\alpha S_i)+(1-D_{2i})Y_i(1-\beta S_i),$$

where $D_{1i} \sim Bernoulli\left(P_1 = \frac{\alpha}{\alpha+\beta}\right)$. The distribution of i^{th} respondent response under the M_2 is given as:

$$R_{M_{2i}} = \begin{cases} Y_i(1+\alpha S_i) & \text{with probability} \quad P_2 = \dfrac{\alpha}{\alpha+\beta} \\ Y(1-\beta S_i) & \text{with probability } (1-P_2) = \dfrac{\beta}{\alpha+\beta} \end{cases}$$

where α and β denote the constants which are determined by the interviewer.

An unbiased mean estimator μ_y of the sensitive variable M_2 is written as:

$$\hat{\mu}_{M_2} = \frac{1}{n}\sum_{i=1}^{n} R_{M_{2i}}$$

The variance of $\hat{\mu}_{M_2}$ is given by:

$$Var(\hat{\mu}_{M_2}) = \frac{1}{n}\left[\sigma_y^{\,2} + \sigma_s^{\,2}\mu'_{2y}\left(P_2\alpha^2 + (1-P_2)\beta^2\right)\right]. \tag{27}$$

For the M_2, the measure of respondent privacy is written as:

$$\nabla_{M_2} = \sigma_s^{\,2}\mu'_{2y}\left(P_2\alpha^2 + (1-P_2)\beta^2\right). \tag{28}$$

6.3 Proposed Model 3 (M_3)

The i^{th} response under M_3 is written as

$$R_{M_{3i}} = D_{3i}Y_i(1+\alpha S_i)+D_{4i}Y_i(1-\beta S_i)+(1-D_{3i}-D_{4i})Y_i,$$

where $D_{3i} \sim Bernoulli\left(p_1 = \frac{\alpha}{\alpha+\beta+\gamma}\right)$ and $D_{4i} \sim Bernoulli\left(p_2 = \frac{\beta}{\alpha+\beta+\gamma}\right)$

The distribution of i^{th} response under the M_3 is given by:

$$R_{M_{3i}} = \begin{cases} Y_i(1+\alpha S_i) & \text{with probability } \; p_1 = \dfrac{\alpha}{\alpha+\beta+\gamma} \\ Y_i(1-\beta S_i) & \text{with probability } \; p_2 = \dfrac{\beta}{\alpha+\beta+\gamma} \\ Y_i & \text{with probability } \; p_3 = \dfrac{\gamma}{\alpha+\beta+\gamma}, \end{cases}$$

where α, β and γ denote the constants determined by the interviewer.

An unbiased mean estimator μ_y of the sensitive variable of M_3 is given as:

$$\hat{\mu}_{M_3} = \frac{1}{n}\sum_{i=1}^{n} R_{M_{3i}}$$

The variance of $\hat{\mu}_{M_3}$ is given by:

$$Var(\hat{\mu}_{M_3}) = \frac{1}{n}\left[\sigma_y^{\,2} + \frac{\alpha\beta(\alpha+\beta)(\sigma_y^{\,2}+\mu_y^{\,2})\sigma_s^{\,2}}{\alpha+\beta+\gamma}\right] \tag{29}$$

For the M_3, the measure of respondent privacy is given by:

$$\nabla_{M_3} = \frac{\alpha\beta(\alpha+\beta)\left(\sigma_y^{\ 2}+\mu_y^{\ 2}\right)\sigma_s^{\ 2}}{\alpha+\beta+\gamma} \tag{30}$$

6.4 Proposed Models as Special Cases of Generalized Framework

The proposed models are not entirely new in isolation, but their formulation under a unified random coefficient interaction framework provides new theoretical insights, including a closed-form variance structure and a clear efficiency–privacy trade-off governed by the second moment of the random coefficient. The proposed models are formulated within a unified random coefficient interaction framework $R_i = Y_i + C_i Y_i S_i$, where the random coefficient C_i governs the level and nature of perturbation. This formulation generalizes existing additive and multiplicative models while introducing an additional layer of randomness through the coefficient structure. In particular, M_3 incorporates a zero-distortion state, allowing truthful responses with positive probability, which is not commonly accommodated in traditional randomized response designs. Furthermore, the proposed framework yields a unified variance expression and a clear efficiency–privacy trade-off governed by $E\left(C_i^2\right)$, providing theoretical insights that are not explicitly available in existing literature. The proposed models can be expressed within the random coefficient class as given in the Table 1 below.

Table 1: The proposed models and their random coefficient C_i.

Model	**Expression**	**Random Coefficient** C_i
M_1	$R_{M_{1i}} = D_{1i}\left(Y_i + Y_i S_i\right)_i + \left(1-D_{1i}\right)\left(Y_i - Y_i S_i\right)$	$2D_{1i}-1$
M_2	$R_{M_{2i}} = D_{2i}\left(Y_i + \alpha Y_i S_i\right) + \left(1-D_{2i}\right)\left(Y_i - \beta Y_i S_i\right)$	$D_{2i}\alpha - \left(1-D_{2i}\right)\beta$
M_3	$R_{M_{3i}} = D_{3i}Y_i\left(1+\alpha S_i\right) + D_{4i}Y_i\left(1-\beta S_i\right) + \left(1-D_{3i}-D_{4i}\right)\left(Y_i\right)$	$D_{3i}\alpha - D_{4i}\beta$

Now, we see that how M_1, M_2 and M_3 differ in structure**.** Recall the common structure $R_i = Y_i + C_i Y_i S_i$**.** So, the only thing that differentiates models is the structure of C_i**.** Model M_1 has a binary symmetric coefficient such as $C_i = 2D_{1i}-1$ and $D_{1i} \sim Bernoulli(P)$**.** The random coefficient $C_i \in \{-1,1\}$. Model M_2 is identified as asymmetric two-point coefficient model where $C_i \in \{\alpha,-\beta\}$**.** The Model M_3 is seen to be a multinomial coefficient model as in this case $C_i = D_{3i}\alpha - D_{4i}\beta$ and $C_i \in \{\alpha,-\beta,0\}$**.** This establishes that the proposed models are structurally different with different level of flexibility. It is important to note that there exist mutual relationships between the models**.**

(i) M_1 as a Special Case of M_2

Set $\alpha = 1, \beta = 1$ Then**:** $C_i = D_{2i}\alpha - \left(1-D_{2i}\right)\beta = D_{2i}(1) - \left(1-D_{2i}\right)(1) = 2D_{2i}-1$

Hence: $M_1 \subset M_2$

(ii) M_2 as a Special Case of M_3

If we set **,** $D_{3i} = D_{2i}$, and $D_{4i} = 1-D_{2i}$ Then: $C_i = D_{3i}\alpha - D_{4i}\beta = D_{2i}\alpha - \left(1-D_{2i}\right)\beta$.

Hence $M_2 \subset M_3$. So, we have the hierarchy as $M_1 \subset M_2 \subset M_3$.

6.5 Comparison of M_1, M_2 and M_3

In this section, comparison of M_1**,** M_2 and M_3 using the new ψ measure is presented. For clarity in presenting the comparison results, each of the proposed models is considered separately in comparison with the same set of seven existing randomized response models considered in section 5.

Let $E_{(j/M_1)} = \dfrac{Var\left(\hat{\mu}_j\right)}{Var\left(\hat{\mu}_{M_t}\right)}$ and $P_{(j/M_t)} = \dfrac{\nabla_{M_t}}{\nabla_j}$**,** $j = W, EH, DP, GS, M, NS, G$ and $t = 1,2,3$ be the relative efficiency and relative privacy, respectively, of the model M_t relative to j^{th} existing models considered in section 5. Then we have the $\log(\psi)$ measure for the competing models as given below.

$$\log\left(\psi_{(j/M_t)}\right) = w_1 \log\left(E_{(j/M_t)}\right) + w_2 \log\left(P_{(j/M_t)}\right). \tag{31}$$

Using (11)-(30) in (31) with $t = 1$, we get the $\log(\psi)$ measure for M_1 relative to existing models. Graphical presentation of the $\log(\psi)$ for different parametric settings is displayed in Figure 1 below.

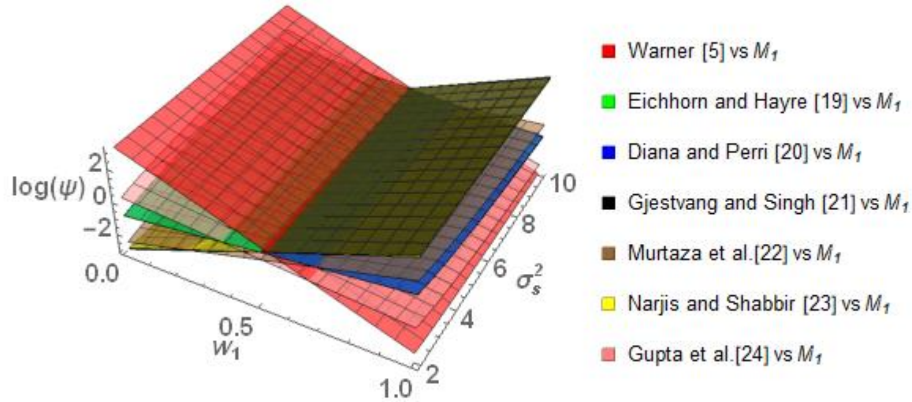

Figure 1: Behavior of $\log(\psi)$) for M_1 relative to existing models for $\sigma_y^2 = 2$, $\mu_y = 5$, $\sigma_t^2 = 5$, $\alpha = 25$, $\beta = 25$, $\gamma = 5$, $P = 0.4$, $P_1 = 0.4$, $P_2 = 0.6$, $A = 0.7$, $W = 0.2$.

The 3D plots in Figure 1 compare the performance of the proposed model M_1 with several existing randomized response models, including those of Warner [5], Eichhorn and Hayre [19], Diana and Perri [20], Gjestvang and Singh [21], Murtaza et al. [22], Narjis and Shabbir [23], and Gupta et al. [24]. The comparison is based on the measure $\log(\psi)$, which combines both efficiency and privacy into a single criterion. In this setup, positive values of $\log(\psi)$ indicate that the proposed model performs better than the competing model.

A careful look at the figure shows a clear pattern. The proposed model M_1 tends to perform better as the value of w_1 increases. For smaller values of w_1, some of the existing models remain competitive, and in a few cases perform slightly better. However, once w_1 moves into the moderate range, the advantage gradually shifts toward M_1. For example, the proposed model begins to outperform the Warner [5] model roughly beyond the mid-range of w_1, and a similar trend can be seen when compared with the models of Eichhorn and Hayre [19], and Diana and Perri [20], particularly when the variance parameter σ_s^2 is not too small.

For the models of Gjestvang and Singh [21] and Narjis and Shabbir [23], the dominance of M_1 is even more noticeable, as the corresponding surfaces lie mostly in the positive region across a wide span of the parameter space. A similar, though slightly less uniform, pattern is observed for the model of Murtaza et al.[22], where the proposed model again performs better for moderate to higher values of w_1. In the case of Gupta et al. 24], the proposed model appears to dominate almost entirely, except possibly when w_1 is very close to zero.

Overall, the figure suggests that the proposed model provides a more balanced performance, particularly when greater weight is given to the combined efficiency–privacy criterion. In practical terms, this means that M_1 becomes increasingly preferable as the emphasis shifts toward joint performance rather than relying on either efficiency or privacy alone.

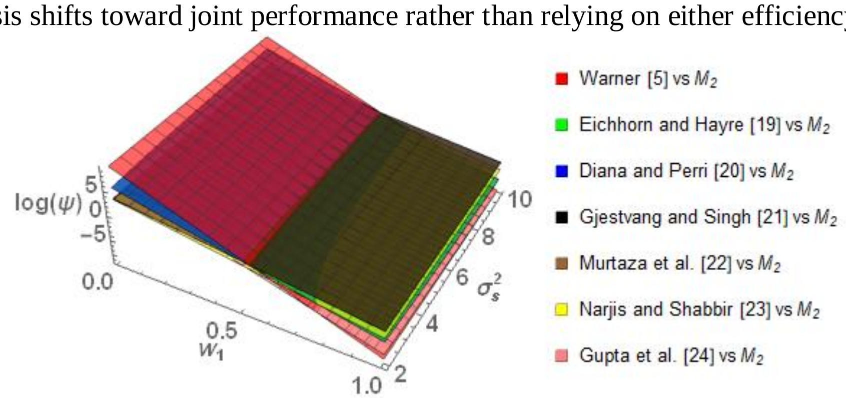

Figure 2: Behavior of $\log(\psi)$) for M_2 relative to existing models for $\sigma_y^2 = 2$, $\mu_y = 5$, $\sigma_t^2 = 5$, $\alpha = 25$, $\beta = 25$, $\gamma = 5$, $P = 0.4$, $P_1 = 0.4$, $P_2 = 0.6$, $A = 0.7$, $W = 0.2$.

The 3D plots in Figure 2 compare the performance of the proposed model M_2 with several well-known randomized response models using the measure ψ. Here, positive values of $\log(\psi)$ indicate that M_2 performs better than the competing model.

One clear pattern that emerges from the figure is that the value of $\log(\psi)$ generally decreases as w_1 increases. This suggests that the relative advantage of M_2 becomes weaker when more weight is assigned to the combined efficiency–privacy component. In fact, for most of the competing models, the surfaces start in the positive region for small values of w_1, but gradually move toward zero and then into the negative region as w_1 increases.

For smaller values of w_1, the proposed model performs quite well. In this range, the surfaces corresponding to all competing models remain above zero, indicating that w_1 consistently outperforms them. This advantage is fairly strong against the Warner model and is also clearly visible for the models of Diana and Perri [20], Eichhorn and Hayre [19], and Gjestvang and Singh [21].

As w_1 increases to moderate levels, the situation begins to change. The gap between M_2 and some competing models narrows, and in certain cases the competing models start to perform better. This is particularly noticeable for the models of Narjis and Shabbir [23], Murtaza et al.[22], and Gupta et al. [24], where the corresponding surfaces approach or fall below zero.

The role of σ_s^2 appears to be less pronounced. While it slightly affects the height and slope of the surfaces, it does not seem to change the overall pattern of dominance. The main shift is clearly driven by w_1.

Overall, the Figure 2 suggests that M_2 is more suitable when the value of w_1 is relatively small. As w_1 increases, its advantage reduces, and other models may become preferable. In this sense, the performance of M_2 appears to be more sensitive to the choice of the weighting parameter compared to what was observed for M_1.

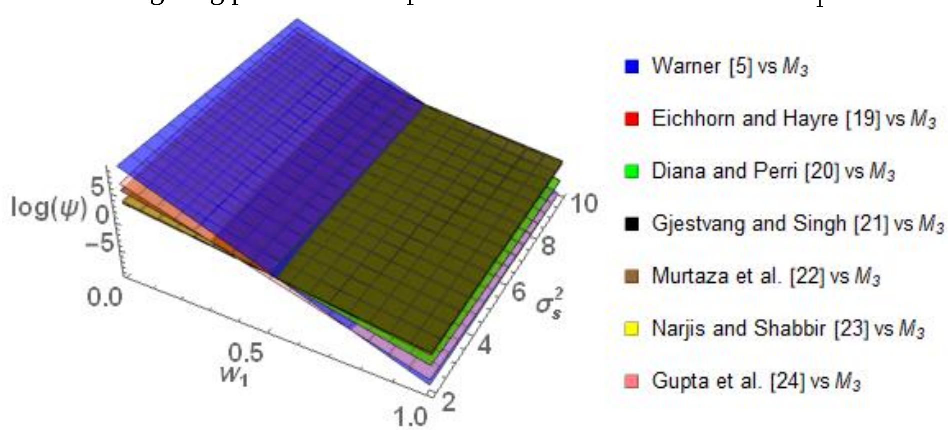

Figure 3: Behavior of $\log(\psi)$) for M_3 relative to existing models for $\sigma_y^2 = 2$, $\mu_y = 5$, $\sigma_t^2 = 5$, $\alpha = 25$, $\beta = 25$, $\gamma = 5$, $P = 0.4$, $P_1 = 0.4$, $P_2 = 0.6$, $A = 0.7$, $W = 0.2$.

The 3D plots in Figure 3 present a comparison of the proposed model M_3 with several existing randomized response models using the combined performance measure ψ. As in the previous figures, positive values of $\log(\psi)$ indicate that M_3 performs better than the competing model.

A noticeable feature of the figure is that $\log(\psi)$ tends to decrease as w_1 increases. For smaller values of w_1, most of the surfaces lie above zero, indicating that M_3 performs better than the competing models in this region. This advantage is quite visible across nearly all models, including Warner [5], Eichhorn and Hayre [19], and Diana and Perri [20].

However, as w_1 increases, the surfaces gradually move downward and, in several cases, cross into the negative region. This suggests that the relative performance of M_3 weakens for moderate to larger values of w_1. In particular, for models such as Gjestvang and Singh [21], Murtaza et al. [22], Narjis and Shabbir [23], and Gupta et al. [24], the advantage of M_3 reduces noticeably and may even reverse in parts of the parameter space.

The effect of σ_s^2 appears to be relatively mild compared to w_1. While variations in σ_s^2 slightly adjust the level of the surfaces, they do not significantly alter the overall pattern. The dominant factor influencing performance remains the weight parameter w_1.

Overall, the Figure 3 suggests that M_3 performs well when w_1 is small, but its advantage diminishes as w_1 increases. This indicates that, similar to M_2, the effectiveness of M_3 is more pronounced under lower weighting of the combined efficiency–privacy component.

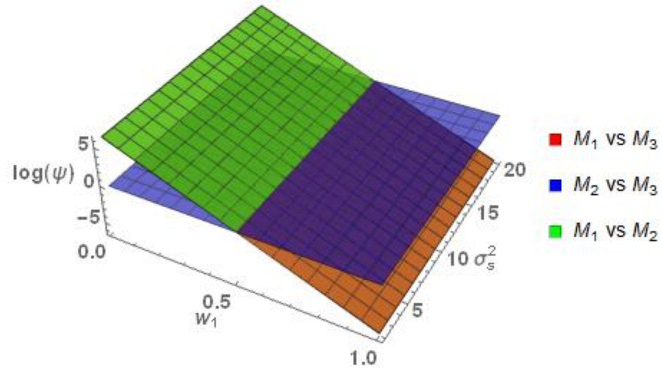

Figure 4: Behavior of $\log(\psi)$) for M_1 vs M_2, M_1 vs M_3 and M_2 vs M_3 for $\sigma_y^2 = 2$, $\mu_y = 5$, $\sigma_t^2 = 5$, $\alpha = 25$, $\beta = 25$, $\gamma = 5$, $P = 0.4$, $P_1 = 0.4$, $P_2 = 0.6$, $A = 0.7$, $W = 0.2$.

The 3D surfaces in the figure 4 present a direct comparison among the three proposed models M_1, M_2 and M_3 using the performance measure ψ. As before, positive values of $\log(\psi)$ indicate that the model in the numerator of the comparison outperforms the one in the denominator.

A clear pattern emerges when observing how the surfaces behave with respect to w_1. The surface corresponding to the comparison between M_1 and M_2 (shown in green) remains largely above zero across most of the parameter space. This suggests that M_1 consistently outperforms M_2 and this advantage becomes more pronounced as w_1 increases.

A similar trend can be seen when comparing M_1 with M_3 (red surface). For moderate to higher values of w_1, the surface lies above zero, indicating that M_1 dominates M_3 in this region. However, for smaller values of w_1, the surface approaches zero and may dip slightly below it, suggesting that the advantage of M_1 is less clear and that M_3 can be competitive in this range.

In contrast, the comparison between M_2 and M_3 (blue surface) shows a more balanced behavior. For smaller values of w_1, the surface tends to lie above zero, indicating that M_2 performs better than M_3 in this region. As w_1 increases, however, the surface declines and approaches or crosses zero, suggesting that the relative advantage between these two models diminishes and may even reverse.

The role of σ_s^2 appears to be secondary in all three comparisons. While it influences the magnitude of $\log(\psi)$, it does not significantly change the overall ordering of the models. The primary driver of the relative performance remains the weight parameter w_1.

Taken together, the pairwise comparisons reveal a clear ordering among the proposed models that depends primarily on the value of the weight parameter w_1. For smaller values of w_1, the models M_2 and M_3 tend to perform better, with M_2 showing a slight advantage over M_3. However, as w_1 increases, the performance of M_1 improves steadily and eventually surpasses both M_2 and M_3.

This indicates that no single model is uniformly superior across all parameter settings. Instead, each model performs best within a specific region of the design space. In particular, M_2 and M_3 are more suitable for lower values of w_1, whereas

M_1 becomes the preferred choice when w_1 is moderate to large. This structured behavior provides useful guidance for model selection based on the desired balance between efficiency and privacy.

In order to have more clarity of the behavior of $\log(\psi)$, different design parameters' settings are considered. The values of $\log(\psi)$ for M_1, M_2 and M_3 are arranged in the Tables 2-9, given below.

Table 2: Values of $\log(\psi)$ for M_1 when $\sigma_y^2 = 2$, $\mu_y = 2$, $\sigma_t^2 = 3$, $\alpha = 15$, $\beta = 15$, $\gamma = 10$, $A = 0.7$, $P = 0.6$, $W = 0.2$, $P_1 = 0.6$, $P_2 = 0.4$.

σ_s^2	w_1	$\log\left(\psi_{(W/M_1)}\right)$	$\log\left(\psi_{(EH/M_1)}\right)$	$\log\left(\psi_{(DP/M_1)}\right)$	$\log\left(\psi_{(GS/M_1)}\right)$	$\log\left(\psi_{(M/M_1)}\right)$	$\log\left(\psi_{(NS/M_1)}\right)$	$\log\left(\psi_{(G/M_1)}\right)$
2	0.05	1.639533	-0.46269	-0.36736	-3.26939	-3.01041	-2.74729	0.420429
2	0.15	1.335081	-0.3664	-0.29114	-2.5595	-2.3579	-2.20917	0.294433
2	0.3	0.878403	-0.22198	-0.17682	-1.49465	-1.37914	-1.40198	0.10544
2	0.45	0.421724	-0.07756	-0.0625	-0.42981	-0.40037	-0.59479	-0.08355
2	0.6	-0.03495	0.066861	0.051819	0.635039	0.578386	0.212393	-0.27255
2	0.75	-0.49163	0.211282	0.16614	1.699883	1.557148	1.019579	-0.46154
2	0.9	-0.94831	0.355704	0.280461	2.764728	2.535909	1.826765	-0.65053
4	0.05	1.628855	0.078659	0.26018	-3.2658	-3.00685	-2.12707	0.829242
4	0.15	1.303045	0.061953	0.205175	-2.54871	-2.34722	-1.71855	0.613038
4	0.3	0.814331	0.036895	0.122668	-1.47308	-1.35779	-1.10577	0.288731
4	0.45	0.325616	0.011837	0.040161	-0.39745	-0.36835	-0.49299	-0.03558
4	0.6	-0.1631	-0.01322	-0.04235	0.678174	0.621081	0.119789	-0.35988
4	0.75	-0.65181	-0.03828	-0.12485	1.753803	1.610516	0.73257	-0.68419
4	0.9	-1.14053	-0.06334	-0.20736	2.829432	2.599951	1.345351	-1.00849
6	0.05	1.624264	0.366217	0.626397	-3.26454	-3.0056	-1.76812	1.019495
6	0.15	1.289274	0.287722	0.492897	-2.54493	-2.34348	-1.43649	0.758675
6	0.3	0.786788	0.169979	0.292647	-1.46551	-1.35029	-0.93905	0.367444
6	0.45	0.284303	0.052235	0.092397	-0.3861	-0.35711	-0.44161	-0.02379
6	0.6	-0.21818	-0.06551	-0.10785	0.693316	0.636075	0.05583	-0.41502
6	0.75	-0.72067	-0.18325	-0.3081	1.77273	1.629259	0.553271	-0.80625
6	0.9	-1.22315	-0.30099	-0.50835	2.852145	2.622443	1.050712	-1.19748
8	0.05	1.6217	0.553458	0.885973	-3.26389	-3.00496	-1.51575	1.131877
8	0.15	1.28158	0.434166	0.696261	-2.543	-2.34156	-1.23874	0.843737
8	0.3	0.7714	0.255228	0.411693	-1.46165	-1.34647	-0.82324	0.411529
8	0.45	0.261221	0.076289	0.127125	-0.3803	-0.35137	-0.40773	-0.02068
8	0.6	-0.24896	-0.10265	-0.15744	0.701041	0.643727	0.007777	-0.45289
8	0.75	-0.75914	-0.28159	-0.44201	1.782387	1.638823	0.423285	-0.8851
8	0.9	-1.26932	-0.46053	-0.72658	2.863732	2.63392	0.838792	-1.31731
10	0.05	1.62006	0.687736	1.087204	-3.2635	-3.00457	-1.32167	1.206595
10	0.15	1.276661	0.538928	0.853667	-2.54182	-2.3404	-1.08691	0.899832
10	0.3	0.761563	0.315717	0.50336	-1.45931	-1.34415	-0.73478	0.439687
10	0.45	0.246465	0.092505	0.153054	-0.37679	-0.34789	-0.38264	-0.02046
10	0.6	-0.26863	-0.13071	-0.19725	0.705727	0.648369	-0.03051	-0.4806
10	0.75	-0.78373	-0.35392	-0.54756	1.788244	1.644625	0.321628	-0.94075
10	0.9	-1.29883	-0.57713	-0.89786	2.870762	2.640882	0.673762	-1.40089

Table 3: Values of $\log(\psi)$ for M_1 when $\sigma_y^2 = 2$, $\mu_y = 5$, $\sigma_t^2 = 5$, $\alpha = 25$, $\beta = 25$, $\gamma = 5$, $A = 0.7$, $P = 0.4$, $W = 0.2$, $P_1 = 0.4$, $P_2 = 0.6$.

σ_s^2	W_1	$\log\left(\psi_{(W/M_1)}\right)$	$\log\left(\psi_{(EH/M_1)}\right)$	$\log\left(\psi_{(DP/M_1)}\right)$	$\log\left(\psi_{(GS/M_1)}\right)$	$\log\left(\psi_{(M/M_1)}\right)$	$\log\left(\psi_{(NS/M_1)}\right)$	$\log\left(\psi_{(G/M_1)}\right)$
2	0.05	2.999092	-0.83899	-0.82574	-2.82946	-2.74367	-2.42957	0.17346
2	0.15	2.405603	-0.65498	-0.64465	-2.20456	-2.13781	-1.99479	0.041421
2	0.3	1.515369	-0.37896	-0.37301	-1.2672	-1.22902	-1.34262	-0.15664
2	0.45	0.625134	-0.10294	-0.10138	-0.32984	-0.32023	-0.69045	-0.3547
2	0.6	-0.2651	0.17307	0.170261	0.607522	0.588555	-0.03828	-0.55276
2	0.75	-1.15533	0.44909	0.441899	1.544881	1.497345	0.613893	-0.75081
2	0.9	-2.04557	0.7251	0.713537	2.48224	2.406135	1.266064	-0.94887
4	0.05	2.985609	-0.22731	-0.20101	-2.8286	-2.74282	-1.8072	0.75657
4	0.15	2.365153	-0.17725	-0.15675	-2.20197	-2.13524	-1.50872	0.496426
4	0.3	1.43447	-0.10216	-0.09035	-1.26203	-1.22388	-1.06101	0.10621
4	0.45	0.503786	-0.02706	-0.02395	-0.32209	-0.31252	-0.6133	-0.28401
4	0.6	-0.4269	0.04803	0.042443	0.617853	0.598839	-0.16558	-0.67422
4	0.75	-1.35758	0.12312	0.10884	1.557795	1.5102	0.282131	-1.06444
4	0.9	-2.28827	0.19821	0.175236	2.497737	2.421561	0.729844	-1.45465
6	0.05	2.980024	0.12504	0.164211	-2.82831	-2.74253	-1.44433	1.082008
6	0.15	2.348398	0.09745	0.12799	-2.2011	-2.13438	-1.22582	0.749736
6	0.3	1.400958	0.05607	0.073659	-1.26029	-1.22215	-0.89805	0.251328
6	0.45	0.453519	0.0147	0.019328	-0.31947	-0.30992	-0.57028	-0.24708
6	0.6	-0.49392	-0.02668	-0.035	0.621341	0.602311	-0.24252	-0.74549
6	0.75	-1.44136	-0.06806	-0.08933	1.562155	1.51454	0.085252	-1.2439
6	0.9	-2.3888	-0.10944	-0.14366	2.502969	2.426768	0.41302	-1.7423
8	0.05	2.97695	0.371424	0.423278	-2.82816	-2.74238	-1.18761	1.302974
8	0.15	2.339175	0.289405	0.329826	-2.20066	-2.13394	-1.02581	0.921492
8	0.3	1.382513	0.166375	0.189648	-1.25941	-1.22127	-0.78311	0.349269
8	0.45	0.425851	0.04336	0.049471	-0.31816	-0.30861	-0.54041	-0.22295
8	0.6	-0.53081	-0.07968	-0.09071	0.623093	0.604055	-0.29771	-0.79518
8	0.75	-1.48747	-0.20271	-0.23088	1.564345	1.51672	-0.05501	-1.3674
8	0.9	-2.44414	-0.32574	-0.37106	2.505597	2.429385	0.187693	-1.93962
10	0.05	2.975	0.559835	0.624199	-2.82808	-2.7423	-0.98903	1.467341
10	0.15	2.333327	0.43613	0.486302	-2.2004	-2.13368	-0.87116	1.049125
10	0.3	1.370817	0.250571	0.279457	-1.25888	-1.22075	-0.69435	0.421801
10	0.45	0.408307	0.065012	0.072611	-0.31737	-0.30782	-0.51754	-0.20552
10	0.6	-0.5542	-0.12055	-0.13423	0.624147	0.605104	-0.34073	-0.83285
10	0.75	-1.51671	-0.30611	-0.34108	1.565662	1.518031	-0.16391	-1.46017
10	0.9	-2.47922	-0.49166	-0.54792	2.507178	2.430958	0.012896	-2.08749

Table 4: Values of $\log(\psi)$ for M_2 when $\sigma_y^2 = 2$, $\mu_y = 2$, $\sigma_t^2 = 3$, $\alpha = 15$, $\beta = 15$, $\gamma = 10$, $A = 0.7$, $P = 0.6$, $W = 0.2$, $P_1 = 0.6$, $P_2 = 0.4$.

σ_s^2	W_1	$\log\left(\psi_{(W/M_2)}\right)$	$\log\left(\psi_{(EH/M_2)}\right)$	$\log\left(\psi_{(DP/M_2)}\right)$	$\log\left(\psi_{(GS/M_2)}\right)$	$\log\left(\psi_{(M/M_2)}\right)$	$\log\left(\psi_{(NS/M2)}\right)$	$\log\left(\psi_{(G/M_2)}\right)$
2	0.05	6.521694	4.514803	4.419476	1.612768	1.871756	1.497833	5.271815
2	0.15	5.149363	3.523138	3.447878	1.254786	1.456384	0.988422	4.016391
2	0.3	3.090866	2.03564	1.990481	0.717812	0.833327	0.224305	2.133254
2	0.45	1.032369	0.548143	0.533084	0.180838	0.21027	-0.53981	0.250118
2	0.6	-1.02613	-0.93936	-0.92431	-0.35614	-0.41279	-1.30393	-1.63302

2	0.75	-3.08462	-2.42685	-2.38171	-0.89311	-1.03584	-2.06805	-3.51615
2	0.9	-5.14312	-3.91435	-3.83911	-1.43008	-1.6589	-2.83216	-5.39929
4	0.05	6.507329	5.138654	4.957133	1.612676	1.871627	1.514812	5.673838
4	0.15	5.106266	4.008396	3.865174	1.254509	1.455997	1.000524	4.314624
4	0.3	3.004672	2.31301	2.227237	0.717259	0.832553	0.229091	2.275804
4	0.45	0.903079	0.617624	0.5893	0.180008	0.209109	-0.54234	0.236984
4	0.6	-1.19852	-1.07776	-1.04864	-0.35724	-0.41434	-1.31378	-1.80184
4	0.75	-3.3001	-2.77315	-2.68658	-0.89449	-1.0377	-2.08521	-3.84066
4	0.9	-5.4017	-4.46853	-4.32451	-1.43174	-1.6612	-2.85664	-5.87948
6	0.05	6.50144	5.50357	5.243399	1.612645	1.87158	1.520543	5.86041
6	0.15	5.08861	4.2922	4.087065	1.254417	1.45586	1.0046	4.449218
6	0.3	2.96937	2.47523	2.352565	0.717074	0.83229	0.230685	2.33243
6	0.45	0.85013	0.65822	0.618065	0.179731	0.20872	-0.54323	0.215642
6	0.6	-1.2691	-1.15878	-1.11644	-0.35761	-0.4148	-1.31714	-1.90115
6	0.75	-3.3883	-2.97579	-2.85094	-0.89495	-1.0384	-2.09106	-4.01793
6	0.9	-5.5076	-4.7928	-4.58544	-1.4323	-1.662	-2.86497	-6.13472
8	0.05	6.49822	5.76249	5.429981	1.61263	1.87156	1.523421	5.970235
8	0.15	5.07894	4.49362	4.231532	1.25437	1.45580	1.006646	4.526611
8	0.3	2.95003	2.59032	2.433859	0.716981	0.83216	0.231482	2.361176
8	0.45	0.82111	0.68702	0.636186	0.179592	0.20852	-0.54368	0.195741
8	0.6	-1.3078	-1.21628	-1.16149	-0.3578	-0.4151	-1.31884	-1.96969
8	0.75	-3.4367	-3.11958	-2.95916	-0.89519	-1.0387	-2.09401	-4.13513
8	0.9	-5.5656	-5.02289	-4.75683	-1.43257	-1.6623	-2.86917	-6.30056
10	0.05	6.49618	5.96332	5.563859	1.612621	1.87154	1.525153	6.043009
10	0.15	5.07282	4.64983	4.335095	1.254343	1.45576	1.007876	4.576874
10	0.3	2.93779	2.67959	2.49195	0.716926	0.83208	0.231961	2.37767
10	0.45	0.80276	0.70935	0.648804	0.179509	0.20841	-0.54395	0.178466
10	0.6	-1.3322	-1.26089	-1.19434	-0.35791	-0.4152	-1.31987	-2.02074
10	0.75	-3.4673	-3.23113	-3.03749	-0.89532	-1.0389	-2.09579	-4.21994
10	0.9	-5.6023	-5.20137	-4.88063	-1.43274	-1.6626	-2.8717	-6.41915

Table 5: Values of $\log(\psi)$ for M_2 when $\sigma_y^2 = 2$, $\mu_y = 5$, $\sigma_t^2 = 5$, $\alpha = 25$, $\beta = 25$, $\gamma = 5$, $A = 0.7$, $P = 0.4$, $W = 0.2$, $P_1 = 0.4$, $P_2 = 0.6$

σ_s^2	W_1	$\log\left(\psi_{(W/M_2)}\right)$	$\log\left(\psi_{(EH/M_2)}\right)$	$\log\left(\psi_{(DP/M_2)}\right)$	$\log\left(\psi_{(GS/M_2)}\right)$	$\log\left(\psi_{(M/M_2)}\right)$	$\log\left(\psi_{(NS/M_2)}\right)$	$\log\left(\psi_{(G/M_2)}\right)$
2	0.05	8.794884	4.970047	4.956801	2.96633	3.202482	2.793837	5.939599
2	0.15	6.917475	3.86722	3.856893	2.307317	2.491044	1.994951	4.464334
2	0.3	4.101362	2.212978	2.207032	1.318797	1.423886	0.796621	2.251436
2	0.45	1.285248	0.558737	0.557171	0.330276	0.356728	-0.40171	0.038539
2	0.6	-1.53086	-1.0955	-1.09269	-0.65824	-0.71043	-1.60004	-2.17436
2	0.75	-4.34698	-2.74975	-2.74255	-1.64676	-1.77759	-2.79837	-4.38726
2	0.9	-7.16309	-4.40399	-4.39241	-2.63528	-2.84474	-3.9967	-6.60015
4	0.05	8.780501	5.593881	5.567581	2.966292	3.202432	2.838636	6.521617
4	0.15	6.874327	4.352427	4.331925	2.307201	2.490892	2.029419	4.916062
4	0.3	4.015066	2.490246	2.47844	1.318566	1.423583	0.815594	2.507731
4	0.45	1.155805	0.628065	0.624956	0.32993	0.356275	-0.39823	0.099399
4	0.6	-1.70346	-1.23412	-1.22853	-0.65871	-0.71103	-1.61206	-2.30893

4	0.75	-4.56272	-3.0963	-3.08201	-1.64734	-1.77834	-2.82588	-4.71726
4	0.9	-7.42198	-4.95848	-4.9355	-2.63598	-2.84565	-4.03971	-7.1256
6	0.05	8.774613	5.9588	5.919633	2.966279	3.202415	2.854078	6.846573
6	0.15	6.856661	4.636254	4.605721	2.307163	2.490842	2.041296	5.167927
6	0.3	3.979734	2.652435	2.634855	1.318489	1.423482	0.822122	2.649957
6	0.45	1.102807	0.668616	0.663988	0.329815	0.356123	-0.39705	0.131988
6	0.6	-1.77412	-1.3152	-1.30688	-0.65886	-0.71124	-1.61623	-2.38598
8	0.75	-4.65105	-3.29902	-3.27775	-1.6475	-1.7786	-2.8354	-4.90395
8	0.9	-7.52797	-5.28284	-5.24861	-2.6362	-2.8459	-4.05457	-7.42192
8	0.05	8.771386	6.21771	6.165861	2.96627	3.20240	2.8619	7.06722
8	0.15	6.846981	4.83763	4.797211	2.30714	2.49081	2.04731	5.338728
8	0.3	3.960374	2.7675	2.744236	1.31845	1.42343	0.825426	2.74599
8	0.45	1.073767	0.69738	0.691262	0.32975	0.35604	-0.39646	0.153251
8	0.6	-1.81284	-1.37274	-1.36171	-0.6589	-0.7113	-1.61834	-2.43949
8	0.75	-4.69945	-3.44286	-3.41469	-1.6476	-1.7787	-2.84023	-5.03223
8	0.9	-7.58605	-5.51298	-5.46766	-2.6363	-2.8461	-4.06211	-7.62496
10	0.05	8.769345	6.41854	6.35418	2.96626	3.20241	2.866626	7.231341
10	0.15	6.840858	4.99383	4.943661	2.30713	2.49080	2.050944	5.465623
10	0.3	3.948128	2.85676	2.827882	1.31842	1.42340	0.827421	2.817046
10	0.45	1.055398	0.71970	0.712103	0.32972	0.35600	-0.3961	0.168469
10	0.6	-1.83733	-1.41736	-1.40368	-0.6589	-0.7114	-1.61962	-2.48011
10	0.75	-4.73006	-3.55443	-3.51945	-1.6476	-1.7788	-2.84315	-5.12868
10	0.9	-7.62279	-5.69149	-5.63523	-2.6363	-2.8462	-4.06667	-7.77726

Table 6: Values of $\log(\psi)$ for M_3 when $\sigma_y^2 = 2$, $\mu_y = 2$, $\sigma_t^2 = 3$, $\alpha = 15$, $\beta = 15$, $\gamma = 10$, $A = 0.7$, $P = 0.6$, $W = 0.2$, $P_1 = 0.6$, $P_2 = 0.4$.

σ_s^2	W_1	$\log\left(\psi_{(W/M_3)}\right)$	$\log\left(\psi_{(EH/M_3)}\right)$	$\log\left(\psi_{(DP/M_3)}\right)$	$\log\left(\psi_{(GS/M_3)}\right)$	$\log\left(\psi_{(M/M_3)}\right)$	$\log\left(\psi_{(NS/M_3)}\right)$	$\log\left(\psi_{(G/M_3)}\right)$
2	0.1	5.605358	3.7888	3.703506	1.203607	1.4339	1.905882	4.413932
2	0.3	2.975719	1.920493	1.875334	0.602665	0.71818	0.955351	2.018108
2	0.5	0.34608	0.052187	0.047162	0.001724	0.002461	0.00482	-0.37772
2	0.7	-2.28356	-1.81612	-1.78101	-0.59922	-0.71326	-0.94571	-2.77354
2	0.9	-4.9132	-3.68443	-3.60918	-1.20016	-1.42898	-1.89624	-5.16937
4	0.1	5.576639	4.343367	4.180996	1.203435	1.433654	1.920946	4.764073
4	0.3	2.889563	2.1979	2.112127	0.602149	0.717443	0.961708	2.160695
4	0.5	0.202486	0.052433	0.043259	0.000863	0.001232	0.002469	-0.44268
4	0.7	-2.48459	-2.09303	-2.02561	-0.60042	-0.71498	-0.95677	-3.04606
4	0.9	-5.17167	-4.2385	-4.09448	-1.20171	-1.43119	-1.91601	-5.64944
6	0.1	5.564878	4.667756	4.435078	1.203377	1.433572	1.92603	4.92466
6	0.3	2.854277	2.360136	2.237467	0.601976	0.717197	0.963845	2.217333
6	0.5	0.143676	0.052516	0.039857	0.000576	0.000822	0.00166	-0.48999
6	0.7	-2.56692	-2.2551	-2.15775	-0.60083	-0.71555	-0.96053	-3.19732
6	0.9	-5.27752	-4.56272	-4.35536	-1.20223	-1.43193	-1.92271	-5.90465
8	0.1	5.558432	4.897909	4.600604	1.203348	1.433531	1.928584	5.018271
8	0.3	2.83494	2.475233	2.318767	0.60189	0.717074	0.964917	2.246085
8	0.5	0.111448	0.052557	0.036931	0.000432	0.000617	0.00125	-0.5261
8	0.7	-2.61204	-2.37012	-2.24491	-0.60103	-0.71584	-0.96242	-3.29829
8	0.9	-5.33554	-4.7928	-4.52674	-1.20248	-1.4323	-1.92608	-6.07047

10	0.1	5.554355	5.076429	4.719326	1.203331	1.433506	1.93012	5.079791
10	0.3	2.822708	2.564505	2.376862	0.601838	0.717	0.965561	2.262582
10	0.5	0.091062	0.052582	0.034398	0.000345	0.000493	0.001002	-0.55463
10	0.7	-2.64058	-2.45934	-2.30807	-0.60115	-0.71601	-0.96356	-3.37184
10	0.9	-5.37223	-4.97127	-4.65053	-1.20264	-1.43252	-1.92812	-6.18904

Table 7: Values of $\log(\psi)$ for M_3 when $\sigma_y^2 = 2$, $\mu_y = 5$, $\sigma_t^2 = 5$, $\alpha = 25$, $\beta = 25$, $\gamma = 5$, $A = 0.7$, $P = 0.4$, $W = 0.2$, $P_1 = 0.4$, $P_2 = 0.6$.

σ_s^2	w_1	$\log\left(\psi_{(W/M_3)}\right)$	$\log\left(\psi_{(EH/M_3)}\right)$	$\log\left(\psi_{(DP/M_3)}\right)$	$\log\left(\psi_{(GS/M_3)}\right)$	$\log\left(\psi_{(M/M_3)}\right)$	$\log\left(\psi_{(NS/M_3)}\right)$	$\log\left(\psi_{(G/M_3)}\right)$
2	0.1	7.77993	4.34238	4.330598	2.560575	2.636839	2.88727	5.125718
2	0.3	4.06323	2.17485	2.168906	1.280671	1.318843	1.44422	2.21331
2	0.5	0.34654	0.00732	0.007214	0.000767	0.000847	0.001169	-0.6991
2	0.7	-3.37015	-2.16021	-2.15448	-1.27914	-1.31715	-1.44188	-3.6115
2	0.9	-7.08685	-4.32774	-4.31617	-2.55904	-2.63515	-2.88493	-6.52391
4	0.1	7.75116	4.89690	4.873504	2.560498	2.636754	2.927131	5.642591
4	0.3	3.97694	2.45212	2.440316	1.280441	1.318589	1.463873	2.469606
4	0.5	0.20271	0.00733	0.007127	0.000384	0.000424	0.000616	-0.70338
4	0.7	-3.57151	-2.43745	-2.42606	-1.27967	-1.31774	-1.46264	-3.87636
4	0.9	-7.34573	-4.88223	-4.85925	-2.55973	-2.63591	-2.9259	-7.04935
6	0.1	7.73938	5.22127	5.186429	2.560472	2.636726	2.940872	5.931001
6	0.3	3.94160	2.61431	2.59673	1.280364	1.318504	1.470645	2.611833
6	0.5	0.14383	0.00734	0.007032	0.000256	0.000282	0.000418	-0.70734
6	0.7	-3.65395	-2.59963	-2.58267	-1.27985	-1.31794	-1.46981	-4.0265
6	0.9	-7.45173	-5.20659	-5.17237	-2.55996	-2.63616	-2.94004	-7.34567
8	0.1	7.73293	5.45142	5.405287	2.56046	2.636712	2.947832	6.126726
8	0.3	3.92225	2.72938	2.706112	1.280326	1.318462	1.474074	2.707865
8	0.5	0.11156	0.00734	0.006936	0.000192	0.000212	0.000316	-0.711
8	0.7	-3.69912	-2.7147	-2.69224	-1.27994	-1.31804	-1.47344	-4.12986
8	0.9	-7.50981	-5.43674	-5.39141	-2.56008	-2.63629	-2.9472	-7.54872
10	0.1	7.72885	5.62994	5.572672	2.560452	2.636703	2.952037	6.272234
10	0.3	3.91000	2.81864	2.789758	1.280303	1.318436	1.476146	2.778922
10	0.5	0.09115	0.00734	0.006843	0.000153	0.000169	0.000254	-0.71439
10	0.7	-3.7277	-2.80395	-2.77607	-1.28	-1.3181	-1.47564	-4.2077
10	0.9	-7.54655	-5.61525	-5.55899	-2.56015	-2.63636	-2.95153	-7.70101

Table 8: Values of $\log(\psi)$ for M_1 vs M_3, M_2 vs M_3 and M_1 vs M_2 when $\sigma_y^2 = 2$, $\mu_y = 2$, $\sigma_t^2 = 3$, $\alpha = 15$, $\beta = 15$, $\gamma = 510$ $A = 0.7$, $P = 0.6$, $W = 0.2$.

σ_s^2	w_1	M_1 vs M_3	M_2 vs M_3	M_1 vs M_2
2	0.05	4.623235	-0.25893	4.882161
2	0.15	3.612867	-0.20141	3.814282
2	0.3	2.097316	-0.11515	2.212463
2	0.45	0.581765	-0.02888	0.610645
2	0.6	-0.93379	0.057388	-0.99117
2	0.75	-2.44934	0.143656	-2.59299
2	0.9	-3.96489	0.229924	-4.19481

4	0.05	4.619554	-0.25892	4.878474
4	0.15	3.601825	-0.2014	3.803221
4	0.3	2.075232	-0.11511	2.190342
4	0.45	0.548639	-0.02882	0.577463
4	0.6	-0.97795	0.057462	-1.03542
4	0.75	-2.50455	0.143748	-2.6483
4	0.9	-4.03114	0.230035	-4.26118
6	0.05	4.618263	-0.25892	4.877181
6	0.15	3.597954	-0.20139	3.799343
6	0.3	2.067489	-0.1151	2.182586
6	0.45	0.537024	-0.02881	0.565829
6	0.6	-0.99344	0.057487	-1.05093
6	0.75	-2.52391	0.143779	-2.66768
6	0.9	-4.05437	0.230072	-4.28444
8	0.05	4.617605	-0.25892	4.876522
8	0.15	3.595979	-0.20139	3.797366
8	0.3	2.06354	-0.11509	2.178631
8	0.45	0.531101	-0.0288	0.559897
8	0.6	-1.00134	0.057499	-1.05884
8	0.75	-2.53378	0.143795	-2.67757
8	0.9	-4.06622	0.23009	-4.29631
10	0.05	4.617206	-0.25892	4.876122
10	0.15	3.594782	-0.20138	3.796167
10	0.3	2.061145	-0.11509	2.176233
10	0.45	0.527508	-0.02879	0.556299
10	0.6	-1.00613	0.057507	-1.06364
10	0.75	-2.53976	0.143804	-2.68357
10	0.9	-4.0734	0.230101	-4.3035

Table 9: Values of $\log(\psi)$ for M_1 vs M_3 , M_2 vs M_3 and M_1 vs M_2 when $\sigma_y^2 = 2$, $\mu_y = 5$, $\sigma_t^2 = 5$, $\alpha = 25,\ \beta = 25,\ \gamma = 5\ A = 0.7$, $P = 0.4$, $W = 0.2$.

σ_s^2	w_1	M_1 vs M_3	M_2 vs M_3	M_1 vs M_2
2	0.05	5.710012	-0.08578	5.795792
2	0.15	4.445154	-0.06672	4.511872
2	0.3	2.547867	-0.03813	2.585993
2	0.45	0.65058	-0.00953	0.660114
2	0.6	-1.24671	0.019058	-1.26577
2	0.75	-3.14399	0.047651	-3.19164
2	0.9	-5.04128	0.076243	-5.11752
4	0.05	5.709113	-0.08578	5.794892
4	0.15	4.442457	-0.06672	4.509174
4	0.3	2.542472	-0.03812	2.580597
4	0.45	0.642487	-0.00953	0.652019
4	0.6	-1.2575	0.01906	-1.27656
4	0.75	-3.15748	0.047653	-3.20514
4	0.9	-5.05747	0.076245	-5.13371
6	0.05	5.70881	-0.08578	5.794589
6	0.15	4.441546	-0.06672	4.508264
6	0.3	2.540651	-0.03812	2.578776

6	0.45	0.639756	-0.00953	0.649288
6	0.6	-1.26114	0.019061	-1.2802
6	0.75	-3.16203	0.047654	-3.20969
6	0.9	-5.06293	0.076246	-5.13918
8	0.05	5.708657	-0.08578	5.794437
8	0.15	4.441089	-0.06672	4.507806
8	0.3	2.539737	-0.03812	2.577861
8	0.45	0.638384	-0.00953	0.647916
8	0.6	-1.26297	0.019061	-1.28203
8	0.75	-3.16432	0.047654	-3.21197
8	0.9	-5.06567	0.076247	-5.14192
10	0.05	5.708566	-0.08578	5.794345
10	0.15	4.440814	-0.06672	4.507531
10	0.3	2.539187	-0.03812	2.577311
10	0.45	0.637559	-0.00953	0.647091
10	0.6	-1.26407	0.019061	-1.28313
10	0.75	-3.1657	0.047654	-3.21335
10	0.9	-5.06732	0.076247	-5.14357

The numerical results presented in Tables 2–9 provide a comprehensive comparison of the proposed models with existing randomized response techniques using the logarithmic performance measure $\log(\psi)$. A consistent pattern emerges across all configurations: for smaller values of the weighting parameter, the proposed models yield positive values of $\log(\psi)$, indicating clear superiority over competing methods. As the weight increases, the performance advantage gradually diminishes, with a transition occurring in the mid-range where competing models begin to perform comparably or slightly better. This transition reflects the inherent trade-off between efficiency and privacy embedded in the design. The effect of the variance parameter is primarily to scale the magnitude of comparison without altering the overall ranking structure.

A more detailed examination reveals that while certain existing models remain competitive over limited regions, the proposed framework offers broader and more stable dominance across practical parameter ranges. Importantly, the pairwise comparisons among the proposed models (Tables 8 and 9) demonstrate a clear and consistent hierarchy, with M_1 emerging as the most robust performer, followed by M_3, while M_2 exhibits relatively weaker performance. This internal consistency further validates the flexibility and effectiveness of the generalized framework, highlighting its ability to generate models with varying strengths suited to different design priorities.

7. Conclusion and future directions

This study develops a generalized framework for randomized response scrambling models, within which a broad class of existing designs can be expressed and systematically compared. As special cases of this unified structure, three new models, M_1, M_2, and M_3, are proposed and examined in detail. To facilitate a meaningful comparison that accounts for both efficiency and privacy, a logarithmic performance measure, $\log(\psi)$, is introduced. The graphical and analytical results show that the proposed models exhibit distinct but complementary strengths: M_2 and M_3 tend to perform better for smaller values of the weighting parameter, while M_1 becomes increasingly preferable as the weight assigned to the combined performance criterion grows. This behavior highlights that no single model is uniformly optimal; rather, the choice of model should depend on the underlying design priorities. Overall, the proposed framework not only unifies several existing approaches but also provides flexibility, improved interpretability, and practical guidance for selecting an appropriate scrambling mechanism under varying conditions.

There are several directions in which this work can be extended. First, the proposed generalized framework can be further explored by identifying additional classes of scrambling models as special cases, which may offer improved performance under specific design settings. Second, a more rigorous theoretical investigation of the proposed performance measure $\log(\psi)$, particularly its properties as a divergence-type criterion, could provide deeper insight into model comparison and optimality. Third, the framework may be extended to more complex survey settings, such as multi-sensitive attributes, stratified sampling, or adaptive designs, where the balance between efficiency and privacy becomes even more critical. In addition, simulation studies under a wider range of distributions and real data applications would help assess the practical robustness of the proposed models. Finally, developing formal decision rules or optimization strategies for selecting the most appropriate model based on design parameters remains an important area for future research.

References

1. Miller, J. D. (1984). *A new survey technique for studying deviant behavior*. Unpublished Ph.D. dissertation, George Washington University. Department of Sociology.
2. Sirken, M. G. (1970). Household surveys with multiplicity. *Journal of the American statistical Association*, *65*(329), 257-266.
3. Raghavarao, D., & Federer, W. T. (1979). Block total response as an alternative to the randomized response method in surveys. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *41*(1), 40-45.
4. Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American statistical association*, 60(309), 63-69.
5. Warner, S. L. (1971). The linear randomized response model. *Journal of the American Statistical Association*, 66(336), 884-888.
6. Greenberg, B. G., Abul-Ela, A. L. A., Simmons, W. R., & Horvitz, D. G. (1969). The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association*, *64*(326), 520-539.
7. Horvitz, D. G., Shah, B. V., & Simmons, W. R. (1967). The unrelated question randomized response model with unequal probabilities. *Journal of the American Statistical Association*, *71*(354), 311–320.
8. Kuk, A. Y. (1990). Asking sensitive questions indirectly. *Biometrika*, 77(2), 436-438.
9. Mangat, N. S. (1994). An improved randomized response strategy. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(1), 93-95.
10. Christofides, T. C. (2003). A generalized randomized response technique. *Metrika*, 57, 195-200.
11. Kim, J. J., & Warde, W. D. (2005). A stratified randomized response model. *Journal of Statistical Planning and Inference*, *130* (2), 341–349.
12. Arnab, R., & Singh, S**.** (2006). An improved quantitative randomized response model. *Journal of Statistical Planning and Inference*, *136* (8), 2820–2830
13. Gupta, S., Shabbir, J., & Sehra, S. K. (2002). A new randomized response model for quantitative data. *Journal of Statistical Theory and Practice*, *2* (4), 449–458.
14. Eichhorn, B. H., & Hayre, L. S. (1983). Scrambled randomized response methods for obtaining sensitive quantitative data. *Journal of Statistical Planning and inference*, 7 (4), 307-316.
15. Le, T. N., Lee, S. M., Tran, P. L., & Li, C. S. (2023). Randomized response techniques: a systematic review from the pioneering work of warner (1965) to the present. *Mathematics*, 11 (7), 1718.
16. Yan, Z., Wang, J., & Lai, J. (2008). An efficiency and protection degree-based comparison among the quantitative randomized response strategies. *Communications in Statistics-Theory and Methods*, 38 (3), 400-408.
17. Gupta, S., Mehta, S., Shabbir, J., & Khalil, S. (2018). A unified measure of respondent privacy and model efficiency in quantitative RRT models. *Journal of Statistical Theory and Practice*, 12, 506-511.
18. Azeem, M. (2023). Introducing a weighted measure of privacy and efficiency for comparison of quantitative randomized response models. *Pakistan Journal of Statistics*, 39 (3), 377-385.
19. Eichhorn, B. H., & Hayre, L. S. (1983). Scrambled randomized response methods for obtaining sensitive quantitative data. *Journal of Statistical Planning and inference*, 7 (4), 307-316.
20. Diana, G., & Perri, P. F. (2011). A class of estimators for quantitative sensitive data. *Statistical Papers*, 52, 633-650.
21. Gjestvang, C. R., & Singh, S. (2009). An improved randomized response model: Estimation of mean. *Journal of Applied Statistics*, 36 (12), 1361-1367.
22. Murtaza, M., Singh, S., & Hussain, Z. (2020). An innovative optimal randomized response model using correlated scrambling variables. *Journal of Statistical Computation and Simulation*, 90 (15), 2823-2839.
23. Narjis, G., & Shabbir, J. (2021). An efficient new scrambled response model for estimating sensitive population mean in successive sampling. *Communications in Statistics-Simulation and Computation*, 52 (11), 5327-5344.
24. Gupta, S., Zhang, J., Khalil, S., & Sapra, P. (2024). Mitigating lack of trust in quantitative randomized response technique models. *Communications in Statistics-Simulation and Computation*, 53 (6), 2624-2632.

Declaration

Conflict of Study: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contribution Statement: Both the authors worked together as a cell group and designed the research; Analyzed and interpreted the data, and wrote the paper.

Availability of Data and Material: Data will be made available by the corresponding author on reasonable request.

Author(s) Bio / Authors' Note

Shoaib Iqbal *is from the Department of Statistics, Faculty of Computing, The Islamia University of Bahawalpur, 63100, Pakistan. Email:* shoaibiqbalqau654@gmail.com

Zawar Hussain *is from the Department of Statistics, Faculty of Computing, The Islamia University of Bahawalpur, 63100, Pakistan. Email:* zhlangah@yahoo.com

Appendix

Derivation of estimator and its variance under M_1

The motivation for M_1 comes from the randomized response techniques developed by Warner [5] and Gjestvang and Singh [38]. The Model 1 we are proposing here uses a more straightforward one-stage randomized response approach using multiplicative scrambling.

Let us assume that there are n number of respondents in a sample selected from the population. Each respondent is given a randomization device to generate two outcomes, each occurring with known probability *P* and *(1-P)*. If the outcome 1 appears, then $Y(1+S)$ will be reported by the respondent, while if the second outcome, then $Y(1-S)$ will be reported by the respondent. The response of the i^{th} respondent through M_1 is:

$$R_{M_{1i}} = D_{1i}\left(Y_i + Y_iS_i\right) + \left(1 - D_{2i}\right)\left(Y_i - Y_iS_i\right).$$

The distribution of this response may be written as

$$R_{M_{1i}} = \begin{cases} Y_i\left(1+S_i\right) & \text{with probability } P_1 \\ Y_i\left(1-S_i\right) & \text{with probability } 1-P_1 \end{cases}$$

Now, the expected response is given by

$$E\left(R_{M_{1i}}\right) = E\left(D_{1i}\right)\left\{E\left(Y_i\right) + E\left(Y_i\right)E\left(S_i\right)\right\} + E\left(1-D_{1i}\right)\left\{E\left(Y_i\right) - E\left(Y_i\right)E\left(S_i\right)\right\}$$

$$E\left(R_{M_{1i}}\right) = \left(P_1\right)E\left(Y_i\right) + \left(1-P_1\right)E\left(Y_i\right)$$

$$E\left(R_{M_{1i}}\right) = E\left(Y_i\right) = \mu_y$$

So, by methods of moment estimation, an unbiased mean estimator of μ_y is given by:

$$\hat{\mu}_{M_1} = \frac{1}{n}\sum_{i=1}^{n} R_{M_{1i}}.$$

Now, by definition, variance of the above estimator is given by

$$Var\left(\hat{\mu}_{M_1}\right) = Var\left(\frac{1}{n}\sum_{i=1}^{n} R_{M_{1i}}\right) = \frac{Var\left(R_{M_{1i}}\right)}{n}.$$

where $Var\left(R_{M_{1i}}\right)$ is given by

$$Var\left(R_{M_{1i}}\right) = E\left(R^2_{M_{1i}}\right) - \left[E\left(R_{M_{1i}}\right)\right]^2.$$

Now, consider

$$\left(R^2_{M_{1i}}\right) = \left[D_{1i}\left(Y_i + Y_iS_i\right)_i + \left(1-D_{1i}\right)\left(Y_i - Y_iS_i\right)\right]^2$$

After applying expectation,

$$E\left(R^2_{M_{1i}}\right) = E\left[D_{1i}\left(Y_i + Y_iS_i\right) + \left(1-D_{1i}\right)\left(Y_i - Y_iS_i\right)\right]^2$$

$$E\left(R^2_{M_{1i}}\right) = E\left(D^2_{1i}\right)\left\{E\left(Y_i^2\right) + E\left(Y_i^2\right)E\left(S_i^{\ 2}\right)\right\} + E\left(1-D_{1i}\right)^2\left[\left\{E\left(Y_i^2\right) + E\left(Y_i^2\right)E\left(S_i^{\ 2}\right)\right\}\right.$$

$$E\left(R^2_{M_{1i}}\right) = \left(P_1\right)\left\{E\left(Y_i^2\right) + E\left(Y_i^2\right)E\left(S_i^{\ 2}\right)\right\} + \left(1-P_1\right)\left[\left\{E\left(Y_i^2\right) + E\left(Y_i^2\right)E\left(S_i^{\ 2}\right)\right\}\right.$$

$$E\left(R^2_{M_{1i}}\right) = \left(\sigma_y^2 + \mu_y^2\right) + \sigma_s^2\left(\sigma_y^2 + \mu_y^2\right)$$

Thus, the variance of $R_{M_{1i}}$ is given by

$$Var\left(R_{M_{1i}}\right)=E\left(R_{M_{1i}}^2\right)-\left(E\left(R_{M_{1i}}\right)\right)^2$$
$$=\left(\sigma_y^2+\mu_y^2\right)+\sigma_s^2\left(\sigma_y^2+\mu_y^2\right)-\mu_y^2$$
$$=\sigma_y^2+\sigma_s^2\left(\sigma_y^2+\mu_y^2\right)$$

Hence, the variance of $\hat{\mu}_{M_1}$ is given by:

$$Var\left(\hat{\mu}_{M_1}\right)=\frac{1}{n}\left[\ \sigma_y^2+\sigma_s^2\mu_{2y}'\ \right]$$

Derivation of estimator and its variance under M_2

The response of the i^{th} respondent through M_2 is:

$$R_{M_{2i}}=D_{2i}\left(Y_i+\alpha Y_iS_i\right)+\left(1-D_{2i}\right)\left(Y_i-\beta Y_iS_i\right)$$

The distribution of i^{th} respondent response under the M_2 is given as:

$$R_{M_{2i}}=\begin{cases} Y_i\left(1+\alpha S_i\right) & \text{with probability } P_2=\dfrac{\alpha}{\alpha+\beta} \\ Y_i\left(1-\beta S_i\right) & \text{with probability } \left(1-P_2\right)=\dfrac{\beta}{\alpha+\beta} \end{cases}$$

Now, the expected response is given by

$$E\left(R_{M_{2i}}\right)=E\left(D_{2i}\right)\left\{E\left(Y_i\right)+\alpha E\left(Y_i\right)E\left(S_i\right)\right\}+E\left(1-D_{2i}\right)\left\{E\left(Y_i\right)-\beta E\left(Y_i\right)E\left(S_i\right)\right\}$$
$$E\left(R_{M_{2i}}\right)=\left(P_2\right)E\left(Y_i\right)+\left(1-P_2\right)E\left(Y_i\right)$$
$$E\left(R_{M_{2i}}\right)=E\left(Y_i\right)=\mu_y$$

So, by methods of moment estimation, an unbiased mean estimator of μ_y is given by:

$$\hat{\mu}_{M_2}=\frac{1}{n}\sum_{i=1}^{n}R_{M_{2i}}\ .$$

Now, by definition, variance of the above estimator is given by

$$Var\left(\hat{\mu}_{M_2}\right)=Var\left(\frac{1}{n}\sum_{i=1}^{n}R_{M_{2i}}\right)=\frac{Var\left(R_{M_{2i}}\right)}{n}.$$

where $Var\left(R_{M_{2i}}\right)$ is given by

$$Var\left(R_{M_{2i}}\right)=E\left(R_{M_{2i}}^2\right)-\left[E\left(R_{M_{2i}}\right)\right]^2.$$

Now, consider

$$\left(R_{M_{2i}}^2\right)=\left[D_{2i}\left(Y_i+\alpha Y_iS_i\right)+\left(1-D_{2i}\right)\left(Y_i-\beta Y_iS_i\right)\right]^2$$

After applying expectation,

$$E\left(R_{M_{2i}}^2\right)=E\left[D_{2i}\left(Y_i+\alpha Y_iS_i\right)+\left(1-D_{2i}\right)\left(Y_i-\beta Y_iS_i\right)\right]^2$$
$$E\left(R_{M_{2i}}^2\right)=E\left(D_{2i}^2\right)\left\{E\left(Y_i^2\right)+\alpha^2E\left(Y_i^2\right)E\left(S_i^2\right)\right\}+E\left(1-D_{2i}\right)^2\left[\left\{E\left(Y_i^2\right)+\beta^2E\left(Y_i^2\right)E\left(S_i^2\right)\right\}\right.$$
$$E\left(R_{M_{2i}}^2\right)=\left(P_2\right)\left\{E\left(Y_i^2\right)+\alpha^2E\left(Y_i^2\right)E\left(S_i^2\right)\right\}+\left(1-P_2\right)\left[\left\{E\left(Y_i^2\right)+\beta^2E\left(Y_i^2\right)E\left(S_i^2\right)\right\}\right.$$
$$E\left(R_{M_{2i}}^2\right)=\left(\sigma_y^2+\mu_y^2\right)+\left(\sigma_y^2+\mu_y^2\right)\left(P_2\alpha^2+\left(1-P_2\right)\beta^2\right)\sigma_s^2$$

Thus, the variance of $R_{M_{2i}}$ is given by

$$Var\left(R_{M_{2i}}\right)=E\left(R^2_{M_{2i}}\right)-\left(E\left(R_{M_{2i}}\right)\right)^2$$
$$=\left(\sigma_y^2+\mu_y^2\right)+\left(\sigma_y^2+\mu_y^2\right)\left(P_2\alpha^2+\left(1-P_2\right)\beta^2\right)\sigma_s^2-\mu_y^2$$
$$=\sigma_y^2+\mu'_{2y}\left(P_2\alpha^2+\left(1-P_2\right)\beta^2\right)\sigma_s^2$$

Hence, the variance of $\hat{\mu}_{M_2}$ is given by:

$$Var\left(\hat{\mu}_{M_2}\right)=\frac{1}{n}\left[\ \sigma_y^2+\mu'_{2y}\left(P_2\alpha^2+\left(1-P_2\right)\beta^2\right)\sigma_s^2\right]$$

Derivation of estimator and its variance under M_3

The i^{th} respondent's response may be written as

$$R_{M_{3i}}=D_{3i}Y_i\left(1+\alpha S_i\right)+D_{4i}Y_i\left(1-\beta S_i\right)+\left(1-D_{2i}-D_{3i}\right)\left(Y_i\right)$$

And its distribution is given by

$$R_{M_{3i}}=\begin{cases} Y_i\left(1+\alpha S_i\right) & \text{with probability } p_1=\dfrac{\alpha}{\alpha+\beta+\gamma} \\ Y_i\left(1-\beta S_i\right) & \text{with probability } p_2=\dfrac{\beta}{\alpha+\beta+\gamma} \\ Y_i & \text{with probability } p_3=\dfrac{\gamma}{\alpha+\beta+\gamma} \end{cases}$$

Now, the expected response is given by

$$E\left(R_{M_{3i}}\right)=E\left(D_{3i}\right)\left\{E\left(Y_i\right)+\alpha E\left(Y_i\right)E\left(S_i\right)\right\}+E\left(D_{4i}\right)\left\{E\left(Y_i\right)-\beta E\left(Y_i\right)E\left(S_i\right)\right\}+E\left(1-D_{3i}-D_{4i}\right)E\left(Y_i\right)$$
$$E\left(R_{M_{3i}}\right)=\left(p_1\right)E\left(Y_i\right)+\left(p_2\right)E\left(Y_i\right)+\left(p_3\right)E\left(Y_i\right)$$
$$E\left(R_{M_{3i}}\right)=E\left(Y_i\right)=\mu_y$$

So, by methods of moment estimation, an unbiased mean estimator of μ_y is given by:

$$\hat{\mu}_{M_3}=\frac{1}{n}\sum_{i=1}^{n}R_{M_{3i}}\,.$$

Now, by definition, variance of the above estimator is given by

$$Var\left(\hat{\mu}_{M_3}\right)=Var\left(\frac{1}{n}\sum_{i=1}^{n}R_{M_{3i}}\right)=\frac{Var\left(R_{M_{3i}}\right)}{n}\,.$$

where $Var\left(R_{M_{3i}}\right)$ is given by

$$Var\left(R_{M_{3i}}\right)=E\left(R^2_{M_{3i}}\right)-\left[E\left(R_{M_{3i}}\right)\right]^2.$$

Now, consider

$$\left(R^2_{M_{3i}}\right)=\left[D_{3i}Y_i\left(1+\alpha S_i\right)+D_{4i}Y_i\left(1-\beta S_i\right)+\left(1-D_{3i}-D_{4i}\right)\left(Y_i\right)\right]^2$$

After applying expectation,

$$E\left(R^2_{M_{3i}}\right)=E\left[D_{3i}\left(Y_i+\alpha Y_iS_i\right)+\left(D_{4i}\right)\left(Y_i-\beta Y_iS_i\right)+\left(1-D_{3i}-D_{4i}\right)\left(Y_i\right)\right]^2$$
$$E\left(R^2_{M_3}\right)=E\left(D^2_{31}\right)\left\{E\left(Y_i^2\right)+\alpha^2E\left(Y_i^2\right)E\left(S_i^{\ 2}\right)\right\}+E\left(D_{4i}\right)^2\left[\left\{E\left(Y_i^2\right)+\beta^2E\left(Y_i^2\right)E\left(S_i^{\ 2}\right)\right\}+E\left(1-D_{3i}-D_{4i}\right)^2E\left(Y_i^2\right)\right.$$
$$E\left(R^2_{M_{3i}}\right)=\left(p_1\right)\left\{E\left(Y_i^2\right)+\alpha^2E\left(Y_i^2\right)E\left(S_i^{\ 2}\right)\right\}+\left(p_2\right)\left[\left\{E\left(Y_i^2\right)+\beta^2E\left(Y_i^2\right)E\left(S_i^{\ 2}\right)\right\}+\left(p_3\right)E\left(Y_i^2\right)\right.$$
$$E\left(R^2_{M_{3i}}\right)=\left(\sigma_y^2+\mu_y^2\right)+\frac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_y^{\ 2}+\mu_y^{\ 2}\right)\sigma_s^{\ 2}}{\alpha+\beta+\gamma}$$

Thus, the variance of $R_{M_{3i}}$ is given by

$$Var\left(R_{M_{3i}}\right)=E\left(R^2_{M_{3i}}\right)-\left(E\left(R_{M_{3i}}\right)\right)^2$$

$$=\left(\sigma_y^2+\mu_y^2\right)+\frac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_y{}^2+\mu_y{}^2\right)\sigma_s{}^2}{\alpha+\beta+\gamma}-\mu_y^2$$

$$=\sigma_y{}^2+\frac{\alpha\beta\left(\alpha+\beta\right)\mu'_{2y}\sigma_s{}^2}{\alpha+\beta+\gamma}$$

Hence, the variance of $\hat{\mu}_{M_3}$ is given by:

$$Var\left(\hat{\mu}_{M_3}\right)=\frac{1}{n}\left[\ \sigma_y{}^2+\frac{\alpha\beta\left(\alpha+\beta\right)\mu'_{2y}\sigma_s{}^2}{\alpha+\beta+\gamma}\right].$$