symmetry

Article

Estimation of Hidden Logits Using Several Randomized Response Techniques

Asma Halim [1,2], Irshad Ahmad Arshad [1], Abdullah Mohammed Alomair [3,*] and Mohammed Ahmed Alomair [3]

1 Department of Statistics, Allama Iqbal Open University, Islamabad 44000, Pakistan
2 Department of Business Administration, Iqra University, Islamabad 44000, Pakistan
3 Department of Quantitative Methods, School of Business, King Faisal University, Al-Ahsa 31982, Saudi Arabia
* Correspondence: ama.alomair@kfu.edu.sa

Abstract: In survey sampling, we aspire to obtain sound and consistent responses, which are not achieved while dealing with sensitive issues. Frequently, respondents give elusive responses to sensitive questions, so we employ randomized response techniques that facilitate finding an appropriate proportion of socially sensitive characteristics. In the present study, we proposed a hidden logit estimation method using Huang, Warner, and Mangat's randomized response techniques. This study depicts that the estimates become closer to the standard logits as the values of p increase. We found that the hidden logit estimates obtained by the Huang randomized response technique were nearer to the parametric values, in contrast to the other existing techniques, and we demonstrate an increase in accuracy as well. The simulation-based AIC and SIC values are used to assess model performance. We found that the Huang model is the best model for the proposed hidden logit method. This paper contributes towards the application of logistic models in the case of sensitive or socially stigmatized issues.

Keywords: randomized response; logit models; logistic regression; sensitive characteristics; hidden logit estimation

check for updates

Citation: Halim, A.; Arshad, I.A.; Alomair, A.M.; Alomair, M.A. Estimation of Hidden Logits Using Several Randomized Response Techniques. *Symmetry* **2023**, *15*, 1636. https://doi.org/10.3390/sym15091636

Academic Editor: Jinyu Li

Received: 2 August 2023
Revised: 18 August 2023
Accepted: 22 August 2023
Published: 24 August 2023

1. Introduction

Survey sampling is a strategy for gathering data about a specific trait of a population based on a sub-part of that population, when there is limited time or cost to observe each person in the entire population. It involves taking samples from the population, analyzing them, and then using the results of the samples to draw final statements, interpretations, and conclusions about the whole population. Survey sampling is a comprehensive and widely used strategy in various areas of data collection. The most compelling goal in data collection is to obtain precise, predictable, and reliable outcomes.

1.1. Survey Sampling and Randomized Response

In conducting surveys, there are many issues of interest when somebody needs to gather data on sensitive or stigmatized issues. The findings might be misleading if we have sensitive characteristics under study, such as scams, the use of illegal drugs or intoxicating beverages, prohibited or unlawful earnings, evading income tax, reserves in the form of prize bonds, number of induced abortions, exploitation of finances, etc.

The main aim in survey sampling is to obtain accurate and reliable data, which mostly fails in the case of sensitive issues. When surveys are conducted with a direct questioning or interviewing method regarding sensitive issues, it is expected to get ambiguous or false results. Issues like scams, prohibited drugs, and tax evasion are considered as sensitive issues. Let us first discuss some real life examples, in which our variable of interest is considered as sensitive or socially stigmatized.

Racism is a major workplace sensitivity issue that refers to making colleagues feel uncomfortable regarding their background or skin tone. If we want to know about people who perpetrate racism, direct questioning methods will fail to elicit authentic responses. Therefore, racism is a highly sensitive issue, especially in business research, as discussed in detail by Geurts [1]. As another example, if our interest is to assess the proportion of users of illicit medication and alcoholic items in a local area where such substances are completely restricted, respondents may hesitate to reveal their sensitive traits. Thus, if an interviewer conducts a direct inquiry such as "do you smoke or take alcoholic drinks?" the respondents will wonder whether or not to share an accurate response to this delicate inquiry about their drinking status with the interviewer.

Harassment is also one of the sensitive issues where we experience difficulty in collecting data. Harassment can be on the basis of race, religion, gender, or national origin. Workplace harassment, in all its forms, is unacceptable. Mirhosseini et al. [2] presented a descriptive analysis of sexual harassment and its coping strategies. Although connections between coworkers might develop over time, they should always adhere to company policy. A coworker with unprofessional motives should never make an employee feel pressured or uneasy. Beyond interpersonal interactions, harassing, intimidating, or bullying another employee needs to be addressed swiftly and firmly. Therefore, if a person is interested in collecting data about harassment in their workplace, the respondents may not provide them with accurate information.

When we are intrigued to know the proportion of individuals who evade income tax by making payments through contacts or nepotism, sensitive issues may arise. A detailed study on factors persuading taxpayers to engage in tax evasion was done by Kassa [3]. If an income tax officer conducts a survey for such purposes, the respondents will likely lie regarding non-payment of income tax due to their fear of punishment and penalties imposed by the authorities or government accountability. Similarly, if we wish to know about the living standards and comforts of individuals in a specific local area, we have to know about their income. Here, on the off chance that we apply the condition that we know their normal pay, and it does not coordinate with their marvelous expectation for everyday comforts, then we need to know the extent of individuals who are engaged with unlawful pay and illegal income. For the most part, respondents conceal their pay and never want to be asked or questioned about their unlawful pay or additional kinds of revenue. Therefore, they may under-report their illegal means of obtaining income to an interviewer who is a stranger.

The respondent frequently wonders whether or not to answer honestly in the case where respondents are straightforwardly presented with these sorts of sensitive inquiries. The respondents experience dread that either their actual response about the sensitive inquiries being posed would be a reason for humiliation or that they would be ridiculed in the general public. Once in a while, they feel that their truthful reaction might draw punishment or their privacy might be violated. The apprehension that the legal framework can be prompt results in either refusal to answer or in evasive answers. Such a scenario might prompt social desirability bias (SDB). Some of the time, we face such conditions when the study variable is sensitive. Sensitive attributes can be the use of drugs, not paying tax, being involved in illegal activities, etc.

Warner [4] discovered a randomized response (RR) survey model for countering the reservations amongst respondents in the case of susceptible or socially stigmatized inquiries; such a technique is very much needed when we want to obtain reliable and authentic data. This method is very effective at lessening SDB up to a massive degree. To develop the self-assurance of and cater to the confidentiality of the respondents, the unrelated question model is recommended by Greenberg et al. [5]. Some striking work associated with the RR model has been done by a variety of researchers. Let us have a brief discussion on a few of them. Moors [6] altered the model of Greenberg [5] in the case of unknown parameters of the population's characteristics. He also calculated the values for the probabilities p_1 and p_2. In the situation when Moors's [6] model fails, three straightforward alternative RR models are suggested by Mahmood et al. [7]. Upon comparing his proposed estimators with those of Greenberg et al. [5], it is evident that his estimators are more efficient. The work of Christofides [8] advances the groundbreaking work of Warner [4] by presenting an alternate randomized response technique (RRT). He also included Warner's [4] approach in his proposed procedure as a special case. Huang [9] proved that his proposed approach is more efficient than a number of widely used RR approaches. His technique is applicable to direct response surveys as well as in RR surveys when we are striving to obtain genuine responses regarding sensitive issues.

The concept of RR introduced by Warner [4] has also been extended by many researchers to identify the deficiencies and propose solutions in his model. Kim and Warde [10] offered some fresh findings on the RR model, where response variables are presumptively distributed via multinomial distribution. They used Hopkins's test using a randomization device to produce estimates, considering it with and without the assumption of truthful responses. Many scholars have offered numerous alternative ways to address the privacy issue in contrast to the Moors [6] RR model. However, their models might result in a significant loss of data information and incur considerable costs in maintaining secrecy. Compared to earlier RR models, Kim and Warde [11] suggested a simpler model, while still maintaining confidentiality using stratified sampling. The stratified Warner's [4] RR approach and the unrelated question RR model were combined by Kim and Elam to create a new RRT [12]. A three-stage stratified RR approach using optimal allocation was proposed by Kim and Chae [13], which expands upon the two-stage stratified RRT developed by Kim and Elam [12]. They demonstrated their suggested RR estimator to be more effective than Kim and Elam's [12] estimator, but it provided less privacy protection. A new, more efficient RR procedure was sought by Mangat and Singh [14] using two randomization devices. Their model was confusing to the respondents as the respondents had to cater two randomizing devices while responding. Therefore, Mangat [15] presented a simpler technique which was more efficient.

The work of Narjis and Shabbir [16] and Hsieh et al. [17] is also to be noted in the case of the use of two-stage RR models in order to find the commonness of a sensitive characteristic. Singh and Singh [18] offered a technique for finding the population fraction of a stigmatized characteristic, making use of very well-known distribution that is negative binomial distribution for his work. Singh et al. [19] projected a three-stage randomized response model, making use of poisson distribution. Halim et al. [20] derived the transition matrices of the conditional misclassification probabilities of multiple above mentioned models and also worked on finding the association of variables while taking RR into account. Jaiswal et al. [21] projected the calibrated estimator of population mean under a unit response condition using inverse linear, logistic, and exponential integrated models.

1.2. Logistic Regression

When we are working on regression models, our top priority is to approximate the parameters of the mathematical model or the function involved. For this specific goal of parameter estimation, numerous techniques can be adopted. Regression has multiple applications in almost every field. Linear and quadratic multiple regression analysis to find the behavior of certain reactions in the chemistry field was carried out by Falodun et al. [22]. Dorugade [23] introduced new ridge parameters for ridge regression. The Bayesian approach was also used by Ateeq et al. [24] to find the recovery time for the patients of a contagious disease. The ordinary least square (OLS) regression is an incredible asset when the variables of interest are continuous; however, in the case of dichotomous variables, OLS is not valid. A few examples of binary variables are: a head shows up on flipping a coin, an individual smoking or not, a medical test having positive or negative outcomes, an individual having ownership of an industry or not, or a corporation coming to a decision to provide additional benefit to their workers. These referenced models result in a "yes" or "no" reaction. One method to evaluate such dichotomous response variables is logistic regression. Many researchers have utilized this technique of logistic regression in many exploration regions, especially in various subjects of psychological and social examinations, such as those by Clark and Beck [25], Waldman et al. [26], and many others. Not only in the social sciences but in numerous other studies in different fields, the predicted variable is usually stigmatized or sensitive, for example, drug addiction, tax evasion, induced abortion, sexual abuse etc. We express the effect of one variable on other variable(s) in terms of odds ratios in logit estimation.

1.3. Logit Estimation in Randomized Response

We need a procedure which can provide complete privacy to the respondent so that they have no fear of being stigmatized, as stated by Corstange [27]: "If the problem is that people have incentives to hide their true opinions or behavior from the interviewer, then our science suffers unless we can develop means to nullify these incentives. Survey respondents may not be willing to reveal their true answers to sensitive questions without foolproof guarantees of anonymity—not only from outside observers such as law enforcement or friends and family, but even from the interviewers themselves." Ordinary logit models are not suitable when the response dichotomous variable relates to a sensitive issue. Corstange [27,28] projected a method which is known as hidden logit to deal with such problems. The hidden logit model is a customized structure of standard logit which standardizes the outcome of a tool or device which is used for randomization. This technique operates to display the genuine likelihood of a "yes" reply as a function of a predictor variable X. The odd ratio is known as:

$$\ln\left(\frac{\pi}{1-\pi}\right) = X\beta \tag{1}$$

Considering π as the likelihood of a "yes" answer, we work to crack any RR model for π and supplant it in standard logits in order to get the hidden logit model. Utilizing an identical condition, we can discover our estimates of logits using ML methodology. We have to create the logit model form in terms of "X" and "β." According to Corstange [27], the RR model consists of the following methodology: If flipping a coin is the randomizing device and if the coin lands on heads, the respondent is asked to say "yes" without clarification; however, if the coin lands on tails, they are supposed to provide a "yes/no" response according to the actual state they possess. Let us consider π as the probability of absolute "yes" and $p$as the real fraction of participants who truly answer "yes," then the probability of a "yes" answer derived by Corstange [27] is provided as:

$$p(y) = \theta = \pi + (1 - \pi)\, p \tag{2}$$

Solving Equation (2) for π and putting its value in Equation (1) to solve for θ, we obtain:

$$\begin{aligned} \pi &= \tfrac{\theta-p}{1-p} \\ \theta &= \tfrac{e^{X_i\beta}+p}{1+e^{X_i\beta}} \end{aligned} \tag{3}$$

Let us regard "y_i" as a dichotomous variable, for which "1" represents a "yes" response and "0" represents a "no" reply. Subsequently the likelihood function of β is provided as:

$$L(\beta|y_i) = \prod_{i=1}^{n} \theta_i^{y_i}(1-\theta_i)^{1-y_i}, \tag{4}$$

The first derivative of Equation (4) is

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{n}\left[y_i\left\{\frac{e^{X_i\beta}}{(p+e^{X_i\beta})}\right\} - (1+e^{-X_i\beta})^{-1}\right]X_i \tag{5}$$

Setting Equation (5) as equivalent to zero maximizes this articulation, yet we cannot solve it scientifically. Therefore, to measure the parameters, this equation is settled numerically.

Hussain and Shabbir [29] and Hussain et al. [30] also employed different RRTs in order to calculate the hidden logits. The same has been accomplished by Halim et al. [31] for Mangat and Singh [14] RRT. Cruff et al. [32] and Chang et al. [33] also worked on logistic regression in different capacities. Hsieh and Perri [34] worked on finding more advanced approaches to check the components that are faced by researchers dealing with two stigmatized variates, taking RRT into account. Our study helps to find the estimates of logistic models in the case of sensitive issues when there is a complex RRT, like in the case of Huang [9]. This study significantly advances the use and evaluation of logistic models in situations when it is challenging to get sincere responses.

The rest of this article has the following sections. In Section 2, the proposed methodology of hidden logit is discussed using three RRTs. Section 3 presents the results and discussions using simulation. The last section states some concluding remarks.

2. Proposed Hidden Logits using Randomized Response Technique

In the past, many researchers have worked on finding the proportion of "yes" responses while working with sensitive or stigmatized issues. Generally, they have represented the probability of a "yes" response in the form of θ and used different RR devices to calculate it. We propose a methodology that involves finding the π of three RRTs and then incorporating it into the log–odd ratio to determine the value of θ, which is called the hidden logit. It is then used to find the estimates and standard errors of hidden logits and compare them with ordinary logits.

2.1. Proposed Hidden Logit Using Huang [9]

Huang [9] launched a survey technique which was very undemanding to approximate the sensitivity of survey inquiry. His recommended technique is better used to compute a fraction, when the participant of the underlying study responds fairly about possessing sensitive characteristic. One point to be noted is that this technique is also equally applied in the case of direct response surveys as in the case of randomized response surveys in order to get authentic responses. His suggested method has been shown to be more effective than several established RR methods. In the case that the respondent chooses "no," the individual is provided with a randomization device with two explanations of having a place or not having a place in a sensitive group possessing certain probabilities, p and $(1 - p)$. Regardless on the small chance that a direct or a randomized response process is utilized, the respondents will not tell a lie. In this example, it is imaginable for the interviewer that respondents will answer sincerely using the provided RR device in the case of a sensitive group and also in the case of the standard direct response procedure. Here, p is the population proportion of individuals who belong to sensitive group "A." Let "T" be the probability that the respondents belonging to sensitive group "A" report the truth. The proportion of a "yes" answer is provided as:

$$p\,(y) = \theta = p\,\pi\,(1 - T) + (1 - p)\,(1 - \pi) \tag{6}$$

Solving Equation (1) for π, we obtain:

$$\begin{aligned} \theta &= p\,\pi - p\,\pi\,T + 1 - p - \pi + p\,\pi \\ \theta &= \pi\,(p - p\,T + p - 1) + 1 - p \\ \pi\,(p - p\,T + p - 1) &= \theta - 1 + p \\ \pi &= \tfrac{\theta - 1 + p}{2p - pT - 1}. \end{aligned} \tag{7}$$

Now we substitute the value of π from (7) in ordinary logits (1) and solve for θ:

$$
\begin{aligned}
\ln\left[\frac{\theta-1+p}{2p-pT-1}\Big/1-\frac{\theta-1+p}{2p-pT-1}\right] &= X_i\beta \\
\ln\left[\frac{\theta-1+p}{p-pT-\theta}\right] &= X_i\beta \\
\theta\left(1+e^{X_i\beta}\right) &= (p-pT)e^{X_i\beta}+1-p \\
\theta &= \frac{(p-pT)e^{X_i\beta}+1-p}{1+e^{X_i\beta}}
\end{aligned} \tag{8}
$$

We can see that, for $p = 1$, $T = 0$, Equation (8) behaves as the ordinary logits. Further calculations to work on the likelihood function are done as shown below.

Let us first find $\partial\theta_i/\partial\beta$:

$$
\begin{aligned}
\frac{\partial\theta_i}{\partial\beta} &= \frac{\left(1+e^{X_i\beta}\right)\frac{\partial}{\partial\beta}\left\{(p-pT)e^{X_i\beta}+1-p\right\}-\left\{(p-pT)e^{X_i\beta}+1-p\right\}\frac{\partial}{\partial\beta}\left(1+e^{X_i\beta}\right)}{\left(1+e^{X_i\beta}\right)^2} \\
\frac{\partial\theta_i}{\partial\beta} &= \frac{e^{X_i\beta}X_i(2p-pT-1)}{\left(1+e^{X_i\beta}\right)^2}.
\end{aligned} \tag{9}
$$

Replacing $\partial\theta_i/\partial\beta$ from (9) and θ from (8) in the derivative of log likelihood function $\partial \ln L/\partial\beta$ from (5), we get:

$$
\begin{aligned}
&= \sum_{i=1}^{n}\left\{\frac{e^{X_i\beta}X_i(2p-pT-1)}{\left(1+e^{X_i\beta}\right)}\left[\frac{y_i}{(p-pT)e^{X_i\beta}+1-p}+\frac{y_i}{1+e^{X_i\beta}-(p-pT)\,e^{X_i\beta}-1+p}-\frac{1}{1+e^{X_i\beta}-(p-pT)\,e^{X_i\beta}-1+p}\right]\right\} \\
\frac{\partial \ln L}{\partial\beta} &= \sum_{i=1}^{n}\left\{\frac{e^{X_i\beta}X_i(2p-pT-1)}{\left\{1+e^{X_i\beta}-(p-pT)\,e^{X_i\beta}-1+p\right\}}\left[\frac{y_i}{\left\{(p-pT)e^{X_i\beta}+1-p\right\}}-\frac{1}{\left(1+e^{X_i\beta}\right)}\right]\right\}
\end{aligned} \tag{10}
$$

Setting Equation (10) as equivalent to zero maximizes this articulation, yet we cannot solve it scientifically. Therefore, to gauge the parameters, this equation is settled numerically.

2.2. Proposed Hidden Logit Using Warner [4]

Warner [4] initiated the RR technique to use with socially stigmatized or sensitive characteristics. He employed a spinner as a RR device, consisting of two mutually exclusive groups. He used a spinner with two proclamations of fitting in to group "A" or to group "B." Here, "A" is the cluster of people possessing a susceptible attribute. In RR, such a security assurance is provided so that the respondent simply responds with a "yes" or "no" relying upon his standing of having a susceptible attribution. Here p and $(1-p)$ are the respective probabilities of whether the spinner will point to the first or second statement. The probability of a "yes" response is given as:

$$p\,(y) = \theta = p\,\pi + (1-p)\,(1-\pi) \tag{11}$$

Getting the value of π from (11), we obtain:

$$\pi = \frac{\theta-(1-p)}{2p-1} \tag{12}$$

Now replacing value of π from (12) in ordinary logits (1) and solving for θ, we obtain:

$$
\begin{aligned}
\ln\left[\frac{\theta-(1-p)}{2p-1}\Big/1-\frac{\theta-(1-p)}{2p-1}\right] &= X_i\beta \\
\ln\left[\frac{\theta-1+p}{p-\theta}\right] &= X_i\beta \\
\theta-1+p &= (p-\theta)e^{X_i\beta} \\
\theta-1+p &= pe^{X_i\beta}-\theta e^{X_i\beta} \\
\theta &= \frac{pe^{X_i\beta}+1-p}{1+e^{X_i\beta}}
\end{aligned} \tag{13}
$$

For $p = 1$, Equation (13) behaves as the standard logits. Further calculations to work on the likelihood function are done as shown below.

Let us first find $\partial\theta_i/\partial\beta$

$$\begin{aligned}
\frac{\partial\theta_i}{\partial\beta} &= \frac{(1+e^{X_i\beta})\frac{\partial}{\partial\beta}\left(pe^{X_i\beta}+1-p\right)-\left(pe^{X_i\beta}+1-p\right)\frac{\partial}{\partial\beta}(1+e^{X_i\beta})}{\left(1+e^{X_i\beta}\right)^2} \\
\frac{\partial\theta_i}{\partial\beta} &= \frac{pe^{X_i\beta}X_i+pe^{2X_i\beta}X_i-pe^{2X_i\beta}X_i-e^{X_i\beta}X_i+pe^{X_i\beta}X_i}{\left(1+e^{X_i\beta}\right)^2} \\
\frac{\partial\theta_i}{\partial\beta} &= \frac{(2p-1)X_ie^{X_i\beta}}{\left(1+e^{X_i\beta}\right)^2}
\end{aligned} \tag{14}$$

Replacing $\partial\theta_i/\partial\beta$ from (14) and θ from (13) in the derivative of log likelihood function $\partial\ln L/\partial\beta$ from (5), we get:

$$\begin{aligned}
\frac{\partial\ln L}{\partial\beta} &= \sum_{i=1}^{n}\left\{y_i\frac{1}{\left(\frac{pe^{X_i\beta}+1-p}{1+e^{X_i\beta}}\right)}\frac{(2p-1)X_ie^{X_i\beta}}{\left(1+e^{X_i\beta}\right)^2}+(1-y_i)\frac{1}{\left(1-\left(\frac{pe^{X_i\beta}+1-p}{1+e^{X_i\beta}}\right)\right)}(-1)\frac{(2p-1)X_ie^{X_i\beta}}{\left(1+e^{X_i\beta}\right)^2}\right\} \\
&= \sum_{i=1}^{n}\left\{\frac{(2p-1)X_ie^{X_i\beta}}{\left(e^{X_i\beta}(1-p)+p\right)}\left(\frac{y_i}{\left(pe^{X_i\beta}+1-p\right)}-\frac{1}{\left(1+e^{X_i\beta}\right)}\right)\right\} \\
\frac{\partial\ln L}{\partial\beta} &= \sum_{i=1}^{n}\left\{\frac{(2p-1)X_ie^{X_i\beta}}{\left(e^{X_i\beta}(1-p)+p\right)}\left(\frac{y_i}{\left(pe^{X_i\beta}+1-p\right)}-\frac{1}{\left(1+e^{X_i\beta}\right)}\right)\right\}
\end{aligned} \tag{15}$$

2.3. Proposed Hidden Logit Using Mangat [15]

The two-stage RRT utilized by Mangat and Singh [14] has been criticized by Mangat [15] for perhaps misleading the respondents when they are reporting. Therefore, Mangat [15] suggested a less complex method to address this problem. Respondents in the suggested technique are instructed to employ Warner's [4] randomized device with fitting in or no in group A, and they must respond "yes" or "no" as per the outcomes of the randomized tool and their genuine standing. The probability of a "yes" response is provided as:

$$p\,(y) = \theta = \pi + (1-\pi)\,(1-p) \tag{16}$$

Solving Equation (16) for the value of π, we obtain:

$$\pi = \frac{\theta-(1-p)}{p} \tag{17}$$

Putting the value of π in standard logit as defined in Equation (1) and working for θ, we obtain:

$$\begin{aligned}
\frac{\theta-1+p}{1-\theta} &= e^{X_i\beta} \\
\theta+\theta e^{X_i\beta} &= e^{X_i\beta}+1-p \\
\theta &= \frac{e^{X_i\beta}+1-p}{1+e^{X_i\beta}}
\end{aligned} \tag{18}$$

For $p = 1$, Equation (18) behaves as the standard logits. Further calculations to work on the likelihood function are done as shown below.

Let us first find $\partial\theta_i/\partial\beta$

$$\begin{aligned}
\frac{\partial\theta_i}{\partial\beta} &= \frac{(1+e^{X_i\beta})\frac{\partial}{\partial\beta}\left(e^{X_i\beta}+1-p\right)-\left(e^{X_i\beta}+1-p\right)\frac{\partial}{\partial\beta}(1+e^{X_i\beta})}{\left(1+e^{X_i\beta}\right)^2} \\
&= \frac{\left(1+e^{X_i\beta}\right)e^{X_i\beta}X_i-\left(e^{X_i\beta}+1-p\right)e^{X_i\beta}X_i}{\left(1+e^{X_i\beta}\right)^2} \\
&= \frac{e^{X_i\beta}X_i+e^{2X_i\beta}X_i-e^{2X_i\beta}X_i-e^{X_i\beta}X_i+pe^{X_i\beta}X_i}{\left(1+e^{X_i\beta}\right)^2} \\
\frac{\partial\theta i}{\partial\beta} &= \frac{pX_ie^{X_i\beta}}{\left(1+e^{X_i\beta}\right)^2}
\end{aligned} \tag{19}$$

Replacing $\partial\theta_i/\partial\beta$ from (19) and θ from (18)in the derivative of the log likelihood function $\partial \ln L/\partial\beta$ from (5), we get:

$$
\begin{aligned}
\frac{\partial \ln L}{\partial \beta} &= \sum_{i=1}^{n}\left[y_i \frac{1}{\theta_i}\frac{\partial\theta_i}{\partial\beta} + (1-y_i)\frac{1}{1-\theta_i}(-1)\frac{\partial\theta_i}{\partial\beta}\right] \\
&= \sum_{i=1}^{n}\left[y_i \frac{1}{\left(\frac{e^{X_i\beta}+1-p}{1+e^{X_i\beta}}\right)}\frac{pX_ie^{X_i\beta}}{\left(1+e^{X_i\beta}\right)^2} + (1-y_i)\frac{1}{\left(1-\left(\frac{e^{X_i\beta}+1-p}{1+e^{X_i\beta}}\right)\right)}(-1)\frac{pX_ie^{X_i\beta}}{\left(1+e^{X_i\beta}\right)^2}\right] \\
&= \sum_{i=1}^{n}\left[y_i \frac{pX_ie^{X_i\beta}}{\left(e^{X_i\beta}+1-p\right)\left(1+e^{X_i\beta}\right)} - \frac{(1-y_i)e^{X_i\beta}X_i}{\left(1+e^{X_i\beta}\right)}\right] \\
&= \sum_{i=1}^{n}\left[y_i\left\{\frac{p+e^{X_i\beta}+1-p}{\left(e^{X_i\beta}+1-p\right)\left(1+e^{-X_i\beta}\right)}\right\} - \left(1+e^{-X_i\beta}\right)^{-1}\right] \\
\frac{\partial \ln L}{\partial \beta} &= \sum_{i=1}^{n}\left[y_i\left\{\frac{e^{X_i\beta}}{\left(e^{X_i\beta}+1-p\right)}\right\} - \left(1+e^{-X_i\beta}\right)^{-1}\right]X_i
\end{aligned}
\tag{20}
$$

As discussed earlier, setting Equations (5), (10), (15) and (20) as equivalent to zero maximizes them, yet we cannot solve them scientifically; therefore, to gauge the parameters, this equation is settled numerically.

3. Simulation Study

For empirical illustration, using the "Eviews" software, sample sizes of 1000 are produced, setting different values of p and T. A model for three regressors is considered for simulation. The data are generated by taking uniform distribution for parameters ranging from -3 to 3. Hence, the logit estimation is carried out using the starting values of β, set as (0, 1, 1, 1), where 0 is a constant or intercept term and 1's are the estimates of β.

This section consists of three parts. In each part, the proposed methodology is adopted to get the estimates of the β's which are represented by b's, with their standard errors, Akaike Information Criterion (AIC), and Schwarz Information Criterion (SIC) values for all three RRTs, respectively.

The AIC and SIC are among the best methods for model selection. They serve as principles for selecting the best model from among many models of interest. The model with the minimum AIC among all the models is considered a good model, indicating that a lower AIC value signifies a better fit. AIC can be calculated using the following formula:

$$\text{AIC} = -2\ \ln(L) + 2k \tag{21}$$

SIC is closely interrelated to AIC but not the same.SIC can be calculated using the following formula:

$$\text{SIC} = -2\ \ln(L) + 2\ \ln\ (N)k \tag{22}$$

3.1. Estimates Using Huang [9] RRT

In the first part, we conduct assessment of the proposed hidden logit model by Huang [9] for various values of p and T. We utilize the proposed hidden logit structure delivered in Equation (8) for this purpose. Tables 1–3 and Figures 1–3 exhibit the estimation of the b's, standard errors (SE), AIC, and SIC values and their graphs, respectively. In Tables 1 and 2, all values show the standard logits, which are acquired by taking p = 1 and T = 0 in the noted RRT.

Table 1. Estimates of b's for different values of p and T using Huang [9] RRT.

T		p									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
0.1	b_0	−0.2315	−0.3557	−0.4683	−0.5968	−37.6483	3.8144	6.7190	0.1959	0.1495	0.2143
	b_1	1.1506	1.2740	1.0982	0.8187	12.9302	26.0445	51.5573	1.1533	1.0205	1.3130
	b_2	1.0181	1.0824	0.9519	0.8025	20.7750	26.5393	52.6652	1.1957	1.0487	1.3005
	b_3	1.1022	1.1938	1.0767	0.8287	8.8708	24.8303	49.4321	1.0586	0.9761	1.2327
0.2	b_0	0.2198	0.3139	0.4217	0.8703	29.0133	−0.9106	−7.7908	−0.5114	−0.4490	−0.2226
	b_1	0.9557	0.8826	0.9293	1.1556	17.8077	22.9154	82.8046	1.9142	1.3616	1.0345
	b_2	0.8906	0.8499	0.9187	1.2254	14.3520	12.8654	51.2237	1.5453	1.2430	0.9949
	b_3	0.9205	0.9042	0.9822	1.1701	13.8383	18.9082	67.6219	1.7264	1.2706	0.9998
0.3	b_0	0.0532	0.0940	0.1065	0.1578	35.5098	−17.8329	−0.6050	−0.2031	−0.0517	−0.1252
	b_1	1.0789	1.1320	1.0845	1.1173	68.2762	3.1502	1.3498	0.9860	1.0210	0.9978
	b_2	0.9792	0.9799	0.8830	0.8698	36.0714	21.9152	1.5595	0.9426	1.0222	1.0171
	b_3	1.0858	1.0905	1.0995	1.2533	75.5237	20.4426	1.5491	0.8958	1.0067	0.9436
0.4	b_0	−0.0372	0.0001	0.0677	0.1156	−0.1714	2.0771	0.0293	0.0828	0.1007	0.1844
	b_1	0.9027	0.8830	0.9823	0.8133	1.3839	20.1394	0.4858	0.6307	0.7131	0.9475
	b_2	0.9873	0.9387	1.0575	0.8932	1.5794	17.1867	0.1139	0.4789	0.6378	0.9445
	b_3	0.8793	0.8528	0.9375	0.7474	1.0533	16.5136	0.2683	0.5596	0.7022	0.9939
0.5	b_0	0.1453	0.2506	0.2688	0.1421	0.4587	45.6701	−4.7114	−15.1090	−0.0787	−0.1730
	b_1	1.0350	1.0837	1.0169	1.1230	0.8571	5.3868	5.1990	30.7226	1.2067	1.0224
	b_2	1.0104	1.0218	0.9685	1.0063	0.9455	12.0369	21.7138	48.5209	1.3344	1.0162
	b_3	0.9213	0.9435	0.9463	1.0235	0.8549	39.1664	18.5134	47.7135	1.2549	1.0472

Table 1. *Cont.*

T		p									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
0.6	b_0	0.0367	0.0177	0.1599	0.2048	0.2079	0.0932	−29.0775	−16.9430	−0.2394	−0.0752
	b_1	0.9194	0.9406	0.9292	0.8878	1.0052	0.7331	−8.4218	20.5701	0.9922	0.9680
	b_2	0.9765	0.9928	0.9477	0.8344	0.9076	0.9249	16.6294	38.8792	1.2645	0.9397
	b_3	0.8855	0.8885	0.7763	0.7206	0.7477	0.5576	5.7839	50.1951	1.4359	0.9764
0.7	b_0	−0.0191	−0.0516	−0.1179	−0.1626	0.0271	0.2667	1.2094	−31.4399	−0.1193	0.0174
	b_1	0.9302	0.9399	0.9953	0.9284	0.9511	0.5609	−0.2200	25.3507	0.8590	1.0268
	b_2	1.0065	1.1064	1.1523	1.0652	1.2451	1.0934	0.9642	−14.2620	0.2901	1.0739
	b_3	0.9348	0.9873	1.0521	0.9505	0.9456	1.0024	0.6781	−3.4737	0.6145	1.0780
0.8	b_0	−0.0824	−0.0556	−0.1036	−0.0945	−0.1405	−0.3094	0.0584	−24.7924	17.0574	−0.0551
	b_1	1.1424	1.0254	1.0048	0.9978	1.0100	1.2652	1.0481	17.3721	26.5109	0.9780
	b_2	1.0503	0.9605	0.9658	0.8524	0.7711	0.9920	0.3900	−7.9325	27.7356	1.1377
	b_3	1.1869	1.1363	1.1683	1.0326	1.0402	1.2088	0.9503	25.8390	29.4190	1.0852
0.9	b_0	0.1479	0.2048	0.1076	0.0432	−0.0520	−0.1903	−0.1274	0.2546	25.1478	−0.0819
	b_1	1.2641	1.0087	1.1408	1.1645	1.4097	1.5891	1.2565	1.4464	17.7483	0.8769
	b_2	0.9889	0.7779	0.8848	0.8781	0.9722	1.0113	0.7659	0.4046	−8.7823	1.0096
	b_3	1.1542	0.8877	1.0750	1.0892	1.3924	1.3874	1.0369	0.8891	20.9405	0.8866

Table 2. The $s.e(b's)$ for different values of p and T using Huang [9] RRT.

T		p									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
0.1	$s.e(b_0)$	0.1187	0.1841	0.2789	0.4818	333.4179	14.2246	16.2981	0.2226	0.1429	0.0829
	$s.e(b_1)$	0.1309	0.2142	0.2832	0.3892	115.1187	98.1811	132.2272	0.2454	0.1452	0.1006
	$s.e(b_2)$	0.1213	0.1916	0.2574	0.3840	184.2341	100.0052	134.6427	0.2550	0.1498	0.1015
	$s.e(b_3)$	0.1263	0.2043	0.2788	0.3860	79.1464	93.9812	126.9071	0.2286	0.1390	0.0953
0.2	$s.e(b_0)$	0.1063	0.1512	0.2376	0.5328	118.9557	6.8797	21.1552	0.3198	0.1807	0.0763
	$s.e(b_1)$	0.1073	0.1425	0.2260	0.5049	72.4404	112.2464	237.2196	0.5280	0.2260	0.0815
	$s.e(b_2)$	0.0993	0.1343	0.2165	0.5120	58.1661	64.1223	147.0931	0.4290	0.2049	0.0769
	$s.e(b_3)$	0.1029	0.1419	0.2297	0.5038	56.4476	92.9373	193.2615	0.4720	0.2089	0.0779
0.3	$s.e(b_0)$	0.1040	0.1462	0.2074	0.3421	174.6854	319.9981	0.7528	0.2780	0.1812	0.0749
	$s.e(b_1)$	0.1161	0.1692	0.2327	0.3915	332.9167	57.3642	0.8839	0.2870	0.1941	0.0786
	$s.e(b_2)$	0.1074	0.1509	0.2001	0.3258	174.9152	388.2287	0.9772	0.2749	0.1921	0.0786
	$s.e(b_3)$	0.1138	0.1606	0.2286	0.4135	367.7805	363.6587	0.9675	0.2645	0.1882	0.0742
0.4	$s.e(b_0)$	0.1049	0.1407	0.2011	0.2766	0.6394	23.0530	0.6072	0.2962	0.1946	0.0759
	$s.e(b_1)$	0.0997	0.1319	0.2026	0.2466	0.8380	199.2709	0.4059	0.2276	0.1604	0.0745
	$s.e(b_2)$	0.1077	0.1394	0.2168	0.2636	0.9542	170.3881	0.3477	0.2054	0.1527	0.0749
	$s.e(b_3)$	0.1003	0.1319	0.2000	0.2387	0.6852	163.3479	0.3764	0.2206	0.1615	0.0778
0.5	$s.e(b_0)$	0.1016	0.1365	0.1840	0.2700	0.4326	203.3125	34.4337	63.9027	0.3159	0.0776
	$s.e(b_1)$	0.1064	0.1461	0.1872	0.3016	0.3746	23.4865	37.7949	136.7667	0.3703	0.0800
	$s.e(b_2)$	0.1036	0.1391	0.1795	0.2756	0.3957	53.2446	156.7400	218.1805	0.3911	0.0782
	$s.e(b_3)$	0.0953	0.1291	0.1732	0.2740	0.3651	174.0958	133.5330	215.8915	0.3695	0.0784

Table 2. *Cont.*

T		p									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
0.6	$s.e(b_0)$	0.0995	0.1295	0.1678	0.2222	0.3430	0.5897	442.9549	72.1898	0.4089	0.0754
	$s.e(b_1)$	0.0962	0.1274	0.1611	0.2043	0.3436	0.4968	128.3952	86.8768	0.4120	0.0758
	$s.e(b_2)$	0.1008	0.1328	0.1657	0.2005	0.3281	0.5698	252.7913	165.1553	0.4913	0.0742
	$s.e(b_3)$	0.0941	0.1230	0.1452	0.1832	0.2890	0.4395	88.0650	212.9795	0.5411	0.0762
0.7	$s.e(b_0)$	0.0946	0.1223	0.1600	0.2052	0.2877	0.4362	1.3614	242.1730	0.4177	0.0747
	$s.e(b_1)$	0.0956	0.1253	0.1702	0.2063	0.2983	0.3275	0.5764	194.6792	0.3729	0.0811
	$s.e(b_2)$	0.1017	0.1409	0.1899	0.2281	0.3613	0.4722	0.9774	109.9377	0.2631	0.0846
	$s.e(b_3)$	0.0970	0.1302	0.1771	0.2104	0.3006	0.4460	0.7676	27.0893	0.3273	0.0845
0.8	$s.e(b_0)$	0.0992	0.1204	0.1516	0.1852	0.2432	0.3849	0.5409	227.5091	79.7597	0.0745
	$s.e(b_1)$	0.1116	0.1252	0.1552	0.1884	0.2480	0.4534	0.5690	158.6066	123.6629	0.0762
	$s.e(b_2)$	0.1071	0.1224	0.1548	0.1725	0.2116	0.3822	0.3542	73.3429	130.3087	0.0864
	$s.e(b_3)$	0.1151	0.1353	0.1740	0.1921	0.2511	0.4335	0.5302	235.3178	137.0930	0.0830
0.9	$s.e(b_0)$	0.1023	0.1142	0.1436	0.1736	0.2332	0.3106	0.3801	0.6612	618.0020	0.0743
	$s.e(b_1)$	0.1256	0.1170	0.1630	0.2009	0.3172	0.4624	0.4653	0.8783	436.6087	0.0697
	$s.e(b_2)$	0.1024	0.0973	0.1334	0.1610	0.2313	0.3144	0.3238	0.4338	221.7021	0.0752
	$s.e(b_3)$	0.1159	0.1068	0.1549	0.1904	0.3131	0.4154	0.4088	0.6551	516.2336	0.0696

Table 3. AIC and SIC for different values of p and T using Huang [9] RRT.

T		p									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
0.1	AIC	0.9767	1.1885	1.3493	1.4232	1.4530	1.4396	1.4026	1.3071	1.1642	0.5164
	SIC	0.9963	1.2081	1.3689	1.4428	1.4726	1.4592	1.4222	1.3268	1.1839	0.5360
0.2	AIC	1.0077	1.2162	1.3387	1.4087	1.4220	1.4119	1.3403	1.2590	1.1075	0.6667
	SIC	1.0273	1.2358	1.3584	1.4283	1.4416	1.4316	1.3600	1.2787	1.1271	0.6863
0.3	AIC	0.9566	1.1628	1.2974	1.3738	1.4119	1.4231	1.3734	1.3128	1.1941	0.7118
	SIC	0.9763	1.1824	1.3170	1.3934	1.4316	1.4427	1.3931	1.3324	1.2138	0.7315
0.4	AIC	1.0285	1.2259	1.3075	1.3710	1.3977	1.3966	1.3772	1.3143	1.2241	0.6822
	SIC	1.0482	1.2455	1.3271	1.3906	1.4173	1.4162	1.3969	1.3340	1.2437	0.7019
0.5	AIC	0.9004	1.0516	1.1936	1.2898	1.3242	1.3399	1.3182	1.2662	1.1505	0.6235
	SIC	0.9200	1.0712	1.2133	1.3095	1.3438	1.3595	1.3379	1.2858	1.1701	0.6431
0.6	AIC	0.9540	1.1140	1.2320	1.3010	1.3250	1.3193	1.2913	1.1898	1.0682	0.6666
	SIC	0.9737	1.1336	1.2516	1.3206	1.3446	1.3390	1.3109	1.2094	1.0878	0.6862
0.7	AIC	0.9364	1.0801	1.1803	1.2579	1.2525	1.2052	1.1275	1.0050	0.9088	0.6575
	SIC	0.9560	1.0997	1.2000	1.2776	1.2721	1.2249	1.1472	1.0247	0.9284	0.6771
0.8	AIC	0.8390	1.0082	1.1159	1.1739	1.1951	1.1587	1.0615	0.9650	0.6606	0.5955
	SIC	0.8586	1.0278	1.1355	1.1936	1.2147	1.1783	1.0812	0.9847	0.6803	0.6151
0.9	AIC	0.8862	1.0515	1.0933	1.1132	1.0906	1.0295	0.9075	0.6574	0.3312	0.6730
	SIC	0.9059	1.0711	1.1129	1.1328	1.1102	1.0491	0.9272	0.6770	0.3509	0.6926

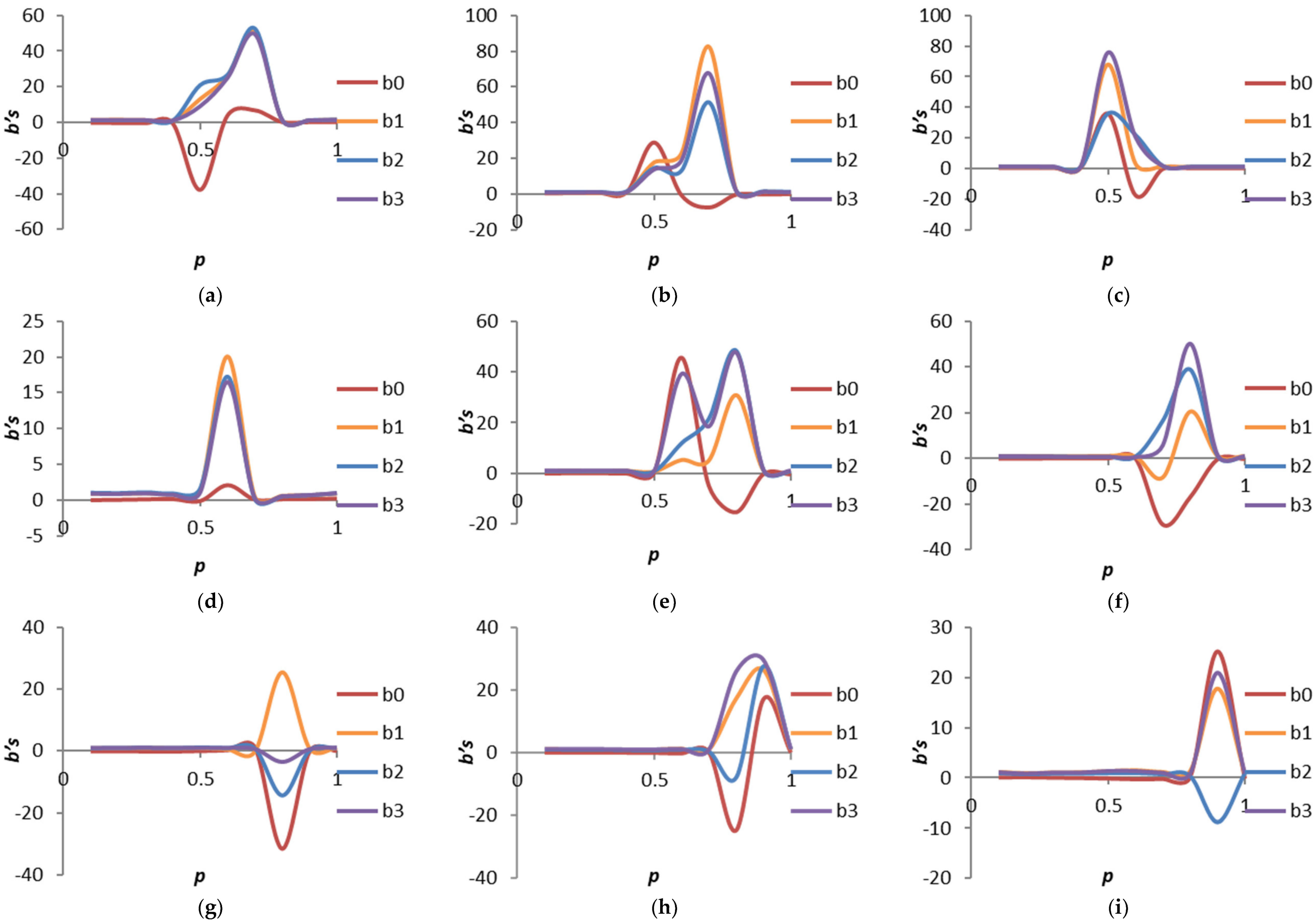

Figure 1. Estimates of b's using Huang [9] RRT. (**a**) T = 0.1; (**b**) T = 0.2; (**c**) T = 0.3; (**d**) T = 0.4; (**e**) T = 0.5; (**f**) T = 0.6; (**g**) T = 0.7; (**h**) T = 0.8; (**i**) T = 0.9.

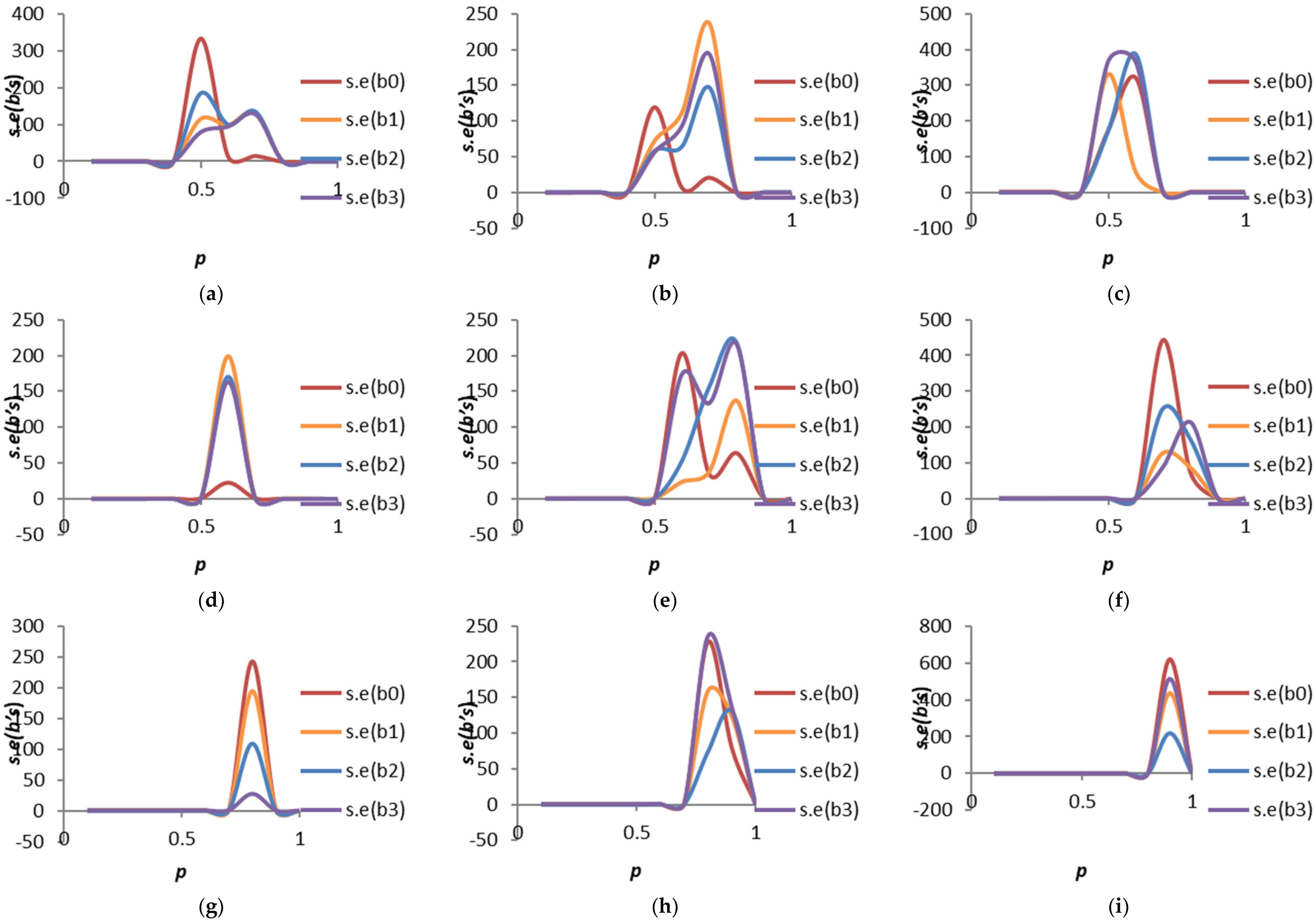

Figure 2. The $s.e(b's)$ using Huang [9] RRT. (**a**) $T = 0.1$; (**b**) $T = 0.2$; (**c**) $T = 0.3$; (**d**) $T = 0.4$; (**e**) $T = 0.5$; (**f**) $T = 0.6$; (**g**) $T = 0.7$; (**h**) $T = 0.8$; (**i**) $T = 0.9$.

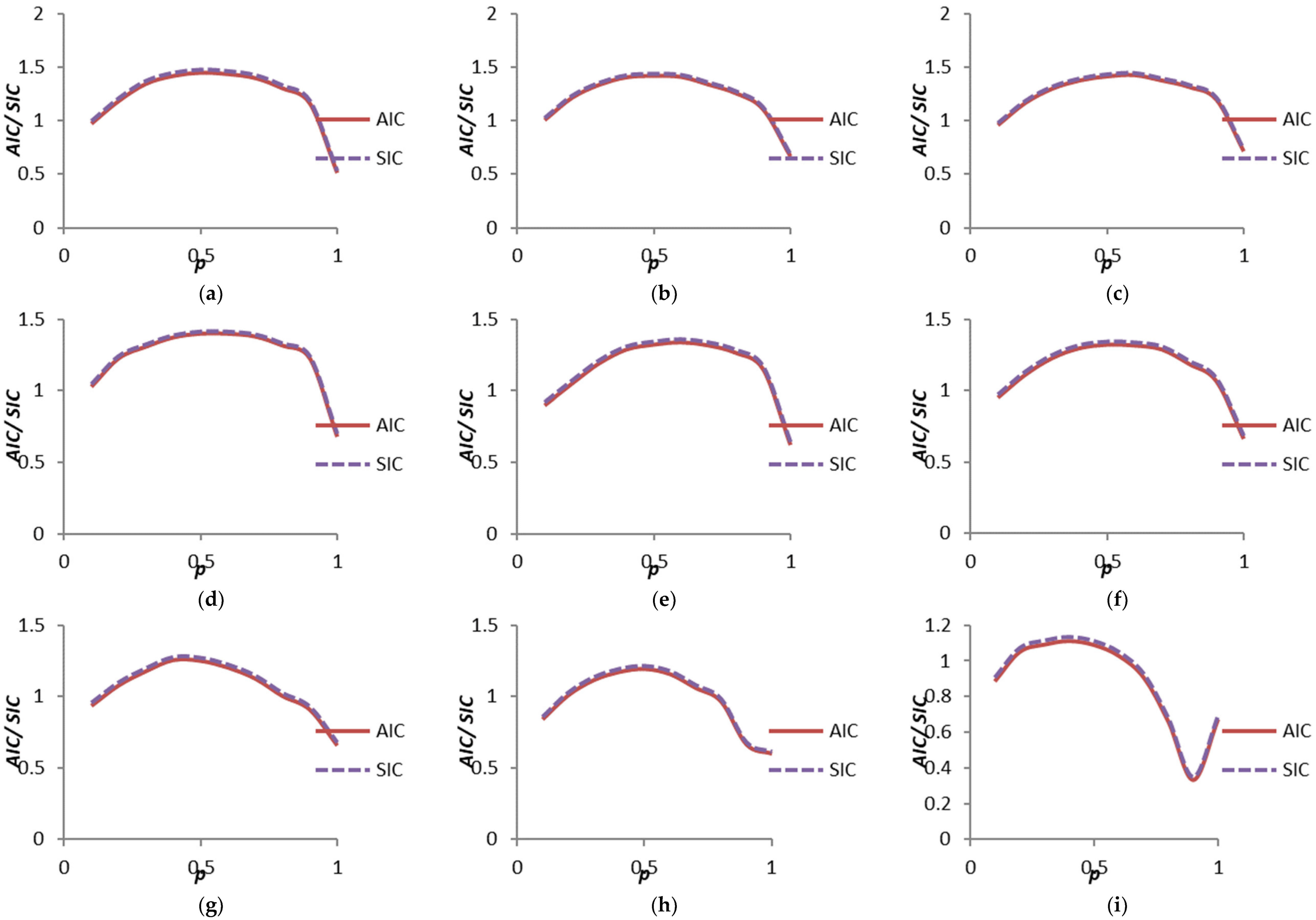

Figure 3. AIC and SIC using Huang [9] RRT. (**a**) $T = 0.1$; (**b**) $T = 0.2$; (**c**) $T = 0.3$; (**d**) $T = 0.4$; (**e**) $T = 0.5$; (**f**) $T = 0.6$; (**g**) $T = 0.7$; (**h**) $T = 0.8$; (**i**) $T = 0.9$.

Results Discussion for Huang [9] RRT

We have calculated the estimates of the b's for the different values of T and p, which are displayed in Table 1 and Figure 1. As an interpretation, we can state that taking the upper values of p, however low the values of T are, these show the same patterns as those of the standard logits. The logit estimates for Huang [9] RRT are altogether not the same as the normal values of the standard logits for $p = 0.5$–0.7 when T is set under 0.5. Nonetheless, these will generally move toward the standard logits when p is greater than 0.5. The values are indeed extremely high for T, between 0.2–0.3 and $p = 0.5$, and as p rises, the values of estimates decline and start approaching towards the standard logits. The same is the case for $T = 0.4$–0.5 and $p = 0.6$. Similar patterns of estimates might be seen when $T = 0.6$–0.7 and $p = 0.7$ are utilized. A very necessary point to be noted here is that the estimated values of the b's in the current technique exhibit extreme and divergent behavior when compared to conventional logits for a very elevated p and T, for which we can say that estimates are not reliable and consistent for the mentioned values of p and T. In all of the tables, the standard logit estimates are shown with $p = 1$. All in all, we might express that, for each increase in p and fall in T, the estimates are moving toward the ordinary logits.

The standard error of estimates for the proposed hidden logit for Huang [9] RRT can be seen in Table 2 and Figure 2. For a greater p and lower T, one can interpret that the standard errors lead to estimates that are fairly alike to regular logits. Elevated values of T and p bring about standard errors that go incredibly amiss from the anticipated values and show an unpredictable way of behaving. When $p = 1$, the standard errors of the estimates of ordinary logits are displayed. As a conclusion, we can state that an elevated p and lower values of T are associated with decreasing standard errors of estimations.

Table 3 and Figure 3 show the AIC and SIC to be more accurate for an elevated p and T. One can observe that these values are smallest for the highest p and T; hence, the best model fit for p and T can be considered as 0.9.

3.2. Estimates Using Warner [4] RRT

In the current section, we conduct an estimation of the proposed hidden logits using Warner [4] RRT, setting various values of p. We use the proposed hidden logit form produced in Equation (13) using the mentioned procedure in Section 2. Tables 4–6 and Figure 4 demonstrate the estimates of the b's, standard errors, AIC, and SIC values and their graphs, respectively. In Tables 4 and 5, the values against p when it is equal to 1 are standard logits, which can be obtained by taking $p = 1$ in Equation (13).

Table 4. Estimates of b's for different values of p using Warner [4] RRT.

b's/p	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9	1
b_0	0.1785	0.1554	0.2245	0.1680	−7.4527	−0.1875	0.1089	0.0252	−0.0415
b_1	0.8971	0.9316	0.9397	0.7655	28.6584	1.2712	0.9870	1.0410	1.0192
b_2	0.9361	0.9373	0.8488	0.6968	36.5850	1.3468	0.9870	1.0420	0.9907
b_3	0.9193	0.8968	0.8565	0.8933	17.1397	0.9436	0.7939	0.9340	0.9356

Table 5. The $s.e(b'$s$)$ for different values of p using Warner [4] RRT.

$s.e(b'$s$)/p$	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9	1
$s.e(b_0)$	0.1135	0.1668	0.2629	0.5182	15.7308	0.2994	0.1662	0.1138	0.0759
$s.e(b_1)$	0.1077	0.1629	0.2559	0.4455	62.8987	0.3702	0.1704	0.1215	0.0799
$s.e(b_2)$	0.1100	0.1622	0.2387	0.4223	80.2608	0.3798	0.1684	0.1200	0.0773
$s.e(b_3)$	0.1072	0.1555	0.2362	0.4779	37.5954	0.2874	0.1446	0.1101	0.0736

Table 6. AIC and SIC for different values of p using Warner [4] RRT.

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
AIC	1.0406	1.2401	1.3631	1.4344	1.4596	1.4222	1.3591	1.2463	0.9934	0.6506
SIC	1.0602	1.2597	1.3827	1.4540	1.4792	1.4418	1.3787	1.2659	1.0130	0.6702

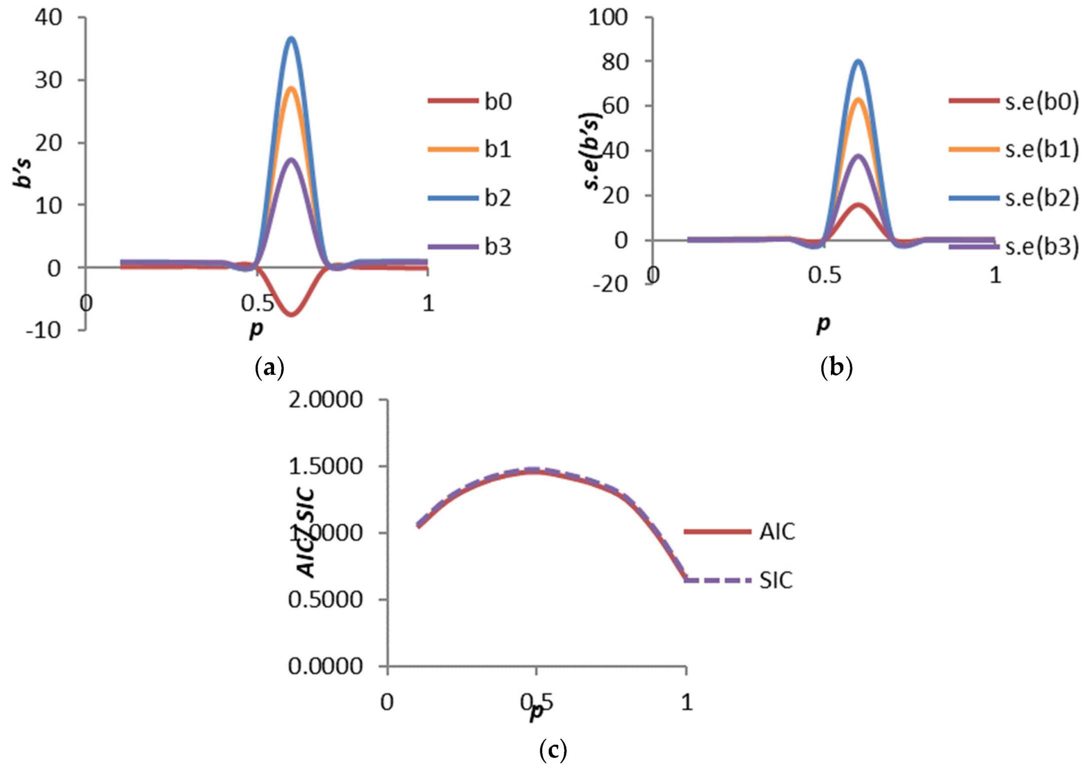

Figure 4. Estimates of ′*s*, the *s.e*(b's), AIC, and SIC using Warner [4] RRT. (**a**) b's; (**b**) *s.e*(b's); (**c**) AIC and SIC.

Results Discussion for Warner [4] RRT

Tables 4–6 are show the comparison of the performance for our proposed method's estimates with ordinary logit estimates. From Figure 4a, we perceive that the estimates of our proposed methodology come back to the estimates of b's, which are fairly close to ordinary logits. We observe a symmetric behavior in the values of b's around $p = 0.5$. For p below 0.5, estimates are found to be increasing and then begin to decrease for p above 0.5. When p is 0.5, the estimates do not exist, as defined by Warner [4]. Setting $p = 0.9$, standard logits and the proposed hidden logit become quite close to the estimates of the standard logits. Since Figure 4b exhibits the same behavior as Figure 4a, the same conclusion may be applied to SE. An extremely important aspect worth highlighting here is that the estimated values of b's and their SE exhibit an extreme and divergent behavior for $p = 0.6$. Therefore, we can say that estimates are not reliable and consistent for $p = 0.6$. Figure 4c illustrates that AIC and SIC begin to decline when higher values of p are taken, and, hence, they are the lowest for $p = 0.9$, which can be regarded as the best model fit for higher values of p.

3.3. Estimates Using Mangat [15] RRT

In the present section, we conduct an estimation of the hidden logits for Mangat [15] RRT, setting various values of p. We use the proposed hidden logit form produced in Equation (18) using the procedure mentio0ned in Section 2 for this purpose. Tables 7–9 show the b's, SE, AIC, and SIC values. Also, Figure 5 depicts their graphical display for different values of p using Mangat [15]. In Tables 7 and 8, under p equals to 1, one can see the standard logit values, which can be obtained by taking p as 1 in Equation (18).

Table 7. Estimates of b's for different values of p using Mangat [15] RRT.

b's/p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
b_0	0.4792	0.2964	0.4808	0.0060	−0.0150	0.2151	0.2160	0.1334	0.0934	0.1061
b_1	1.3541	1.1228	1.2025	1.2514	1.1177	1.2235	1.1079	1.0846	1.0959	1.1219
b_2	1.3213	1.0538	1.0382	1.3235	1.3531	1.3384	1.1475	1.0975	1.1971	1.2074
b_3	1.4329	1.1512	1.2126	1.3275	1.1771	1.1937	1.0225	1.0266	1.0959	1.1199

Table 8. The $s.e(b'$s) for different values of p using Mangat [15] RRT.

$s.e(b'$s)/p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$s.e(b_0)$	0.6178	0.3658	0.2980	0.2475	0.2006	0.1713	0.1358	0.1136	0.0955	0.07918
$s.e(b_1)$	0.7534	0.4075	0.3338	0.3074	0.2233	0.2036	0.1500	0.1253	0.1059	0.08944
$s.e(b_2)$	0.7130	0.3709	0.2903	0.2997	0.2455	0.2082	0.1473	0.1201	0.1072	0.08949
$s.e(b_3)$	0.7678	0.4018	0.3278	0.3083	0.2220	0.1932	0.1377	0.1166	0.1019	0.086

Table 9. AIC and SIC for different values of p using Mangat [15] RRT.

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
AIC	0.2729	0.3836	0.6667	0.8767	0.9410	0.9392	0.9323	0.8907	0.7197	0.5381
SIC	0.2533	0.4032	0.6863	0.8963	0.9606	0.9588	0.9519	0.9103	0.7393	0.5578

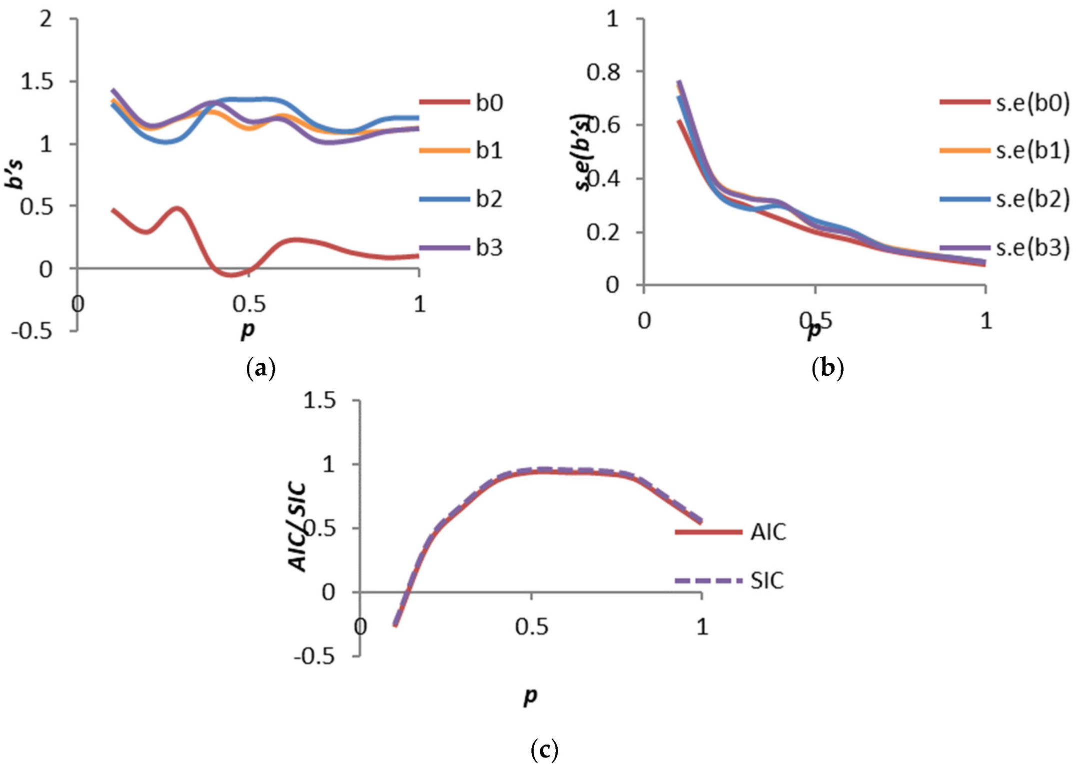

Figure 5. Estimates of b's, the $s.e(b'$s), AIC, and SIC using Mangat [15] RRT. (**a**) b's; (**b**) $s.e(b'$s); (**c**) AIC and SIC.

Results Discussion for Mangat [15] RRT

Tables 7–9 are made to show the comparison of performance for our proposed method's estimates for Mangat [15] RRT with ordinary logit estimates. From Figure 5a, we identify that the estimates of our proposed methodology show random behavior for diverse values of p. The estimates move towards standard logits setting $p = 0.9$. Figure 5b depicts that the values of SE decrease as p increases. Hence, no extreme pattern is observed. Figure 5c depicts that the AIC and SIC are the minimum for the lowest value of p, so taking $p = 0.1$can be considered as the best model fit.

4. Conclusions

This research depicts the results of the proposed hidden logit using different RRTs. A comparison for each RRT has already been conducted for ordinary and proposed hidden logits in Section 3. It can be observed that the calculated values converge to the ordinary logits when setting a higher value of p for all RRTs and lower values of T using Huang [9] RRT. We discovered that, as p increases, in all three RRTs, the hidden logit estimates approach the population parametric values. Additionally, we observe that the standard errors of the estimates decreases as p rises, and they are the least for $p = 1$. As a comparison between all three used RRTs, we can state that estimates for our proposed methodology using Huang [9] RRT are more accurate and are nearer to the population estimates. Thus, in the case of sensitive parameters, hidden logit estimation adopting Huang [9] RRT is more suitable to generate accurate estimates of population proportions. The AIC and SIC values also lead us to the same conclusion. The model with the lowest AIC and SIC is considered the best model. The values of AIC and SIC for Huang [9] are 0.3312 and 0.3509, which are found to be the least among all the above-discussed models. Therefore, Huang [9] can be considered the best model fit for higher values of p. This study makes a significant contribution to the application and evaluation of logistic models when investigating a sensitive topic and finding honest responses is difficult. We suggest that the proposed methodology of hidden logits can be employed with other RRTs.

Author Contributions: Conceptualization, A.H.; Methodology, A.H., I.A.A., A.M.A. and M.A.A.; Software, A.H. and A.M.A.; Validation, A.H.; Formal analysis, A.H. and M.A.A.; Investigation, A.H.; Resources, A.H.; Data curation, A.H.; Writing—original draft, A.H., I.A.A., A.M.A. and M.A.A.; Writing—review & editing, A.H., I.A.A., A.M.A. and M.A.A.; Visualization, A.H.; Supervision, I.A.A.; Project administration, A.M.A.; Funding acquisition, A.M.A. and M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. 3943].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the relevant data information is available within the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Geurts, M.D. Using a randomized response research design to eliminate non-response and response biases in business research. *J. Acad. Mark. Sci.* **1980**, *8*, 83–91. [CrossRef]
2. Mirhosseini, Z.; Pakdel, P.; Ebrahimi, M. Women, Sexual Harassment, and Coping Strategies: A Descriptive Analysis. *Iran. Rehabil. J.* **2023**, *21*, 117–126. [CrossRef]
3. Kassa, E.T. Factors influencing taxpayers to engage in tax evasion: Evidence from Woldia City administration micro, small, and large enterprise taxpayers. *J. Innov. Entrep.* **2021**, *10*, 8. [CrossRef]
4. Warner, S.L. Randomized response: A survey technique for eliminating evasive answer bias. *J. Am. Stat. Assoc.* **1965**, *60*, 63–69. [CrossRef]
5. Greenberg, B.G.; Abul-Ela AL, A.; Simmons, W.R.; Horvitz, D.G. The unrelated question randomized response model: Theoretical framework. *J. Am. Stat. Assoc.* **1969**, *64*, 520–539. [CrossRef]

6. Moors, J.J.A. Optimization of the unrelated question randomized response model. *J. Am. Stat. Assoc.* **1971**, *66*, 627–629. [CrossRef]
7. Mahmood, M.; Singh, S.; Horn, S. On the confidentiality guaranteed under randomized response sampling: A comparison with several new techniques. *Biom. J.* **1998**, *40*, 237–242. [CrossRef]
8. Christofides, T.C. A generalized randomized response technique. *Metrika* **2003**, *57*, 195–200. [CrossRef]
9. Huang, K.C. A survey technique for estimating the proportion and sensitivity in a dichotomous finite population. *Stat. Neerl.* **2004**, *58*, 75–82. [CrossRef]
10. Kim, J.M.; Warde, W.D. Some new results on the multinomial randomized response model. *Commun. Stat. Theory Methods* **2005**, *34*, 847–856. [CrossRef]
11. Kim, J.M.; Warde, W.D. A Mixed Randomized Response Model. *J. Stat. Plan. Inference* **2005**, *133*, 211–221. [CrossRef]
12. Kim, J.M.; Elam, M.E. A stratified unrelated question randomized response model. *Stat. Pap.* **2007**, *48*, 215–233. [CrossRef]
13. Kim, J.M.; Chae, S.S. Three-stage stratified randomize response model. *Korean J. Appl. Stat.* **2010**, *23*, 533–543. [CrossRef]
14. Mangat, N.S.; Singh, R. An alternative randomized response procedure. *Biometrika* **1990**, *77*, 439–442. [CrossRef]
15. Mangat, N.S. An improved randomized response strategy. *J. R. Stat. Soc. B* **1994**, *56*, 93–95. [CrossRef]
16. Narjis, G.; Shabbir, J. An improved two-stage randomized response model for estimating the proportion of sensitive attribute. *Sociol. Methods Res.* **2023**, *52*, 335–355. [CrossRef]
17. Hsieh, S.H.; Lee, S.M.; Li, C.S. A two-stage multilevel randomized response technique with proportional odds models and missing covariates. *Sociol. Methods Res.* **2022**, *51*, 439–461. [CrossRef]
18. Singh, G.; Singh, C. Proficient randomized response model based on blank card strategy to estimate the sensitive parameter under negative binomial distribution. *Ain Shams Eng. J.* **2022**, *13*, 101611. [CrossRef]
19. Singh, C.; Singh, G.N.; Kim, J.M. A randomized response model for sensitive attribute with privacy measure using Poisson distribution. *Ain Shams Eng. J.* **2021**, *12*, 4051–4061. [CrossRef]
20. Halim, A.; Arshad, I.A.; Haroon, S.; Shair, W. Effect of Misclassification on Test of Independence Using Different Randomized Response Techniques. *J. Policy Res.* **2022**, *8*, 427–438. [CrossRef]
21. Jaiswal, A.K.; Usman, M.; Singh, G.N. New calibration estimation procedure in the presence of unit non response. *Ain Shams Eng. J.* **2023**, *14*, 101910. [CrossRef]
22. Falodun, B.O.; Ige, E.O. Linear and quadratic multiple regressions analysis on magneto-thermal and chemical reactions on the Casson-Williamson nanofluids boundary layer flow under Soret-Dufour mechanism. *Arab. J. Basic Appl. Sci.* **2022**, *29*, 269–286. [CrossRef]
23. Dorugade, A.V. New ridge parameters for ridge regression. *J. Assoc. Arab. Univ. Basic Appl. Sci.* **2014**, *15*, 94–99. [CrossRef]
24. Ateeq, K.; Altaf, S.S.; Aslam, M. Modeling and analysis of recovery time for the COVID-19 patients: A Bayesian approach. *Arab. J. Basic Appl. Sci.* **2023**, *30*, 1–12. [CrossRef]
25. Clark, D.A.; Beck, A.T. Personality factors in dysphoria: A psychometric refinement of Beck's Sociotropy-Autonomy Scale. *J. Psychopathol. Behav. Assess.* **1986**, *13*, 369–388. [CrossRef]
26. Waldman, I.D.; Robinson, B.F.; Rowe, D.C. A logistic regression based extension of the TDT for continuous and categorical traits. *Ann. Hum. Genet.* **1999**, *63*, 329–340. [CrossRef]
27. Corstange, D. Sensitive questions, truthful responses? Randomized response and hidden logit as a procedure to estimate it. In Proceedings of the Annual Meeting of the American Political Science Association, Washington, DC, USA, 2–5 September 1986. Available online: http://www.umich.edu/~dancorst (accessed on 17 December 2021).
28. Corstange, D. Sensitive Questions, Truthful Answers? Modeling the List Experiment with LISTIT. *Political Anal.* **2009**, *17*, 45–63. [CrossRef]
29. Hussain, Z.; Shabbir, J. Logit estimation using Warner's randomized response model. *J. Mod. Appl. Stat. Methods* **2008**, *7*, 140–151. [CrossRef]
30. Hussain, Z.; Anjum, S.; Shabbir, J. Improved logit estimation through Mangat randomized response model. *Int. J. Bus. Soc. Sci.* **2011**, *2*, 179–188.
31. Halim, A.; Arshad, I.A.; Haroon, S.; Shair, W. A Comparative Study of Modified Hidden Logits Using Randomized Response Techniques. *J. Policy Res.* **2022**, *8*, 447–461. [CrossRef]
32. Cruff, M.J.L.F.; Böckenholt, U.; Van der Heijden, P.G.M.; Frank, L.E. Chapter 18—A Review of Regression Procedures for Randomized Response Data, Including Univariate and Multivariate Logistic Regression, the Proportional Odds Model and Item Response Model, and Self-Protective Responses. In *Handbook of Statistics*; Chaudhuri, A., Christofides, T.C., Rao, C.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; pp. 287–315. [CrossRef]
33. Chang, P.C.; Pho, K.H.; Lee, S.M.; Li, C.S. Estimation of parameters of logistic regression for two stage randomized response technique. *Comput. Stat.* **2021**, *36*, 2111–2133. [CrossRef]
34. Hsieh, S.H.; Perri, P.F. A Logistic Regression Extension for the Randomized Response Simple and Crossed Models: Theoretical Results and Empirical Evidence. *Sociol. Methods Res.* **2020**, *51*, 1244–1281. [CrossRef]