SAPRA, PUJITA. Ph.D. Accounting for Lack of Trust in Randomized Response Models. (2023)
Directed by Dr. Sat Gupta. 177 pp.

This study addresses a key assumption made while using traditional Randomized Response models in survey sampling when the question being asked pertains to a sensitive topic. It is traditionally assumed that under a randomized response framework, survey participants have no further reason to lie due to privacy concerns. We demonstrate that if this assumption is not true and even if a small proportion of respondents do not trust the RRT model being used in a survey, we get considerably biased estimates. We also propose alternative binary and quantitative models that account for respondents' lack of trust in traditional RRT models. These proposed models are mixtures of traditional RRT models and in one particular case mixture of an RRT model with an encryption technique, commonly used in the computer science domain. We also incorporate optionality into these models which helps improve the model efficiency. We evaluate the overall model performance using a combined measure of privacy and efficiency. Both theoretical and empirical results confirm that accounting for lack of trust helps us obtain more reliable results when survey respondents may not trust the RRT model used. Simulation studies have also been conducted to verify theoretical results. For sensitive mean estimation, we also propose estimators that utilize the auxiliary information and are more efficient compared to the ordinary mean estimator that does not utilize the auxiliary information.

ACCOUNTING FOR LACK OF TRUST IN RANDOMIZED RESPONSE MODELS

by

Pujita Sapra

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro

2023

Approved by

_____

Sat Gupta

Committee Chair

I dedicate this dissertation to my brother, who believed that I could be a great actor, a lawyer, an IAS officer and now graciously deals with my choice to be a statistician.

APPROVAL PAGE

This dissertation written by Pujita Sapra has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____

Sat Gupta

Committee Members _____

Sadia Khalil

_____

Somya Mohanty

_____

Scott Richter

_____

Haimeng Zhang

_____

Date of Acceptance by Committee

_____

Date of Final Oral Examination

ACKNOWLEDGMENTS

This dissertation would not have taken the shape it did without the contributions made by several people in various forms and ways.

First and foremost, I would like to extend my gratitude towards my advisor and mentor, Dr Sat Gupta, for presenting his pure enthusiasm for this topic and encouraging me to work with him on it. He was extremely patient and generous with his guidance, suggestions and most importantly, his time. Through various stages of this work, he showed up as the mentor I needed, eventually making me confident in my ability to work on research independently. His thorough and prompt feedback for my work has helped me stay focused and productive even during the most challenging phases of my research.

Next, I would like to express my gratitude to all the committee members, Dr Sadia Khalil, Dr Somya Mohanty, Dr Scott Richter and Dr Haimeng Zhang for their feedback and direction to help me improve my work. In particular, I would like to thank Dr Somya Mohanty for helping me identify and work on a very unique problem that brought about a new aspect to my work. Furthermore, I would like to thank Dr Sadia Khalil, who has also been my collaborator, for her endless encouragement and wise counsel on various projects that I have presented in this dissertation. Furthermore, I would like to thank Dr Qi Zhang, a UNCG alumna, for her time and her insightful comments and suggestions as they helped me approach my research problems in a more efficient manner.

I would also like to express my appreciation for all my professors at UNCG, who have played a critical role in my professional grooming as a statistician. Moreover, I would like to acknowledge my research collaborators, Maxwell Lovig, Sumaita Rahman, Joia Zhang and Nathaniel Mersy, who have been a part of the REU program hosted

by the Department of Mathematics and Statistics, for being an incredible team for my initial years in research. I would also like to thank my friends Kanika and Vaibhav who helped me navigate fundamental roadblocks during my research by offering me their thoughts based on years of industrial expertise.

My appreciation goes out to my fellow graduate students in my program who helped form a healthy community for brainstorming and helping each other resolve issues pertaining to our projects, LaTex syntax, research and struggles with teaching all the while standing in the hallways of the Brown Building. I would also like to thank my friends, both inside and outside the program, who were a great support, be it in studying together for exams, working on research, being the unofficial therapists and walking buddies for a much-needed *brain reset*. Finally, I would like to thank my family for their encouragement and support throughout my studies.

Table of Contents

List of Tables . . . . . **xi**

List of Figures . . . . . **xiii**

I. Introduction **1**

I.1. Respondent Privacy Protection in Sensitive Question Surveys 4

I.2. Social Desirability Bias 4

I.3. Approaches to Circumvent Social Desirability Bias 5

I.3.1. Bogus Pipeline Technique (BPL) 5

I.3.2. Unmatched Count Technique (UCT) 6

I.3.3. Social Desirability Bias (SDB) Scale 7

I.3.4. Randomized Response Technique (RRT) 8

I.3.5. Data Encryption Techniques 9

I.4. Outline of the Dissertation 11

II. Literature Review **14**

II.1. Estimation of Sensitive Trait Prevalence using Binary RRT Models . 16

II.1.1. Warner's Indirect Question Model [1965] 17

II.1.2. Greenberg's Unrelated Question Model [1969] 19

II.2. Homomorphic Encryption Techniques 21

II.2.1. Various Types of Homomorphic Encryption 23
II.2.2. Paillier Encryption Protocol and Algorithm 24
II.3. Sensitive Mean Estimation using Quantitative RRT Models 25
II.3.1. Warner's (1971) and Pollock & Beck's (1976) Quantitative RRT Model 26
II.3.2. Greenberg's Quantitative Model (1971) 27
II.3.3. Eichhorn and Hayre (1983) 29
II.3.4. Diana and Perri Linear Combination Model (2011) 30
II.4. Optional RRT Models 32
II.4.1. Gutpa et al.(2002) Optional Multiplicative RRT Model 33
II.4.2. Gupta et al. Optional Additive RRT Model (2010) 34
II.5. Use of Auxiliary Information in Quantitative RRT Surveys 36
II.5.1. Mean Estimation using Auxiliary Variables 37
II.5.2. Mean Estimation using Auxiliary Variables under RRT Models 40

III. Mixture Binary RRT Models with a Unified Measure of Privacy and Efficiency 53
III.1. Accounting for Lack of Trust in RRT Models 54
III.2. Proposed Mixture Binary RRT Model 58
III.2.1. Mixture Binary RRT Model with Unaccounted Untruthfulness 59
III.2.2. Efficiency under Mixture Binary RRT Model with Unaccounted Untruthfulness 60
III.2.3. Mixture Model Accounting for Untruthfulness 61
III.2.4. Efficiency under Mixture Binary RRT Model with Accounted Untruthfulness 63

III.2.5. Introduction to Privacy under Binary RRT Models 64
III.2.6. Privacy under Mixture Binary RRT Model with Accounted Untruthfulness 65
III.2.7. Unified Measure of Privacy and Efficiency 67
III.3. Simulation Study 68
III.3.1. Impact of untruthfulness on the unified measure 72
III.4. Concluding Chapter Remarks 73

IV. Optional Mixture Binary RRT Model with a Unified Measure of Privacy and Efficiency 75
IV.1. Proposed Optional Mixture Binary RRT Model 76
IV.1.1. Accounting for Lack of Trust in Optional Mixture RRT Model 81
IV.1.2. Efficiency of Optional Mixture Binary RRT Model 82
IV.1.3. Privacy of the Optional Mixture RRT Model 84
IV.2. Simulation Study on the Optional Mixture Binary RRT Model 85
IV.3. Concluding Chapter Remarks 89

V. Hybrid (Encryption + RRT) Model 90
V.1. Paillier Encryption Protocol 91
V.2. Additive Homomorphism Property of Paillier Encryption Scheme . . . 93
V.3. Example of Paillier Encryption Application 95
V.4. Hybrid (Paillier + Warner RRT) Model 97
V.4.1. Proposed Hybrid (Paillier + Warner RRT) Model 99
V.5. Simulation Study on Hybrid Model 103
V.5.1. Paillier Encryption Simulation-Example for Binary Responses 103
V.6. Concluding Chapter Remarks 107

VI. Mitigating Lack of Trust in Quantitative RRT Models 109

VI.1. Background for Optional Enhanced Trust Model 111

VI.1.1. Efficiency of RRT Models 111

VI.1.2. Privacy Level in Quantitative RRT Models 111

VI.1.3. Combined Measure for Efficiency and Privacy 111

VI.1.4. Warner Additive Model (1971) 112

VI.1.5. A Linear Combination Model (2011) 113

VI.2. Estimation of the Mean and Sensitivity Level using Optional Enhanced Trust (OET) Model 115

VI.3. Ratio Estimator of the Mean for the OET Model 119

VI.4. Regression Estimator of the Mean for the OET Model 123

VI.5. Simulation Study 128

VI.6. Concluding Chapter Remarks 138

VII. Generalized Mean Estimator under the Optional Enhanced Trust Model 140

VII.1. Mean Estimation using the Optional Enhanced Trust (OET) Model . 141

VII.2. Ratio Estimator of the Mean for the OET Model 144

VII.3. Regression Estimator of the Mean for the OET Model 146

VII.4. A Generalized Estimator of the Mean for the OET Model 149

VII.5. Simulation Study 152

VII.6. Concluding Chapter Remarks 162

VIII. Concluding Remarks and Future Directions 164

VIII.1. General Discussion of Work and Remarks 164

VIII.2. Future Directions 166

References **167**

A. List of Publications **177**

List of Tables

I.1. Marlowe-Crowne Short Form with 13-Items proposed by Reynolds (1982)[52] 8

III.1. Estimates averaged over 10000 simulations with Greenberg model with untruthfulness, $n = 500$, $\pi_x = 0.3$, $\pi_y = 0.1$ 57

III.2. Theoretical (bold) and empirical values based on 10000 iterations $n = 500$, $\pi_X = 0.4$, $\pi_Y = 0.1$, $\pi_{y0} = 0.1$, $p_0 = 0.7$ 70

IV.1. Simulation Results: Estimator performance when one wrongly assumes $A = 1$ $(N = 10000, n = 500, \pi_x = 0.4)$ 80

IV.2. Simulation Results Mixture ORRT: N=1000, n=500, $\pi_x = 0.4$, $\pi_y = 0.1$, $\pi_{y0w} = 0.1$, $p_{0w} = 0.7$, $\pi_{y0a} = 0.15$, $p_{0a} = 0.75$ 87

V.1. Paillier Encryption Example for Binary Responses 104

V.2. Simulation results for Hybrid model: $Iterations = 10000$, $n = 500$, $\pi = 0.3$ for various levels of mixture (α) and the Warner's model parameter (p). 107

VI.1. Simulation results (for estimating sensitive mean μ_Y and sensitivity level W (Theoretical (**bold**) and Empirical Measures): $Iterations = 10000$, $N = 5000, n = 500$, $\mu_Y = 10$, $\rho_{YX} = 0.9$ for various levels of trust (A) and the sensitivity level (W). 132

VI.2. Simulation results (for estimating sensitive mean μ_Y and sensitivity level W: $Iterations = 10000$, $N = 5000, n = 500$, $\mu_Y = 9$, $\rho_{YX} = 0.6$ for various levels of trust (A) and the sensitivity level (W). 135

VII.1. Simulation results for estimating sensitive mean μ_Y only (Theoretical (**bold**) and Empirical Measures): $Iterations = 10000$, $N = 5000, n = 500$, $\mu_Y = 10$, $\rho_{YX} = 0.9$, $k = 1$, $g = 1$, $a = 1$ and $b = 0$ for various levels of trust (A) and the sensitivity level (W). 155

VII.2. Simulation results for estimating sensitive mean μ_Y only (Theoretical (**bold**) and Empirical Measures): $Iterations = 10000$, $N = 5000, n = 500$, $\mu_Y = 10$, $\rho_{YX} = 0.6$, $k = 1$, $g = 1$, $a = 1$ and $b = 0$ for various levels of trust (A) and the sensitivity level (W). 159

List of Figures

I.1. Asymmetric encryption and decryption process [27] 10

III.1. Lack of Trust and Untruthfulness in Greenberg Model 56

III.2. Mixture RRT Model 59

III.3. Unified Measure (M) for different choices of p and q $\pi_X = 0.3$,$\pi_Y = 0.1$, $n = 500$, $A = 0.8$, $a = 1$, $b = 1$ (a) Untruthfulness is not accounted for. (b) Untruthfulness is accounted for. 73

IV.1. Optional Mixture Binary RRT Model 77

V.1. General Asymmetric Encryption Process 92

V.2. Simplified Paillier Encryption Process 94

V.3. Election Results 96

V.4. Hybrid Encryption-RRT Model 98

V.5. Hybrid Model using Paillier Encryption & Warner's Indirect Question RRT Model 99

VI.1. Optional Enhanced Trust Model 116

Chapter I: Introduction

The primary focus of many statistical studies is on the estimation of various population parameters. However, conducting a thorough evaluation of the entire population of interest is often not practically possible under time and resource constraints. Therefore, researchers often resort to conducting sample studies to help make inferences about the population being studied. Unfortunately, samples drawn for the purpose of such studies may not always be an appropriate representation of the true population of interest due to sampling and non-sampling errors which are potentially present in surveys.

Sampling errors get introduced in a survey due to imperfect representation of the population being studied. On the other hand, non-sampling errors are caused by aspects that are not associated with the process and design of a sample survey such as subject non-responses and intentional or unintentional misreporting[1]. Although one can often reduce sampling errors by increasing the sample size, it may not always be possible. Therefore, researchers often need to settle with a reasonable size of the sample that aligns with constraints associated with effort, time and cost, and efficiency in estimation for the population parameter of interest(1)[31]. Another cause of concern for researchers while conducting a survey study is the non-sampling errors. There are numerous sources of non-sampling errors including but not limited to non-response and untruthful responses by the respondents(2) [51].

If the sample drawn for a survey study is a good representation of the population of interest, one can assume the inferences made based on such a sample would be reasonable. However, if the sample drawn is not representative of the population under study, the estimates computed and the inferences made would both be unreliable.

Several sampling techniques have been studied and implemented over the years such as simple random sampling (SRS), stratified sampling, cluster sampling, two-stage sampling etc. Depending on the study objective and the population of interest, one can make use of one of these different sampling techniques to draw a representative sample. One of the most basic sampling techniques is the SRS. Under this technique, each individual in the population has an equal chance of being selected as a sample unit. The sample units are randomly selected from the population after being identified in the population. This method is used when the list of all subjects from the population being studied, i.e. the sampling frame, is accessible to the researchers(3)[12]. If we do not choose an appropriate sampling technique, we might end up with a biased sample.

In addition to the sampling techniques, we must be careful about the method of conducting the survey based on the study objective and the potential nature of the population being scrutinized. These decisions must be made while optimizing for cost constraints for a study. We can use different survey mediums such as electronic surveys, phone surveys, mail surveys and in-person interviews. Although electronic, mail and phone surveys can be a cheaper alternative for collecting survey responses, they tend to have a very high non-response rate. This can often lead to a non-response bias if the respondents who choose not to participate are systematically different from those who do. For example, suppose the survey question is "Have you consumed any illegal drugs in the past week?". In such surveys, it is possible that the majority of those who do not respond to the survey are also people who have actually consumed

illegal drugs and were probably concerned about having to divulge this information. Hence, such a phenomenon can also lead to participation bias as the majority of people who choose to participate in the survey have a specific kind of response only. For instance, when one is asked to take a feedback survey at the end of a customer service call, people who have extremely good or extremely bad opinions are usually the ones that end up taking the survey. This can leave us with survey responses that are not reflective of how the responses are distributed in the population. Therefore, researchers must be wary of low survey participation rates and untruthful responses as they can be sources of considerable bias, especially when the survey is on a topic that respondents might consider sensitive.

Face-to-face interviews are more expensive for the researcher but have a higher response rate. However, when the question being asked is on a sensitive subject, it can lead to social desirability bias (SDB). SDB refers to the tendency of survey respondents to give socially acceptable and favorable responses, even if they are untruthful. For example, if someone is asked "Have you tested positive for any sexually transmitted disease in the last year?" or "Have you ever incorrectly reported your income to the IRS?", they would probably report a response of "No" even if that is not the truth because otherwise they might be worried that the surveyor would not view them as *socially acceptable* or worse, report them to authorities, if applicable, which could result in a legal repercussion against them. Therefore, for surveyors working on sensitive data collection, non-response and untruthful responses are two major concerns.

This dissertation will address practical issues encountered in surveys with questions on sensitive topics. We talk about methods that address the root causes of untruthful responses in sensitive question surveys such as social desirability bias and lack of

trust in privacy protection under the survey method being used. For all the studies discussed in this dissertation, we have considered samples drawn by the SRS technique.

I.1 Respondent Privacy Protection in Sensitive Question Surveys

When respondents participate in sensitive question surveys, the concern for their privacy can influence their behavior in the survey. The concern for privacy protection is so high that either the sampled respondents might decline to participate, or worse, they might lie when they do participate. This can worsen the extent of social desirability bias (SDB) when the survey is being conducted as an in-person interview. Moreover, the surveyors have an ethical and professional obligation to protect respondent privacy(4) [49]. Traditional direct survey questioning is not appropriate for sensitive question surveys. Therefore, researchers must consider methods specifically designed to reliably collect sensitive data in a survey while protecting respondent privacy.

I.2 Social Desirability Bias

A common objective for researchers in social and behavioral science is often to estimate the prevalence of sensitive behavior. However, human beings have a tendency to want to appear more altruistic and society-oriented than they actually are(5)[7]. When survey participants modify their true response before reporting for impression management (to look better to others) or self-deception (to feel good about themselves) we end up with social desirability bias (SDB) in our survey responses(6)[39]. This happens when respondents try to adhere to the social status quo to come across as socially

acceptable individuals or when respondents are afraid of unfavorable consequences against them. Surveys that ask questions on topics that are often considered taboo in a society such as sexual activities, illegal behavior such as social fraud or unsocial attitude such as racism often renders responses that are distorted due to SDB(7)[35]. Such responses lead to unreliable estimates.

SDB is one of the many sources of bias that surveyors encounter. Other typical sources of bias can be evasive responses, non-responses, selection or coverage issues, voluntary responses, etc. These biases are issues of concern as they cause sample estimates to systematically either over-estimate or under-estimate the population parameters of interest.

I.3 Approaches to Circumvent Social Desirability Bias

Statisticians, sociologists and psychologists have come up with various techniques to encourage survey respondents to participate in surveys and to provide truthful responses by guaranteeing them confidentiality. A few such methods are the Bogus Pipe Line (BPL), Unmatched Count Technique (UCT), SDB scale, Randomized Response Techniques and Encryption.

I.3.1 Bogus Pipeline Technique (BPL)

The Bogus Pipeline method was proposed by Jones and Sigall 1971 [28](8) based on the idea that if the respondents can be convinced that a physiological monitoring device is able to measure both the amplitude and direction of emotional response, their

subsequent attempts to predict what the machine says about their attitudes should be uncontaminated by many biases that obscure paper-and-pencil measures. This method involves tricking the respondents into believing that the untruthful responses can be detected by a mere decoy of a physiological monitoring device such as a polygraph machine. If the respondents are convinced that the device monitoring them is a working lie detector, the responses through this method would not get affected by many of the potential biases because respondents would not want to possibly *lose face* if they were to be caught while lying. Under this method, no attempt is made to protect the respondents' privacy.

I.3.2 Unmatched Count Technique (UCT)

The Unmatched Count technique (UCT) or Item Count Technique (ICT) was proposed by Raghavarao and Federer 1979(9). Under this method, survey participants are randomly assigned into two groups. One of the groups receives a set of non-sensitive questions or "items" while the other group receives this list with an additional sensitive item/question. For example, if the first group receives a list with four non-sensitive items, the second group will receive the list with these four non-sensitive items and a fifth item that is sensitive. Participants in both groups are then asked to report the number of items that are applicable to them. Respondents are only required to report the count of items applicable to them without needing to disclose what items are applicable. The prevalence of sensitive behavior is estimated by calculating the difference in mean count reported in the two groups[26](10). However, it should be noted that despite the random assignment, some of the differences between the two groups may be a function of differences between the two groups not entirely related to the number of subjects in the second group who consider the additional sensitive item

[9](11).

I.3.3 Social Desirability Bias (SDB) Scale

Social Desirability Bias (SDB) scales are used to measure an individual's tendency to present themselves as more socially acceptable than they actually might be. This technique assumes that the reason for a respondent to lie is to look more socially acceptable by downplaying any negative behavior.

Crowne-Marlowe Social Desirability Bias (MCSDB) was proposed by Crowne and Marlowe (1960)[8](12) which uses a 33-item questionnaire based on personal attitude for the subjects. Each item can be answered with a "True" or a "False" and has a socially acceptable response depending on how the statement has been phrased. The subjects are administered this questionnaire. For each item, if the respondent selects the socially acceptable response, they earn a score of one, else they get zero as the score. The higher an individual's score, the greater their tendency to come across as socially desirable. Although the MCSDB method helps in evaluating the tendency of an individual to be socially desirable, the 33-item questionnaire in addition to the sensitive question survey can cause response fatigue in the survey participants. Therefore, Reynolds (1982) [52](13) proposed that all individuals can be scored based on their likelihood of giving a socially desirable response based on 11, 12 and 13-item condensed versions of the MCSDB scale questionnaire which can be used to approximate results obtained from the MCSDB scale. They showed that with a little over one-third of the items on the original MCSDB scale, the 13-item form recommended by them would provide a brief easy-to-administer social desirability measure[52]. The questions on this 13-item form, along with the socially desirable responses (in parentheses), are listed in Table I.1.

Table I.1. Marlowe-Crowne Short Form with 13-Items proposed by Reynolds (1982)[52]

S.No	Item	
1	It is sometimes hard for me to go on with my work if I am not encouraged.	(False)
2	I sometimes feel resentful when I don't get my way.	(False)
3	On a few occasions, I have given up doing something because I thought too little of my ability.	(False)
4	There have been times when I felt like rebelling against people in authority even though I knew they were right.	(False)
5	No matter who I'm talking to, I'm always a good listener.	(True)
6	There have been occasions when I took advantage of someone.	(False)
7	I'm always willing to admit it when I make a mistake.	(True)
8	I sometimes try to get even rather than forgive and forget.	(False)
9	I am always courteous, even to people who are disagreeable.	(True)
10	I have never been irked when people expressed ideas very different from my own.	(True)
11	There have been times when I was quite jealous of the good fortune of others.	(False)
12	I am sometimes irritated by people who ask favors of me.	(False)
13	I have never deliberately said something that hurt someone's feelings.	(True)

Once obtained, the SDB scores can then be utilized as a covariate along with the reported responses in the direct sensitive question survey to reflect a more truthful picture in research studies that rely on self-reported measures of personality, individual characteristics and behavior.

I.3.4 Randomized Response Technique (RRT)

The Randomized Response Technique (RRT) method allows a survey administrator to collect sensitive data from individuals by asking them to add some noise to their original response. The noise to be added depends on the question and potential response types and is determined prior to conducting the data collection stage. This helps to

protect respondent privacy and improves honest response behavior by eliminating any potential social desirability bias (SDB) and embarrassment. RRT was first introduced by Warner (1965)[62] to reduce or eliminate under-reporting of sensitive behaviors (Scheers (1992)[55]. Several more field studies, such as the one by Chow et al. (1979)[5], Chaloupka (1985)[3] and Chhabra et al. (2016)[4], have also confirmed the efficacy of RRT models in acquiring information from the respondents. A study was conducted by Kwan et al.(2010)[36] on software piracy and their results confirmed that the respondents that responded to the question using RRT were more willing to provide true responses about the behaviors of software piracy.

We will discuss various aspects of RRT models in Chapter II.

I.3.5 Data Encryption Techniques

In the modern day and age of cloud computing and connected servers, sensitive data are being collected and stored online on a very large scale. Data Encryption is a traditional data security method that uses a cipher to secure highly confidential information especially the kind that is shared or collected online.

Traditionally, there are two major types of encryption- Asymmetric and Symmetric Encryption. Here, Figure I.1 shows an example of asymmetric encryption protocol, which is usually conducted in three steps. The first and most important stage is the key generation in which a public and private key pair is generated. The confidential data can be encrypted using the public key in the next stage. The encrypted information can then be sent over to the intended organization or person where it can be decrypted using the private key.

- Homomorphic Encryption (HE): Homomorphic encryption is a special kind

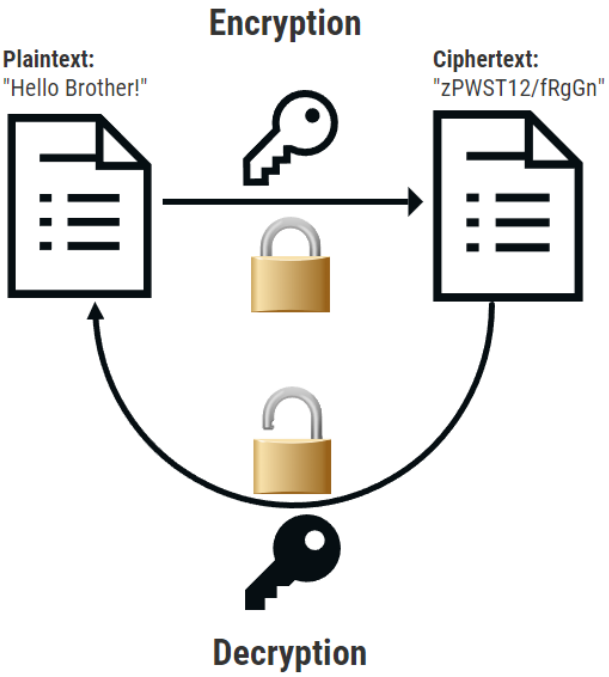

Figure I.1. Asymmetric encryption and decryption process [27]

of encryption mechanism that can resolve the security and privacy issues [56] and also allows the administrators to conduct certain mathematical operations to be carried out on the ciphertext, instead of on the actual data itself [6]. Therefore, it facilitates some computations (multiplication, addition, etc.) to be performed on the encrypted data without requiring decryption. The result of such computations, when decrypted, is the same as what one would get if the operation were performed on the original data. This is a critical property of homomorphic encryption.

- Partial Homomorphic Encryption (PHE): An encryption protocol is considered partially homomorphic if it only preserves the additive operation or the multiplicative operation over the ciphertext but not both. This means that, unlike

a (fully) homomorphic encryption protocol, one can perform either addition or multiplication over the encrypted data and decrypt it to sum/product of original values but both operations cannot be successfully implemented simultaneously on the ciphertext. Some examples of partially homomorphic cryptosystems are RSA (multiplicative homomorphism), ElGamal (multiplicative homomorphism) and Paillier (additive homomorphism) [6].

The HE and PHE methods can be utilized when the survey question has a binary response (Yes/No) or when the survey question has a non-negative integer response. This method has been further discussed in Chapter V.

The primary focus of this dissertation is on the theory and application of RRT models. In particular, we want to address the issue and the effect of the lack of trust in RRT models, and methods to mitigate those effects. We propose various mixtures of traditional RRT models that address and mitigate the respondents' lack of trust in the traditional methods used in surveys on sensitive topics.

I.4 Outline of the Dissertation

Chapter I provides a brief introduction to the issue of Social Desirability Bias (SDB) encountered in sensitive question surveys and methods that exist in literature to help address these issues such as the Bogus Pipeline (BPL) method, the Unmatched Count Technique (UCT), SDB scale, Data Encryption and Randomized Response Technique (RRT).

Chapter II provides a detailed background and literature review of some important RRT models, including both binary and quantitative models. Additionally, this chapter will present a detailed review of Optional RRT models, sensitive mean estimation

in the presence of auxiliary information and accounting for the lack of trust in RRT surveys. The concepts highlighted in this chapter serve as the foundation of the work carried out in the subsequent chapters.

Chapter III presents a mixture of Warner's Indirect Question model [62] and Greenberg's Unrelated Question model [16]. The efficiency and the privacy protection offered under this binary RRT model are evaluated and compared to the traditional models using a unified measure of privacy and efficiency. This model accounts for a lack of respondents' trust. These models are compared both theoretically and through a simulation study. The match between the theoretical and the empirical results is also noted.

Chapter IV presents an extension of the work discussed in Chapter III by including the aspect of optionality. The efficiency and the privacy protection offered under this optional binary mixture RRT model are evaluated and compared to the traditional models using a unified measure of privacy and efficiency. These models are compared both theoretically and through a simulation study. The match between the theoretical and the empirical results is also noted.

Chapter V presents a model which is a hybrid mixture of RRT and Paillier encryption techniques. Potential benefits in terms of efficiency and privacy by the inclusion of encryption with RRT have also been discussed. A simulation study is also conducted to show the efficiency performance of the hybrid mixture of Warner's Indirect Question model [62] and the Paillier encryption technique relative to a pure Warner's Indirect Question model [62].

Chapter VI presents mean estimators based on a quantitative model that accounts for respondents' lack of trust in the absence and the presence of non-sensitive auxiliary information under various scenarios using a split-sample technique which allows

simultaneous estimation of the sensitivity level of the survey question. A simulation study is also conducted to show and compare the performance of the various sensitive mean estimators.

Chapter VII presents only mean estimators for the sensitive study variable under the quantitative model introduced in Chapter VI both in the absence and the presence of non-sensitive auxiliary information without estimating the sensitivity level of the survey question. This decreases the sampling burden. We also introduce a generalized estimator for the mean of the sensitive study variable in the presence of auxiliary information. Unlike Chapter VI, all the sensitive mean estimators were obtained without splitting the sample.

Chapter VIII presents a general discussion of the research introduced in this dissertation. It also summarizes the most significant findings and some future directions for the work presented in this dissertation.

Chapter II: Literature Review

As discussed earlier, a major cause of non-response and untruthful responses in sensitive question surveys is social desirability bias (SDB). Over the years, several methods have been recommended to address SDB.

The primary focus of this dissertation will be on RRT models and related topics. A popular method used by psychologists is the Bogus Pipeline (BPL) which convinces the respondents that any truthful response can be detected by a lie detector machine when in fact there is no way for the BPL to detect such phenomena. However, this method is not mathematically credible nor does it offer any privacy to the survey participants. Although SDB scales and unmatched count technique (UCT) are relatively better in accounting for SDB, they still put the respondents under the pressure of revealing the truth to some extent. SDB scales provide a measure of an individual's tendency to give a more socially acceptable answer which can be helpful but it still puts respondents under the pressure to reveal their responses through direct questioning. Although UCT merely asks respondents to reveal the number of items that apply to them, it still puts respondents against direct questioning without offering anything more to protect their privacy. However, the Randomized Response Technique (RRT) is a method that shifts away from a direct-questioning survey set-up to offer a more secure data collection in terms of respondent privacy.

Randomized response technique (RRT) is a method that is based on adding noise

to a respondent's true response such that the survey administrator cannot know their true response, thereby helping to curb the tendency to lie. This traditional RRT method makes use of a randomizing mechanism that decides whether a respondent's final response would include the noise or not. Moreover, the aggregate responses can be unscrambled to estimate the sensitive population parameters but there is no way to unscramble the individual responses. Therefore, this method helps in providing privacy to the respondents with reasonable efficiency in the estimation process.

One can broadly classify RRT models as follows:

1. Based on the type of the survey question response:

    - Binary RRT Models- When the survey question can have "Yes" or "No" i.e. binary responses.
    - Quantitative RRT Models- When the survey question can have quantitative responses.

2. Based on the rules which are followed by the respondents in scrambling their true response before reporting it:

    - Full RRT- all respondents scramble their responses.
    - Partial RRT- some respondents scramble and some do not scramble their response but the choice to do so is made by the researcher using a random process.
    - Optional RRT- if respondents find the question sensitive they scramble their response, otherwise they report unscrambled response. The choice to scramble or not is made by the respondent.

[62] was the first to propose the randomized response method as a survey technique to reduce potential bias due to non-response and social desirability during survey studies of sensitive behavior prevalence and beliefs [2]. In the years since, several variants of RRT models have emerged motivated by the need to improve the estimator efficiency, respondent privacy, or both.

In this chapter, various types of RRT models and related techniques have been reviewed.

II.1 Estimation of Sensitive Trait Prevalence using Binary RRT Models

Binary Randomized Response Technique (RRT) models are used when the survey question is on a sensitive topic and requires a binary response. Binary responses can be of various forms such as "Yes/No", "Agree/Disagree", "Support/Do not Support", etc. The most common form of binary response questions usually requires a "Yes" or a "No" response. For example, when asked "Have you ever violated the academic integrity policy in college?", a respondent can either respond with a "Yes" or with a "No".

The foundation of the RRT involves creating an indirect questioning survey set-up such that at no point is the respondent expected to respond directly to the question based on a sensitive topic. To this effect, a randomizing device can be placed facing only the respondent and prompts the respondent to respond with a certain set of rules which scramble the individual responses. The interviewer would not be able to see what prompts are provided and hence is unable to distinguish or unscramble the reported responses thereby ensuring a much higher level of privacy compared to many

other methods considered for sensitive data collection such as the Bogus Pipeline Technique, the Unmatched Count Technique and the SDB scale method.

II.1.1 Warner's Indirect Question Model [1965]

The first technique to curb the extent of untruthful responses in sensitive surveys was introduced by S L Warner in 1965 [62] for questions with only binary responses (Yes/No). He proposed a set-up where all respondents were randomly prompted to respond to the sensitive question directly or indirectly while the survey investigator would be completely oblivious to which question was being responded to. Hence, the investigator would only know the final response (Yes/No) without knowing whether the respondent was prompted to respond to the direct or the indirect question, thereby protecting their privacy. Therefore, the key is to randomly divide the survey participants into two groups and ask the question of interest either rephrased directly or indirectly depending on which group they get assigned to.

Suppose we are interested in estimating the prevalence of academic integrity policy violations at a university. Then the survey respondents would be prompted to respond to one out of the following two statements:

1. I violated the academic integrity policy in one or more courses last semester (Group A).

2. I did not violate the academic integrity policy in any course last semester (Group B).

The randomization device that prompts all respondents could be a random spinner on a board that points at the letter A with a fixed probability p and at the letter

B with a probability $(1-p)$ thus deciding which question a particular respondent is supposed to respond to. The surveyor cannot see which question the respondent has been prompted to respond to. The respondents simply report a "Yes" or a "No" truthfully to the question they have been prompted to respond to.

Alternatively, we could have a deck of n cards with a proportion p of cards that state "I violated the academic integrity policy in one or more courses last semester" while the remaining cards have "I did not violate the academic integrity policy in any course last semester". The deck of cards is well shuffled and respondents pick a card, respond truthfully to the statement or the question on the card and replace the card back into the deck to be shuffled again before the next respondent is asked to go through the same process. Again, the surveyor cannot see which card has been picked up by the respondents and simply records the reported "Yes" or "No". It must be noted that when a respondent from Group A reports a "Yes" it indicates that they have violated the academic integrity policy. However, when a respondent from Group B reports a "Yes", it indicates that the subject has not violated the academic integrity policy. Thus a reported "Yes" response alone does not necessarily mean that the respondent has violated the academic integrity policy.

Let π be the true proportion of individuals in the population who have engaged in sensitive behavior or have a sensitive trait. Let p be the proportion of subjects in the sample that got assigned to Group A, i.e. subjects that were asked the sensitive question phrased directly. The value of p is fixed by the researcher before a simple random sample of n subjects is drawn from the population. Suppose that n_1 out of the n subjects reported a "Yes" as their response. However, not all of these n_1 respondents have indulged in sensitive behavior.

If P_y is the probability of a reported "Yes" response then,

$$P_y = p\pi + (1-p)(1-\pi). \tag{II.1}$$

Rearranging equation (II.1) Warner (1965)[62] proposed the unbiased estimator

$$\widehat{\pi} = \frac{\frac{n_1}{n} - (1-p)}{2p-1} = \frac{\widehat{P_y} - (1-p)}{2p-1}, p \neq \frac{1}{2}. \tag{II.2}$$

Based on the fact that

$$Var(\widehat{P_y}) = \frac{\pi(1-\pi)}{n}, \tag{II.3}$$

the variance of this estimator for a simple random sample with replacement is given by

$$Var(\widehat{\pi}) = \frac{\pi(1-\pi)}{n} + \frac{p(1-p)}{n(2p-1)^2}. \tag{II.4}$$

Here, the term $\frac{p(1-p)}{n(2p-1)^2}$ is the penalty added on for introducing noise through the RRT model. In order to minimize this penalty, and thus the variance, a large sample size n should be chosen. The proportion of subjects that get assigned to Group A (p) should be considerably different from 0.5 or p should be fixed at a value closer to 0 or 1.

II.1.2 Greenberg's Unrelated Question Model [1969]

A few years after Warner introduced RRT as a method for sensitive question surveys, Greenberg et al. presented a theoretical framework for the Unrelated Question RRT model[62]. They found that based only on the evasive answer bias and not on the response rate, the [62] technique usually requires a substantial amount of lying before

it becomes worthwhile. Based on the recommendation by Walt R. Simmons, [16] felt that by providing the respondent with the opportunity of replying to one of two questions in which one question is completely innocuous and unrelated to the stigmatizing attribute, the respondent might be more truthful.

In this unrelated-question model, the sample subjects are assigned to one of two groups using a randomization device similar to what one might use for the Warner's Indirect Question model [62]. The individual is then asked the sensitive question or the innocuous question depending upon which group they get assigned to. A pair of such sensitive and innocuous questions could be:

1. "Have you been diagnosed with an STD in the past 12 months?" (Group A)

2. "Were you born in the month of June?" (Group B)

Although it is not necessary, the unrelated question is chosen such that the distribution of the responses to this question can be determined/is known prior to the survey. For instance, the probability of a June birth is approximately $\frac{30}{365}$. Here we can refer to having a diagnosis of an STD as the sensitive trait and a June birth to be an innocuous trait.

Suppose π_x is the true proportion of individuals in the population who have been diagnosed with an STD in the past 12 months and π_y is the true proportion of individuals in the population that were born in the month of June.

Now

$$P_y = p\pi_x + (1-p)\pi_y. \tag{II.5}$$

This gives us the unbiased estimator

$$\widehat{\pi_x} = \frac{\widehat{P_y} - (1-p)\pi_y}{p}. \tag{II.6}$$

The variance of this estimator is given by

$$Var(\widehat{\pi_x}) = \frac{P_y(1-P_y)}{np^2}. \tag{II.7}$$

Greenberg et al. (1969)[16] showed that, for $p > \frac{1}{3}$, this method was in fact more efficient than [62]. Unlike Warner's model, under the unrelated question model, some respondents may experience added reassurance they can be asked to answer a potentially innocuous question which can help improve respondent cooperation(24)[24].

II.2 Homomorphic Encryption Techniques

Encryption is a method used to ensure data confidentiality with regard to communication and data storage(26)[15]. In the current day and age of cloud computing and connected servers, sensitive data is being collected and stored online on a very large scale. Data Encryption is a traditional data security method that uses existing conventional algorithms, i.e. a cipher, to secure highly confidential information. Thus, working cryptosystems usually involve three stages:

1. Key Generation: First and the most important stage requires us to generate both the encryption and decryption keys. The encryption key can be private (symmetric encryption) or public (asymmetric encryption) but the decryption key is kept private in order to ensure data privacy.

2. Encryption: The true response or plaintext is encrypted by the cipher into encrypted data or the ciphertext.

3. Decryption: Once the encrypted information has been sent over to the intended organization/person, it can be decrypted using a private decryption key.

Suppose, we collect private data through a traditional asymmetric encryption protocol in a survey with a binary response question and encrypt every individual response into a set of ciphertexts. Traditionally in order to make use of this collection of encrypted responses or to do any computation based on the true responses, one would have to have the private decryption key to first decrypt the ciphertext back into the plaintext for all individuals. Then the required computation can be performed on the set of retrieved plaintext. However, this would expose the private data of all the respondents. In a survey, where such a binary question being asked is on a sensitive topic, it can make the respondents concerned that their true responses can ultimately be uncovered and the collected data might get heavily affected by social desirability bias.

However, homomorphic encryption techniques have the unique property which allows one to simply carry out the necessary computation on the set ciphertext. The result of the computation performed would also be in the encrypted form. When this encrypted result is decrypted using the private decryption key, the resulting plaintext is the same as the outcome one would have obtained if the same computation was performed on the plaintext obtained from all individuals. Hence, one can potentially perform complex computations on the ciphertext by only decrypting the final result rather than every individual's encrypted data. Therefore, homomorphic encryption techniques give us the special ability to add two encrypted numbers and obtain the true sum of the plaintext responses without requiring individual plaintext responses(27)[47].

A few of the applications of Homomorphic encryption have been listed below [37].

- E-Cash
- E-Voting
- Private Information Retrieval
- Cloud Computing

II.2.1 Various Types of Homomorphic Encryption

1. Homomorphic Encryption (HE): Homomorphic encryption is a special kind of encryption mechanism that can resolve security and privacy issues [56]. It allows the administrators to conduct certain mathematical operations to be carried out on ciphertext, instead of on the actual data itself [6]. Therefore, it facilitates some computations (multiplication, addition, etc.) to be performed on the encrypted data without decryption. The result of such computations, when decrypted, is the same as what one would get if the operation were performed on the original data. This is a critical property of homomorphic encryption.

   Let P be a cryptosystem with encryption function E. Suppose x_i is plaintext or the true response for the i^{th} sample unit and c_i be the ciphertext for the i^{th} sample unit such that $E(x_i) = c_i$. Let Δ be some operation.

   - Additive Homomorphism: P is an additively homomorphic cryptosystem if and only if:

     $\exists \Delta : E(x_1) \Delta E(x_2) = E(x_1 + x_2)$.

   - Multiplicative Homomorphism: P is a multiplicatively homomorphic cryptosystem if and only if:

     $\exists \Delta : E(x_1) \Delta E(x_2) = E(x_1 . x_2)$.

[44]

2. Partial Homomorphic Encryption (PHE): An encryption protocol is considered partially homomorphic if it only preserves the additive operation or the multiplicative operation over the ciphertext but not both. This means that, unlike a (fully) homomorphic encryption protocol, one can either perform addition or multiplication over the encrypted data and decrypt it to either the sum or the product of original values but both operations cannot be simultaneously implemented on the ciphertext. Some examples of partially homomorphic cryptosystems are RSA (multiplicative homomorphism)[53], Elgamal (multiplicative homomorphism)[13] and Paillier (additive homomorphism)[48].

The HE and PHE methods can be utilized when the survey question has a binary response ("Yes"/"No") or when the survey question has a non-negative integer response. Although encryption techniques or protocols can be of many kinds, we restrict our discussion to the Paillier encryption technique which has additive homomorphism. This property can be utilized for the estimation of sensitive trait prevalence in the population based on a survey with a sensitive question that has "Yes"/1 or "No"/0 responses, stored after encryption using the Paillier encryption technique. Paillier encryption technique will be discussed in detail in Chapter V.

II.2.2 Paillier Encryption Protocol and Algorithm

Paillier encryption technique, a partially homomorphic encryption (PHE) protocol, offers the same amount of privacy as HE but only allows for addition to be performed on the ciphertext. This encryption scheme was proposed by Pascal Paillier (1999)[48]. The special characteristic of this encryption scheme is its additive homomorphism.

Under the Paillier encryption scheme. When we decrypt the product of two encrypted values, the result of this computation is the same as the sum of their corresponding decrypted values or plaintexts.

This property of the Paillier encryption scheme can be utilized by organizations, that outsource the processing of sensitive data to a third party, to allow for certain helpful computations to be performed of encrypted data, without a need to expose the sensitive data to a third-party organization which may or may not be trustworthy. In Chapter V, we discuss the algorithm(s) used for the Paillier encryption scheme, an example to help understand its application. We also present a discussion on the proposed work where we leverage the strengths of this scheme to help improve upon full RRT surveys.

II.3 Sensitive Mean Estimation using Quantitative RRT Models

Various binary RRT models described in section II.2 can prove to be a helpful tool in scrambling the true responses before they are reported when the sensitive question being asked in a survey has a binary response. This protects respondent privacy and helps us get an efficient estimate of the sensitive trait prevalence. However, it is possible that we are also interested in estimating the mean of a sensitive trait rather than its prevalence. Unlike the binary RRT, the true response needs to be scrambled in a different manner with the introduction of noise. For instance, for the sensitive question *"How many times in the past month have you consumed illegal drugs?"* one could respond with many possible non-negative integers. Hence the noise introduced must be such that it has a quantitative shift in the true responses. For this purpose,

several quantitative RRT models have been introduced over the past several decades. A few of the traditional quantitative RRT models have been discussed in this section.

II.3.1 Warner's (1971) and Pollock & Beck's (1976) Quantitative RRT Model

Warner[63] proposed a modified quantitative RRT model in 1971 and this work was further extended by Pollock and Beck [50]. Warner[63] proposed introducing a random additive noise to the true responses of the survey participants in order to scramble and conceal the individual responses before they are reported to the surveyor. Let Y be the sensitive variable with unknown mean μ_Y and unknown variance σ_Y^2 and S be the scrambling variable that would be added to each individual's response. The scrambling variable S is independent of the sensitive variable Y and has a known mean θ and a known variance σ_S^2. Let Z be the reported response. Then

$$Z = Y + S. \tag{II.8}$$

The expected reported response is given by

$$E(Z) = E(Y) + E(S) = \mu_Y + \mu_S. \tag{II.9}$$

Then an unbiased estimator for the mean of the sensitive variable Y is given by

$$\widehat{\mu_Y} = \bar{z} - \mu_S. \tag{II.10}$$

Therefore, if we generate the values of the scrambling variable S such that $\mu_S = 0$ then we can estimate the sensitive mean by simply computing the sample mean of

reported responses $\bar{z}$.

The variance of this estimator, under simple random sampling without replacement, is given by

$$Var(\widehat{\mu_Y}) = Var(\bar{z}) = \frac{\sigma_Z^2}{n} = \frac{\sigma_Y^2}{n} + \frac{\sigma_S^2}{n}. \tag{II.11}$$

Here, $\frac{\sigma_S^2}{n}$ is the penalty for introducing noise through this RRT model.

In practice, suppose that we are interested in estimating the mean number of times a college student has been diagnosed with an STD in a period of 2 years. We can draw a simple random sample of size n. In order to help each of the n individuals scramble their response Y_i, we can ask them to pick a card from a well-shuffled deck. Each of the cards in this deck would have a number S_i. The numbers on all the cards are generated from a distribution with a known mean θ and known variance σ_S^2. The card drawn and the number on it are concealed from the surveyor. The respondents are asked to report the sum of their true response and the number that was written on the card they pick randomly i.e. $Y_i + S_i = Z_i$. These reported Z_is are then used to estimate the sensitive mean μ_Y using the estimator shown in equation(II.10).

II.3.2 Greenberg's Quantitative Model (1971)

Greenberg et al. (1971)[17] extended the idea of noise added to the cases where the sensitive survey question has quantitative responses. The respondents are still prompted by a randomizer to respond to the sensitive question or the unrelated question. The unrelated question is now chosen such that the responses are roughly in the same order as those for the sensitive question. For instance, if the sensitive question is "How many times in the past month have you consumed illegal drugs?",

one could choose an unrelated question such as "How old are you?".

In this model, a fixed proportion p of sample respondents are prompted to respond to the sensitive question with response A and the remaining proportion $(1-p)$ of the respondents are prompted to respond to the innocuous question with response B where the mean of the non-sensitive or the innocuous variable B is known (μ_B).

Let a simple random sample of size n be drawn from the population with replacement and let the reported response be Z. Then we have

$$Z = \begin{cases} A & with\ \ probability\ \ p \\ B & with\ \ probability\ \ 1-p. \end{cases} \tag{II.12}$$

Then the expected reported response is given by

$$E(Z) = p\mu_A + (1-p)\mu_B. \tag{II.13}$$

Using this, we obtain the estimator

$$\widehat{\mu_A} = \frac{\bar{z} - (1-p)\mu_B}{p}. \tag{II.14}$$

We can estimate the mean of sensitive variable A by computing the mean of the reported response i.e. $\bar{z}$.

The variance of this estimator is given by

$$Var(\widehat{\mu_A}) = \frac{\sigma_Z^2}{np^2}. \tag{II.15}$$

II.3.3 Eichhorn and Hayre (1983)

Warner (1971)[63] had also suggested the use of multiplicative models. However, Eichhorn and Hayere (1983)[11] pursued this in greater detail.

Eichhorn and Hayere (1983)[11] proposed introducing a multiplicative noise to help respondents conceal their true responses before they report them to the surveyor. Therefore, instead of adding the random number on a card picked from a well-shuffled deck, respondents need to multiply their true response with the random number selected from a known distribution and report the product of these two numbers divided by the mean of the multiplicative noise.

Suppose Y is the sensitive variable with unknown mean μ_Y and unknown variance σ_Y^2 and T be the scrambling variable that would be multiplied by each individual's response. The scrambling variable T is independent of the sensitive variable Y and has a known mean $\mu_T = E(T)$ and a known variance σ_T^2. Let Z be the reported response. Then,

$$Z = \frac{YT}{\mu_T}. \tag{II.16}$$

It is common to assume $\mu_T = 1$. This gives us the unbiased estimator

$$\widehat{\mu_Y} = \bar{z}. \tag{II.17}$$

Therefore, we can estimate the mean of the sensitive variable Y by computing the sample mean of the reported responses i.e. $\bar{z}$. The variance of this estimator is given by

$$Var(\widehat{\mu_Y}) = \frac{1}{n}\left[\sigma_Y^2 + \frac{\sigma_T^2(\sigma_Y^2 + \mu_Y^2)}{\mu_T^2}\right] \tag{II.18}$$

Note that this model requires the respondents to perform the arithmetic correctly. However, the respondents may not either know how to perform the multiplication or may not want to perform the multiplication as a part of a survey. This can lead to both intentional and unintentional misreporting.

It must also be noted that this multiplicative model is not the best in terms of protecting respondent privacy. When the true response to the sensitive question is 0, the reported response would always be 0. For instance, if the sensitive question is "How many times in the past month have you consumed illegal drugs?" and a respondent has not indulged in drug abuse, their reported response would be zero since the random noise is being multiplied by the true response zero. This implies that the respondents who reported a non-zero value might have indulged in sensitive behavior to some extent. This could cause the respondents to not trust the model and could discourage them from participating truthfully in the survey.

II.3.4 Diana and Perri Linear Combination Model (2011)

As a researcher, one would want to get as efficient an estimate as possible. However, when we conduct a survey on a sensitive topic, we have the ethical responsibility to ensure respondent privacy as well. Based on a review of various quantitative RRT models existing in literature at the time, [10] conducted a comparative study of several variations of those models and proposed a general scrambling model for survey questions with quantitative responses in sensitive question surveys. As both the multiplicative model and the additive model have their pros and cons with respect

to privacy and efficiency, they proposed combining the additive and multiplicative approaches to optimize for the privacy protection and efficiency trade-off. The primary goal of this model was to encourage the respondents to participate because of additional privacy protection provided by the second scrambling variable.

Let T be a scrambling variable with mean μ_T and variance σ_T^2. Let S be another scrambling variable with mean μ_S and variance σ_S^2. Both T and S are independent of the sensitive study variable Y which has an unknown mean μ_Y and an unknown variance of σ_Y^2. Then the linear combination model introduced by Diana and Perri [10] is given by

$$Z = TY + S. \tag{II.19}$$

It is commonly assumed that $\mu_T = 1$ and $\mu_S = 0$. Then the expected value and the variance of the reported response Z, under simple random sampling without replacement, are given by

$$E(Z) = \mu_Y \tag{II.20}$$

and

$$Var(Z) = \sigma_S^2(\mu_Y^2 + \sigma_Y^2) + \sigma_Y^2 + \sigma_T^2. \tag{II.21}$$

If $\mu_T = 1$ and $\mu_S = 0$, then an unbiased estimator of the mean of the sensitive study variable Y is given by

$$\widehat{\mu_Y} = \frac{\bar{Z} - \mu_S}{\mu_T} = \bar{z}. \tag{II.22}$$

The variance for this estimator is given by

$$Var(\widehat{\mu_Y}) = \frac{1}{n}[\sigma_S^2(\mu_Y^2 + \sigma_Y^2) + \sigma_Y^2 + \sigma_T^2]. \tag{II.23}$$

II.4 Optional RRT Models

The RRT models described in Section II.1 and II.3 assume that all respondents would find the survey question sensitive and force every respondent to report a scrambled response. However, this may not be true. It is reasonable to assume that not every individual in the population may find the survey question sensitive. Therefore, there may not be a need to force every sample respondent to conceal their response. Hence the Optional RRT models acknowledge that sensitivity is subjective and in fact varies from person to person based on many reasons- the question or the subject of the question being an important one. The true proportion of individuals in the population that find a question sensitive, irrespective of the reason, is referred to as the sensitivity level of that question.

The key difference between the RRT models introduced in Sections II.1 and II.3 and the Optional RRT models is that the survey participants decide if they are comfortable providing a true response or if they wish to conceal it with some noise due to concerns about the protection of their privacy. This is not an option for survey participants when we use full RRT models. Therefore, it is intuitive that the data collected from a survey that uses Optional RRT would be able to capture a greater element of truth as opposed to the data from a survey using a full RRT model where even those respondents who are willing to share their true responses are not able to do so.

II.4.1 Gutpa et al.(2002) Optional Multiplicative RRT Model

Following the idea that sensitivity is subjective, Gupta et al. (2002)[19] proposed a modified version of the model proposed by Eichhorn and Hayre (1983)[11] multiplicative RRT model. This optional RRT model allows respondents, who are comfortable reporting their true responses unaltered, to do so. The respondents who find the question sensitive and would prefer to conceal their responses have the option to scramble their responses using a multiplicative noise.

Let Y be the sensitive variable with an unknown mean μ_Y and unknown variance σ_Y^2. The proportion of individuals in the population who would find the question sensitive, i.e. the sensitivity level of the question is denoted by W. Note that now sensitivity level (W) is a parameter we can estimate in addition to the mean of the sensitive variable of interest (μ_Y). Let T be the multiplicative scrambling variable that would be provided as an option to each survey participant which they use only if they find the question sensitive. The scrambling variable T is independent of the sensitive variable Y and has a known mean $\mu_T = E(T)$ and a known variance σ_T^2. Let Z be the reported response. Then under this model, the reported response Z is given by

$$Z = \begin{cases} Y & with\ probability\ W \\ YT & with\ probability\ 1 - W. \end{cases} \tag{II.24}$$

If $\mu_T = 1$ then the expected value of the reported response Z is given by

$$E(Z) = E(Y)(1 - W) + E(YT)W = \mu_Y(1 - W) + \mu_Y\mu_T W = \mu_Y. \tag{II.25}$$

Therefore, If $\mu_T = 1$ the estimator for the sensitive mean μ_Y is given by the sample mean of the reported responses,

$$\widehat{\mu_Y} = \bar{Z}. \tag{II.26}$$

If $\mu_T = 1$, the variance of this unbiased estimator $\widehat{\mu_Y}$ is given by

$$Var(\widehat{\mu_Y}) = \frac{1}{n}[\sigma_Y^2 + W\sigma_T^2(\sigma_Y^2 + \mu_Y^2)]. \tag{II.27}$$

One may note that $Var(\widehat{\mu_Y})$ increases with W, and hence there is a gain in efficiency in the optional models when compared to the non-optional model where $W = 1$.

Gupta et al.(2002) [19] also proposed an estimator for the sensitivity level W which is given by

$$\widehat{W} = \frac{\frac{1}{n}\Sigma_{i=1}^n log(Z_i) - log(\frac{1}{n}\Sigma_{i=1}^n Z_i)}{E[log(T)]}. \tag{II.28}$$

Thus Gupta et al.(2002)[19] showed that by not forcing every respondent to scramble their responses, the estimator efficiency improves when compared to the Eichhorn and Hayre (1983)[11] estimator.

II.4.2 Gupta et al. Optional Additive RRT Model (2010)

As stated earlier in section II.3.3, the multiplicative scrambling model compromises respondent anonymity. This happens because when the true response is zero, so is the reported response irrespective of the multiplicative scrambling variable value. This indicates that respondents with non-zero reported responses have indulged in sensitive

behavior to some extent. Another issue with the multiplicative scrambling model is that some respondents may either not like to multiply or may not know how to multiply the scrambling variable correctly. Under such a circumstance, the respondents still provide incorrect responses. [57] showed that this case is more dangerous than not using the scrambled response.

To help address these problems and also to estimate the sensitivity level W without using any approximations, [22] proposed an additive Optional RRT (ORRT) model using a split-sample approach such that different additive scrambling variables are used for each sub-sample. Using this method, we split the sample respondents into two sub-samples. Let Y be the sensitive variable with an unknown mean μ_Y and unknown variance σ_Y^2. Let S_i be the additive noise used by respondents in the i^{th} sub-sample with a known mean of μ_{S_i} and a known variance of $\sigma_{S_i}^2$. Let W be the true sensitivity level of the question in the population and Z_i be the reported response in the i^{th} sub-sample ($i = 1, 2$). Then following the concept of optionality in RRT models, if the respondents find the survey question sensitive, they will provide a scrambled response. Otherwise, they give the true unscrambled response. Under this model, the reported response Z is given by

$$Z_i = \begin{cases} Y + S_i & with\ probability\ W \\ Y & with\ probability\ 1 - W \end{cases}, \tag{II.29}$$

where $i = 1, 2$.

Then the expected value and the variance of the reported response for the two sub-samples are given by

$$E(Z_i) = \mu_Y + \mu_{S_i} W, \tag{II.30}$$

$$Var(Z_i) = \sigma_Y^2 + \sigma_{S_i}^2 W + \mu_{S_i}^2 W(1-W), \tag{II.31}$$

where $\mu_{S_i} = E(S_i)$, $(i = 1, 2)$.

The unbiased estimators for the mean of sensitive variable Y and the sensitivity level W are given by

$$\widehat{\mu_Y} = \frac{\mu_{S_1}\bar{Z}_2 - \mu_{S_2}\bar{Z}_1}{\mu_{S_1} - \mu_{S_2}} \tag{II.32}$$

and

$$\widehat{W} = \frac{\bar{Z}_1 - \bar{Z}_2}{\mu_{S_1} - \mu_{S_2}}. \tag{II.33}$$

The corresponding variances for the estimators of μ_Y and W, under simple random sampling without replacement, are given by

$$Var(\widehat{\mu_Y}) = \frac{1}{(\mu_{S_2} - \mu_{S_1})^2}\left[\mu_{S_2}^2 \frac{\sigma_{Z_1}^2}{n_1} + \mu_{S_1}^2 \frac{\sigma_{Z_2}^2}{n_2}\right] \tag{II.34}$$

and

$$Var(\widehat{W}) = \frac{1}{(\mu_{S_2} - \mu_{S_1})^2}\left[\frac{\sigma_{Z_1}^2}{n_1} + \frac{\sigma_{Z_2}^2}{n_2}\right], \mu_{S_1} \neq \mu_{S_2}. \tag{II.35}$$

II.5 Use of Auxiliary Information in Quantitative RRT Surveys

As discussed previously, RRT models are an effective method to collect data on sensitive topics as they protect the privacy of the respondents. Several field studies,

such as the one by Chow et al. (1979)[5], Chaloupka (1985)[3] and Chhabra et al. (2016)[4], have also confirmed the efficacy of RRT models in acquiring information from the respondents. Another such study was conducted by Kwan et al. (2010)[36] on software piracy and their results confirmed that the respondents that responded to the question using RRT were more willing to provide a true response about the behaviors on software piracy. Moreover, we know that the accuracy of our results is enhanced when we use mean estimators that make use of auxiliary information (e.g. ratio, regression, product estimators, etc.) rather than the ordinary mean estimator when the auxiliary variable is highly correlated with the variable of interest. Therefore, if we have a situation where we have information on a non-sensitive variable that is highly correlated with the sensitive variable, we can take simultaneous advantage of RRT methodology and the auxiliary information.

In this section, we introduce a few estimators from the literature that make use of strongly correlated auxiliary variables and improve upon the estimation of the sensitive variable compared to corresponding estimators that do not utilize the auxiliary information.

II.5.1 Mean Estimation using Auxiliary Variables

Ratio and regression estimators are examples of the use of auxiliary information in mean estimation. In some situations, we may know the value for the auxiliary variable X for the entire population. Let X_i be the sample values for the auxiliary variable X and Y_i be the sample values for the study variable Y, $(i = 1, 2....n)$. We consider the ratio estimator in those situations when the study variable Y and the auxiliary variable X have a roughly linear relationship through the origin. Therefore in such a case, it is reasonable to assume that when sample value X_i is zero the sample value Y_i

will be zero [60]. We may assume the model

$$Y = RX + \epsilon, \tag{II.36}$$

where $E[\epsilon] = 0$. Let μ_X be the true known mean of the auxiliary variable and μ_Y be the unknown mean of the study variable. Then,

$$\mu_Y = R\mu_X. \tag{II.37}$$

Then the ratio estimator of the population mean μ_Y is given by

$$\widehat{\mu_r} = r\mu_x, \tag{II.38}$$

where the sample ratio r is given by

$$r = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i} = \frac{\bar{y}}{\bar{x}} \tag{II.39}$$

Here, r shown in equation (II.39) is an estimate of the population ratio $R = \frac{\mu_Y}{\mu_X}$.

The approximate mean square error or variance of the ratio estimator is given by,

$$var(\hat{\mu_r}) \approx \left(\frac{N-n}{N}\right)\frac{\sigma_r^2}{n} \tag{II.40}$$

where,

$$\sigma_r^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - RX_i)^2 \tag{II.41}$$

and

$$R = \frac{\Sigma_{i=1}^{N} Y_i}{\Sigma_{i=1}^{N} X_i} = \frac{\tau_y}{\tau_x} = \frac{\mu_Y}{\mu_X}. \tag{II.42}$$

The traditional estimator of the mean square error of the ratio estimator is given by,

$$\widehat{Var(\hat{\mu}_r)} = \left(\frac{N-n}{N}\right)\frac{s_r^2}{n}, \tag{II.43}$$

where

$$s_r^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - rX_i)^2. \tag{II.44}$$

When the auxiliary variable X is approximately linearly related to the study variable Y, but Y_i is not zero when X_i is zero, using a linear regression estimator is more appropriate than using a ratio estimator [60].

Using the same notations introduced for ratio estimators, the (linear) regression estimator for the population mean μ is given by,

$$\widehat{\mu_L} = a + b\mu_x = \bar{y} + b(\mu_x - \bar{x}). \tag{II.45}$$

The value of b gives the slope and a gives the y-intercept of a straight line fitted to the data by least squares.

where,

$$b = \frac{\sum_{i=1}^{n}(X_i-\bar{x})(Y_i-\bar{y})}{\sum_{i=1}^{n}(X_i-\bar{x})^2}$$
$$a = \bar{y} - b\bar{x}$$

The variance of this regression estimator can be approximated by,

$$Var(\widehat{\mu_L}) \approx \frac{N-n}{Nn(N-1)} \sum_{i=1}^{N} (Y_i - A - BX_i)^2 \tag{II.46}$$

where,

$$B = \frac{\sum_{i=1}^{N}(X_i-\mu_x)(Y_i-\mu)}{\sum_{i=1}^{N}(X_i-\mu_x)^2}$$
$$A = \mu - B\mu_x$$

An estimator for this variance is given by,

$$\widehat{Var(\widehat{\mu_L})} = \frac{N-n}{Nn(n-2)} \sum_{i=1}^{n} (Y_i - a - bX_i)^2 \tag{II.47}$$

II.5.2 Mean Estimation using Auxiliary Variables under RRT Models

In Section II.5.1, we discussed the ordinary ratio and regression estimators when both the study variable Y and the auxiliary variable X are non-sensitive. These estimators are more efficient in situations in which highly correlated auxiliary information X is available on every unit in the population. However, in sensitive question surveys, the study variable Y is sensitive. Normally, one would expect X to be also sensitive. However, Sousa et al. (2010)[58] pointed out that it is possible that Y is sensitive and X is non-sensitive. Since the information on X is available on all population units, it is not a variable of direct interest. However, if this auxiliary variable X is highly correlated with our sensitive study variable Y, we can utilize this auxiliary information to improve the estimation of the sensitive mean μ_Y. Thus we can derive ratio and regression estimators under various RRT models and sampling techniques to improve the sensitive mean estimation.

Ratio Estimation under RRT Models

Sousa et al. (2010)[58] proposed a ratio estimator under the additive RRT model[50]. They estimated the mean of the sensitive variable using an improved estimator based on a non-optional RRT model by utilizing a non-sensitive auxiliary variable. Let, Y be the sensitive study variable that can not be observed directly and X be the non-sensitive auxiliary variable (positively correlated with Y). Also, let S be scrambling variable (independent of both X and Y) with mean $\mu_S = 0$ and variance σ_S^2. Let μ_X be the known true population mean and σ_X^2 be the known variance of the non-sensitive auxiliary variable X. Let μ_Y be the unknown true population mean and σ_Y^2 be the unknown variance of the sensitive study variable Y. Then we know that under the non-optional Pollock and Beck (1976) model the reported response Z is given by

$$Z = Y + S.$$

Assuming, $E(S) = 0$, we get $E(Z) = E(Y)$ and the unbiased ordinary estimator under this RRT model is given by,

$$\widehat{\mu_o} = \bar{z}, \tag{II.48}$$

and the MSE of this ordinary mean estimator is given by

$$MSE(\widehat{\mu_o}) = \lambda(\sigma_Y^2 + \sigma_S^2), \tag{II.49}$$

where $\lambda = \frac{(N-n)}{Nn}$ and N and n are the size of the finite population and the simple random sample drawn from it respectively.

Based on this ordinary mean estimator, Sousa et al. (2010)[58] proposed the ratio estimator for the mean of the sensitive variable Y given by

$$\widehat{\mu_R} = \bar{z}\left(\frac{\mu_X}{\bar{x}}\right), \tag{II.50}$$

where $\bar{z}$ is the sample mean of reported responses and $\bar{x}$ is the sample mean of an auxiliary variable.

The mean squared error (MSE) of the ratio estimator, correct up to the first-order approximation, is given by

$$MSE(\widehat{\mu_R}) \approx \lambda\mu_z^2(C_x^2 + C_z^2 - 2\rho_{zx}C_zC_x), \tag{II.51}$$

where $\lambda = \frac{N-n}{nN}$, $C_z = S_z/\mu_z$ & $C_x = S_x/\mu_x$ are the coefficients of variation for Z and X respectively and $\rho_{zx} = S_{zx}/S_zS_x$ is the population correlation coefficient between Z and X.

It was observed that $MSE(\widehat{\mu_R}) < MSE(\widehat{\mu}_o)$ if

$$\rho > \frac{1}{2}\frac{C_X}{C_Y}\sqrt{1 + \frac{\sigma_S^2}{\sigma_Y^2}}. \tag{II.52}$$

Sousa et al. (2010)[58] compared the efficiency of the ratio estimator, in terms of its MSE in equation (II.51), with the MSE of the ordinary mean estimator from equation (II.49), it was established that the estimator proposed by Sousa et al. (2010)[58] is more efficient.

Ratio Estimation under Optional RRT Models

Sousa et al. (2010)[58] improved the estimation of the sensitive mean through their proposed ratio estimator under the non-optional Pollock and Beck (1976)[50] model. However, Gupta et al. (2002)[19] established that forcing all respondents to scramble their responses, irrespective of whether they find the question sensitive or not, hurts

the efficiency of the model. Gupta et al. (2014)[18] modified the work done by Sousa et al. (2010)[58] by introducing the optionality element into their model and by using the split-sample technique. If W is the sensitivity level of the survey question i.e. the proportion of individuals in the population that find the question sensitive, then according to Gupta et al. (2010), the reported response Z_i in the i^{th} sub-sample is given by

$$Z_i = \begin{cases} Y + S_i & with\ probability\ W \\ Y & with\ probability\ 1 - W \end{cases}, \tag{II.53}$$

where $i = 1, 2$.

This work was further improved by Kalucha et al.(2015)[29]. They proposed two ratio estimators of a finite population for the sensitive mean using the Optional RRT model shown in equation (II.53). These estimators were called the additive and the multiplicative ratio estimators.

This additive ratio estimator is given by

$$\widehat{\mu_{AR}} = \frac{1}{2}\left[\frac{\mu_{S_2}\bar{z_1} - \mu_{S_1}\bar{z_2}}{\mu_{S_2} - \mu_{S_1}}\right]\left[\frac{\mu_X}{\bar{x_1}} + \frac{\mu_X}{\bar{x_2}}\right], \tag{II.54}$$

The MSE for this additive ratio estimator, correct up to the first-order approximation, is given by

$$\begin{aligned} MSE(\widehat{\mu_{AR}}) \approx \lambda_1\left[\left(\frac{\mu_{S_2}}{\mu_{S_2} - \mu_{S_1}}\right)^2\sigma^2_{Z_1} + \frac{1}{4}\mu_Y^2 C_X^2 - \mu_Y\rho_{YX}\sigma_Y\left(\frac{\mu_{S_2}}{\mu_{S_2} - \mu_{S_1}}\right)\right] \\ + \lambda_2\left[\left(\frac{\mu_{S_1}}{\mu_{S_2} - \mu_{S_1}}\right)^2\sigma^2_{Z_2} + \frac{1}{4}\mu_Y^2 C_X^2 - \mu_Y\rho_{YX}\sigma_Y\left(\frac{\mu_{S_1}}{\mu_{S_2} - \mu_{S_1}}\right)\right]. \end{aligned} \tag{II.55}$$

The multiplicative ratio estimator is given by

$$\widehat{\mu_{MR}} = \left[\frac{\mu_{S_2}\bar{z_1} - \mu_{S_1}\bar{z_2}}{\mu_{S_2} - \mu_{S_1}}\right]\left[\frac{\mu_X}{\bar{x_1}}\right]\left[\frac{\mu_X}{\bar{x_2}}\right]. \tag{II.56}$$

The MSE for this multiplicative ratio estimator, correct up to the first-order approximation is given by

$$\begin{aligned} MSE(\widehat{\mu_{MR}}) \approx \lambda_1\left[\left(\frac{\mu_{S_2}}{\mu_{S_2} - \mu_{S_1}}\right)^2\sigma_{Z_1}^2 + \mu_Y^2 C_X^2 - 2\mu_Y\rho_{YX}\sigma_Y\left(\frac{\mu_{S_2}}{\mu_{S_2} - \mu_{S_1}}\right)\right] \\ + \lambda_2\left[\left(\frac{\mu_{S_1}}{\mu_{S_2} - \mu_{S_1}}\right)^2\sigma_{Z_2}^2 + \mu_Y^2 C_X^2 - 2\mu_Y\rho_{YX}\sigma_Y\left(\frac{\mu_{S_1}}{\mu_{S_2} - \mu_{S_1}}\right)\right] \end{aligned} \tag{II.57}$$

The geometric mean ratio estimator was introduced by Zhang et al. (2019)[68], which improves the multiplicative ratio estimator in equation (II.56). The geometric mean ratio estimator is given by

$$\widehat{\mu_{GMR}} = \left[\frac{\mu_{S_2}\bar{z_1} - \mu_{S_1}\bar{z_2}}{\mu_{S_2} - \mu_{S_1}}\right]\sqrt{\left(\frac{\mu_X}{\bar{x_1}}\right)\left(\frac{\mu_X}{\bar{x_2}}\right)}. \tag{II.58}$$

The MSE for this geometric mean ratio estimator, correct up to the first order of approximation, is given by

$$\begin{aligned} MSE^{(1)}(\widehat{\mu_{GMR}}) \approx \lambda_1\left[\left(\frac{\mu_{S_2}}{\mu_{S_2} - \mu_{S_1}}\right)^2\sigma_{Z_1}^2 + \frac{1}{4}\mu_Y^2 C_X^2 - \mu_Y\rho_{YX}\sigma_Y\left(\frac{\mu_{S_2}}{\mu_{S_2} - \mu_{S_1}}\right)\right] \\ + \lambda_2\left[\left(\frac{\mu_{S_1}}{\mu_{S_2} - \mu_{S_1}}\right)^2\sigma_{Z_2}^2 + \frac{1}{4}\mu_Y^2 C_X^2 - \mu_Y\rho_{YX}\sigma_Y\left(\frac{\mu_{S_1}}{\mu_{S_2} - \mu_{S_1}}\right)\right]. \end{aligned} \tag{II.59}$$

Zhang et al.(2019)[68] established that the geometric mean estimator is always more efficient than the multiplicative ratio estimator and it is more efficient than the ordinary RRT mean estimator when the correlation coefficient between X and Y is greater than $\frac{1}{2}$. They also show that the geometric mean estimator is approximately as efficient as the additive ratio estimator.

Regression Estimation under RRT Models

Gupta et al. (2012)[23] proposed a regression estimator to estimate the mean of the sensitive variable when we have, a non-sensitive but highly correlated, auxiliary information available for every unit in the population. This work was done under the non-optional Pollock and Beck (1976)[50] model. If the sensitive variable Y and auxiliary variable X have a linear relationship, this regression estimator is given by

$$\widehat{\mu_{Reg}} = \bar{z} + \hat{\beta}_{ZX}(\mu_X - \bar{x}), \tag{II.60}$$

where $\hat{\beta}_{ZX} = \frac{s_{ZX}}{s_x^2}$ is the sample estimate of the regression coefficient between reported response $Z = Y + S$ and auxiliary variable X.

The Bias and the MSE of this estimate, accurate up to the first order of approximation, are given by

$$Bias(\widehat{\mu_{Reg}}) \approx -\beta_{ZX}\lambda\left[\frac{\mu_{12}}{\mu_{11}} - \frac{\mu_{03}}{\mu_{02}}\right], \tag{II.61}$$

and

$$MSE(\widehat{\mu_{Reg}}) \approx \lambda\mu_Y^2 C_Z^2(1-\rho_{ZX}^2) = \lambda\sigma_Y^2\left[\left(1+\frac{\sigma_S^2}{\sigma_Y^2}\right) - \rho_{YX}^2\right]. \tag{II.62}$$

Here, $\beta_{ZX} = \frac{\sigma_{ZX}}{\sigma_X^2} = \frac{\sigma_{YX}}{\sigma_X^2} = \rho_{YX}\frac{\sigma_Y}{\sigma_X} = \beta_{YX}$ and $\rho_{ZX} = \frac{\rho_{YX}}{\sqrt{1+\frac{\sigma_S^2}{\sigma_Y^2}}}$. Also, we have $\lambda = \left(\frac{1-f}{n}\right)$ and $\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(Z_i - \mu_Z)(X_i - \mu_X)$.

The following can be noted regarding the performance of the regression estimator given by equation (II.60)

- The RRT regression estimator(II.60) is more efficient than the RRT sample mean estimator if $\rho_{YX}^2 > 0$, and
- The RRT regression estimator(II.60) is more efficient than the RRT ratio estimator(II.50) if $(C_X - C_Z\rho_{ZX})^2 > 0$.

Considering expressions accurate for first-order approximations, these two conditions are always true. Hence, this regression estimator performs better than both the ordinary RRT estimator(II.48) and the RRT ratio estimator(II.50).

Regression Estimation under Optional RRT Models

Gupta et al.(2017) [20] also proposed a modified version of the estimator proposed by Gupta et al. (2012)[23] under the optional version of the Pollock and Beck (1976)[50] model. They use the split sample approach to obtain this regression estimator for the sensitive mean, as given by

$$\widehat{\mu_{Areg}} = \left(\frac{\theta_2\bar{z_1} - \theta_1\bar{z_2}}{\theta_2 - \theta_1}\right) + [\widehat{\beta_{Z_1X_1}}(\mu_X - \bar{x_1}) + \widehat{\beta_{Z_2X_2}}(\mu_X - \bar{x_2})]\left(\frac{1}{2}\right). \tag{II.63}$$

Here, $\bar{z_i}$ and $\bar{x_i}$ $(i = 1, 2)$ are the sample mean for the reported response and the auxiliary information in the i^{th} sub-sample and $\widehat{\beta_{Z_iX_i}}(i = 1, 2)$ are the sample regression coefficients between Z_i and X_i.

The MSE for this estimator (II.63), correct up to the first order of approximation, is given by

$$MSE(\widehat{\mu_{Areg}}) \approx \left(\frac{1-f_1}{n_1}\right)\left[\left(\frac{\theta_2}{\theta_2-\theta_1}\right)^2\sigma^2_{Z_1} + \frac{1}{4}\beta^2_{Z_1X}\sigma^2_{X_1} - \left(\frac{\theta_2}{\theta_2-\theta_1}\right)\beta_{Z_1X}\sigma_{Z_1X_1}\right] + \left(\frac{1-f_2}{n_2}\right)\left[\left(\frac{\theta_1}{\theta_2-\theta_1}\right)^2\sigma^2_{Z_2} + \frac{1}{4}\beta^2_{Z_2X}\sigma^2_{X_2} + \left(\frac{\theta_1}{\theta_2-\theta_1}\right)\beta_{Z_2X}\sigma_{Z_2X_2}\right], \tag{II.64}$$

where, $\theta_1 \neq \theta_2$ and

$$\sigma^2_{Z_i} = \sigma^2_Y + W\sigma^2_{S_i} + \theta_i^2 W(1-W), i = 1,2 \ \textit{and}$$
$$\beta_{Z_iX} = \frac{\sigma_{Z_iX}}{\sigma^2_X} = \frac{\sigma_{YX}}{\sigma^2_X} = \frac{\rho_{YX}\sigma_Y}{\sigma_X}, i = 1,2.$$

Also, $\sigma_{Z_iX_i} = \sigma_{YX_i}\sigma_{YX} = \rho_{YX}\sigma_Y\sigma_X$, $\sigma^2_{X_i} = \sigma^2_X$ and $\rho_{Z_iX} = \frac{\rho_{YX}\sigma_Y}{\sigma_{Z_i}}$ for $i = 1,2$. Therefore, the MSE for the regression estimator shown in equation (II.63), correct up to the first order of approximation, can be re-written as

$$MSE^{(1)}(\widehat{\mu_{Areg}}) \approx \frac{1}{(\theta_2-\theta_1)^2}\left[\theta_2^2\left(\frac{1-f_1}{n_1}\right)\sigma^2_{Z_1} + \theta_1^2\left(\frac{1-f_2}{n_2}\right)\sigma^2_{Z_2}\right] + \frac{\rho^2_{YX}\sigma^2_Y}{4}\alpha - \rho^2_{YX}\sigma^2_Y\beta, \tag{II.65}$$

where $\theta_1 \neq \theta_2$, $\alpha = \left(\frac{1-f_1}{n_1}\right) + \left(\frac{1-f_2}{n_2}\right)$ and $\beta = \left(\frac{1-f_1}{n_1}\right)\left(\frac{\theta_2}{\theta_2-\theta_1}\right) - \left(\frac{1-f_2}{n_2}\right)\left(\frac{\theta_1}{\theta_2-\theta_1}\right)$.

The following can be noted regarding the performance of the regression estimator given by equation (II.63):

- The regression estimator proposed in equation (II.63) is more efficient than the ordinary optional RRT estimator if

$$\rho_{YX}^2 \sigma_Y^2 \left(\frac{\alpha}{4} - \beta \right) < 0, \tag{II.66}$$

- The regression estimator proposed in equation (II.63) is more efficient than the optional RRT ratio estimator in equation (II.54) if

$$\rho_{YX} < \frac{\alpha}{4\beta - \alpha} \ when \ C_Y \approx C_X. \tag{II.67}$$

These conditions are always true when the two sub-samples are of equal size i.e. $n_1 = n_2$.

Generalized Estimators under RRT Models

Many researchers have proposed various combined generalizations of the ratio and the regression estimators. Some of these works include Perri (2003), Kadilar and Cingi (2004) and Nangsue (2009). Gupta et al. (2012)[23] also proposed a generalized regression-cum-ratio estimator to see if one could improve any aspect of sensitive mean estimation from the generalization of the regression estimators described previously. This estimator was also proposed for a non-optional Pollock and Beck (1976)[50] model and is given by

$$\widehat{\mu_{GRR}} = [k_1 \bar{z} + k_2(\mu_X - \bar{x})] \left[\frac{\mu_X}{\bar{x}} \right], \tag{II.68}$$

where k_1 and k_2 are unknown parameters. The bias for this estimator, accurate up to the first-order approximation, is given by

$$Bias(\widehat{\mu_{GRR}}) \approx (k_1 - 1)\mu_Z + \lambda k_1 \mu_Z (C_X^2 - \rho_{ZX} C_Z C_X) + \lambda k_2 \mu_X C_X^2. \tag{II.69}$$

The optimal values of k_1 and k_2, and the corresponding minimum value of the MSE of this estimator are given by

$$k_{1(opt)} = \frac{1 - \lambda C_X^2}{1 - \lambda[C_X^2 - C_Z^2(1 - \rho_{ZX}^2)]}, \tag{II.70}$$

$$k_{2(opt)} = \frac{\mu_Y}{\mu_X}\left[1 + k_{1(opt)}\left(\frac{\rho_{ZX} C_Z}{C_X} - 2\right)\right], and \tag{II.71}$$

$$MSE(\widehat{\mu_{GRR}})_{min} \approx \mu_Y^2 \frac{\lambda C_Z^2(1 - \rho_{ZX}^2)(1 - \lambda C_X^2)}{\lambda C_Z^2(1 - \rho_{ZX}^2) + (1 - \lambda C_X^2)}. \tag{II.72}$$

The following can be noted regarding the performance of the regression estimator given by equation (II.68) with minimum MSE:

- The generalized regression-cum-ratio estimator (II.68) is more efficient than the ordinary RRT estimator if

$$\lambda(\sigma_Y^2 + \sigma_S^2) > 0 \tag{II.73}$$

- The generalized regression-cum-ratio estimator (II.68) is more efficient than the RRT ratio estimator if

$$\left(\frac{C_X}{C_Z} - \rho_{ZX}\right)^2 + \frac{\lambda C_Z^2(1 - \rho_{ZX}^2)}{\lambda C_Z^2(1 - \rho_{ZX}^2) + (1 - \lambda C_X^2)} > 0. \tag{II.74}$$

- The generalized regression-cum-ratio estimator (II.68) is more efficient than the

RRT regression estimator if

$$\lambda C_Z^2(1-\rho_{ZX}^2) > 0. \tag{II.75}$$

These conditions are always true. Therefore, the generalized regression-cum-ratio estimator with optimal coefficients is always better than the ordinary RRT sample mean, RRT regression and RRT ratio estimators.

Generalized RRT Estimator by Khalil et al. (2018)

Measurement errors are a type of non-sampling error that occurs quite commonly in any survey. These can be a source of concern for researchers, especially when the survey question is on a sensitive topic and the quality of the data collected could get severely impacted by other elements such as social desirability bias (SDB) and intentional misreporting. Hence Khalil et al. (2018)[32] proposed a generalized estimator accounting for measurement errors on both the reported response Z and the non-sensitive auxiliary variable X which is positively correlated with the sensitive study variable Y. This work was done under the non-optional Pollock and Beck (1976)[50] model and is given by

$$\widehat{\mu_Z} = \left[\bar{z} + k(\mu_X - \bar{x})\right]\left(\frac{\mu_D}{\bar{d}}\right)^g, \tag{II.76}$$

where $\bar{d} = \lambda(\alpha\bar{x}+\beta) + (1-\lambda)(\alpha\mu_X+\beta)$ and $\mu_D = \alpha\mu_X + \beta$. Here k and g are suitable constants. λ is an unknown constant which is determined from the optimality conditions. Further, α and β are known parameters of the auxiliary variable X. Various series of estimators can be obtained by using different values of g, k, λ, α and β. Using $g = 1$ will generate a series of ratio estimators and using $g = -1$ will

generate a series of product estimators. It must be noted that we do not discuss product estimators in this dissertation.

If we consider the scenario with no measurement errors on Z and X, then the MSE of this generalized estimator(II.76) is given by

$$MSE^*(\widehat{\mu_Z}) \approx \theta[\sigma_S^2 + g^2\lambda^2R^2\sigma_X^2 + k^2\sigma_X^2 - 2g\lambda R\rho_{ZX}\sigma_Z\sigma_X + 2g\lambda kR\sigma_X^2], \quad \text{(II.77)}$$

where $R = \alpha\mu_Z/(\alpha\mu_Z + \beta)$. By minimizing the MSE shown in equation (II.77), the optimal value of λ is given by

$$\lambda_{opt} = \frac{\rho_{ZX}\sigma_Z\sigma_X - k\sigma_X^2}{gR\sigma_X^2}. \quad \text{(II.78)}$$

Thus, the minimum value of MSE for this generalized estimator, obtained by substituting the optimum value of λ is given by

$$MSE^*_{min}(\widehat{\mu_Z}) \approx \theta\left[\sigma_Z^2 - \frac{\rho_{ZX}^2\sigma_Z^2\sigma_X^2}{\sigma_X^2}\right] = \theta\sigma_Z^2(1 - \rho_{ZX}^2). \quad \text{(II.79)}$$

It can be noted that the minimum value of the MSE for this generalized estimator when there are no measurement errors is the same as the approximate variance of the linear regression estimator. The following can also be noted regarding the performance of the generalized estimator given by equation (II.76) when it has minimum MSE when $\lambda = \lambda_{opt}$:

- The generalized RRT estimator (II.76) is more efficient than the ordinary RRT estimator if

$$\frac{\rho_{ZX}^2\sigma_Z^2\sigma_X^2}{\sigma_X^2} > 0 \ \textit{or} \ \rho_{ZX}^2\sigma_Z^2 > 0. \quad \text{(II.80)}$$

- The generalized RRT estimator (II.76) is more efficient than the RRT ratio estimator as proposed by Sousa et al. (2010)[58] if

$$(R\sigma_X - \rho_{ZX}\sigma_Z)^2 > 0. \tag{II.81}$$

Chapter III: Mixture Binary RRT Models with a Unified Measure of Privacy and Efficiency

As discussed in Chapter I, researchers in social sciences often want to collect data on sensitive topics through a survey. Since mail and telephone surveys have a poor response rate, they often resort to in-person interviews for collecting their data. However, asking sensitive questions, face-to-face could lead to a high non-response rate and intentional misreporting, i.e. untruthful responses, due to social desirability bias (SDB). One method to circumvent the SDB in such a survey is to implement an appropriate RRT model such as Warner's indirect question model[62] or the binary unrelated question model[16]. However, Young et al. (2019)[67] showed that the estimates of sensitive trait prevalence are negatively biased if respondents are not convinced about their privacy being protected. This would occur since a lot less proportion of respondents would admit to having engaged in the sensitive behaviour because they may not be convinced about how well their response is protected from the surveyor. In section II.1, two of the most common binary RRT models were introduced. However, both these models assume that once the respondents know that they are participating in an RRT survey, they have no reason to lie. However, this may not be true. Although many more models have been proposed since for the estimation of the sensitive trait prevalence, in this chapter, we primarily focus on

Warner's indirect question model[62], Greenberg's unrelated question model[16] and the modified unrelated question model based on the concept of accounting for lack of trust. The idea of accounting for lack of trust was first proposed by Young et al. (2019)[67] in the context of Greenberg's unrelated question model (1969)[16].

In Section III.1, we first introduce a modified Greenberg et al. (1969)[16] model that accounts for the lack of trust in the model. Next in Section III.2 the proposed Mixture Binary RRT model will be introduced and its performance in terms of efficiency and privacy will be summarized. Section III.3 will present the simulation results and Section III.4 will provide concluding remarks for this Chapter[1].

III.1 Accounting for Lack of Trust in RRT Models

One of the most common consequences of asking a sensitive question in a survey is untruthful responses or misreporting. This means that sensitivity to a question could make a survey participant give an answer that does not conform to their true status with respect to that sensitive question, resulting in inaccurate answers, a form of measurement error [65]. This could be both intentional and unintentional. However, Tourangeau and Yan (2007)[61] argue that misreporting to sensitive questions largely arises from survey respondents editing their answers prior to reporting them in order to avoid embarrassing themselves. This behavior is a source of response errors unique to sensitive questions. Tourangeau and Yan (2007)[61] summarized that survey respondents tend to over-report socially desirable traits and behavior (volunteer work, benevolent acts, participation in election voting, etc.) and they tend to under-report

[1]This is an Accepted Manuscript of an article published by Taylor & Francis in Journal of Communications in Statistics - Simulation and Computation on 24 April 2021, available online: https://www.tandfonline.com/doi/full/10.1080/03610918.2021.1914092 and this chapter has a version of this work.

socially undesirable traits and behavior such as indulgence in illegal activities, drug & alcohol abuse, etc.

Yet most of the traditional RRT models operate under the assumption that the respondent is convinced about their privacy protection and trusts the model and hence only responds truthfully. However, despite the additional privacy provided by RRT models, some respondents may still provide an untruthful response [67]. Therefore, the estimation of sensitive trait prevalence or mean estimation of sensitive variables in a population while accounting for the level of untruthfulness is not as straightforward as one might think.

Young et al. (2019)[67] studied the impact of untruthful responses and misreporting in cases where the sensitive question has binary responses. They looked at this problem in the context of Greenberg et al. (1969)[16] model. They showed that even if a small proportion of respondents are providing untruthful responses in a survey, considerable bias is introduced in the estimates. Following their work, we consider Greenberg et al.(1969)[16] under the possible scenarios regarding the respondents' trust in the RRT model (Figure III.1).

Let π_X be the true prevalence of the sensitive trait in the population and let π_Y be the true prevalence of the non-sensitive unrelated trait in the population. If p is the proportion of sample respondents that are asked the sensitive question under the Unrelated Question model, then the probability of a "Yes" response (P_y^*) for the model shown in Figure III.1 is given by

$$P_y^* = p\pi_x A + (1-p)\pi_y. \tag{III.1}$$

Here A is the proportion of individuals that trust the model i.e. level of truthfulness.

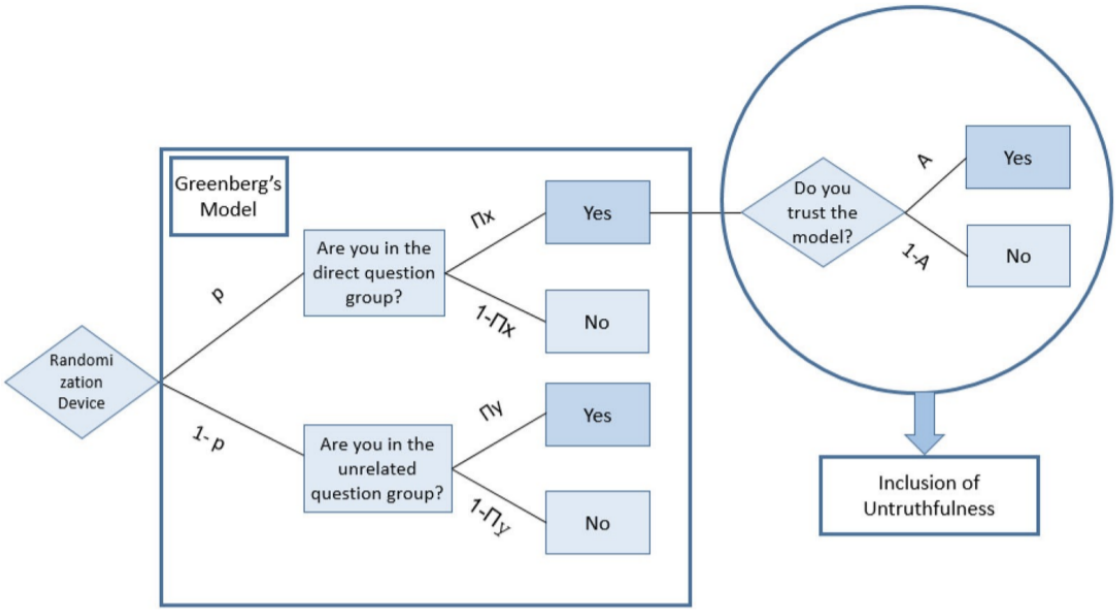

Figure III.1. Lack of Trust and Untruthfulness in Greenberg Model

If one ignores untruthfulness (i.e. $A = 1$), one would end up using the Greenberg et al. (1969)[16] estimator given by

$$\hat{\pi}_x^* = \frac{\hat{P}_y^* - (1-p)\pi_y}{p}. \tag{III.2}$$

Therefore, if untruthful responding (i.e. A) is ignored

$$E(\hat{\pi}_x^*) = \frac{E(\hat{P}_y^*) - (1-p)\pi_y}{p} = \frac{p\pi_x A + (1-p)\pi_y - (1-p)\pi_y}{p} = A\pi_x. \tag{III.3}$$

This then leads to a bias in the estimator given by

$$Bias(\hat{\pi}_x^*) = E[\hat{\pi}_x^*] - \pi_x = A\pi_x - \pi_x = \pi_x(A-1). \tag{III.4}$$

Hence, according to Young et al. (2019)[67], when a researcher fails to account for an untruthful response in the Greenberg et al.(1969) [16] model, a bias is introduced into the estimator. In Lovig et al.(2021) [41], we examined the impact of untruthfulness in the Greenberg's Unrelated Question model[16] by running extensive simulations. The results of this simulation study have been shown in Table III.1.

Table III.1. Estimates averaged over 10000 simulations with Greenberg model with untruthfulness, $n = 500$, $\pi_x = 0.3$, $\pi_y = 0.1$

p	**1 − p**	**A**	$\hat{\pi}_{\mathbf{x}}$	$\widehat{\mathbf{MSE}}$	**MSE**
0.5	0.5	1	0.29984	0.00128	0.00128
0.5	0.5	0.9	0.27023	0.00209	0.00128
0.5	0.5	0.8	0.23967	0.00478	0.00128
0.7	0.3	1	0.30023	0.00076	0.00075
0.7	0.3	0.9	0.27006	0.0016	0.00075
0.7	0.3	0.8	0.24035	0.00424	0.00075
0.9	0.1	1	0.30028	0.0005	0.0005
0.9	0.1	0.9	0.26992	0.00137	0.0005
0.9	0.1	0.8	0.23972	0.00407	0.0005

We noted that when $A < 1$, i.e. when respondents do not trust the model, our estimates show significant errors. Therefore, some level of untruthful responses is an important consideration and can significantly impact the reliability of the RRT estimators.

Moreover, results by Young et al.(2019)[67] showed that when the sensitive trait prevalence is high or when the proportion of people responding dishonestly is high, their model demonstrates an efficiency higher than that for [16]. Further, they showed that when the proportion of untruthful responses is high or the prevalence of the sensitive trait is high, the proposed model can offer a large reduction in sample size while achieving the same efficiency as other models.

Thus, Young et al.(2019)[67] established that even if only a small number of

respondents lie, a significant bias may be introduced to the model. A detailed description and analysis has been presented in Chapters III, IV and VI, of our proposed models that address the lack of trust in traditional binary and quantitative RRT models. In Chapter III, we restrict our discussion to account for the lack of trust in binary RRT models.

III.2 Proposed Mixture Binary RRT Model

In Chapter II, we described two binary RRT models - the Warner's Indirect Question Model[62] and Greenberg's Unrelated Question model[16]. Of these two models, the unrelated question model is more efficient if the proportion of respondents asked the sensitive question, phrased directly, is greater than 1/3. Moreover, as few respondents under the Greenberg model are faced with the sensitive question altogether, it can put respondents more at ease. However, under the indirect question model, all respondents have to address the sensitive question irrespective of whether it is phrased directly or indirectly which puts more respondents in discomfort. Moreover, as stated earlier, according to Young et al.(2019)[67], when a researcher fails to account for an untruthful response in the Greenberg's Unrelated Question model[16] model, a bias is introduced into the estimator.

With this background, in Lovig et al.(2021)[41], we propose a model for asking a sensitive question with a binary response in a survey. This proposed Mixture binary RRT model (Figure III.2) uses the elements of both the Indirect Question model[62] and the Unrelated Question model[16]. The work was done both with and without accounting for untruthful responses.

Let π_X be the true prevalence of the sensitive trait in the population and π_Y be

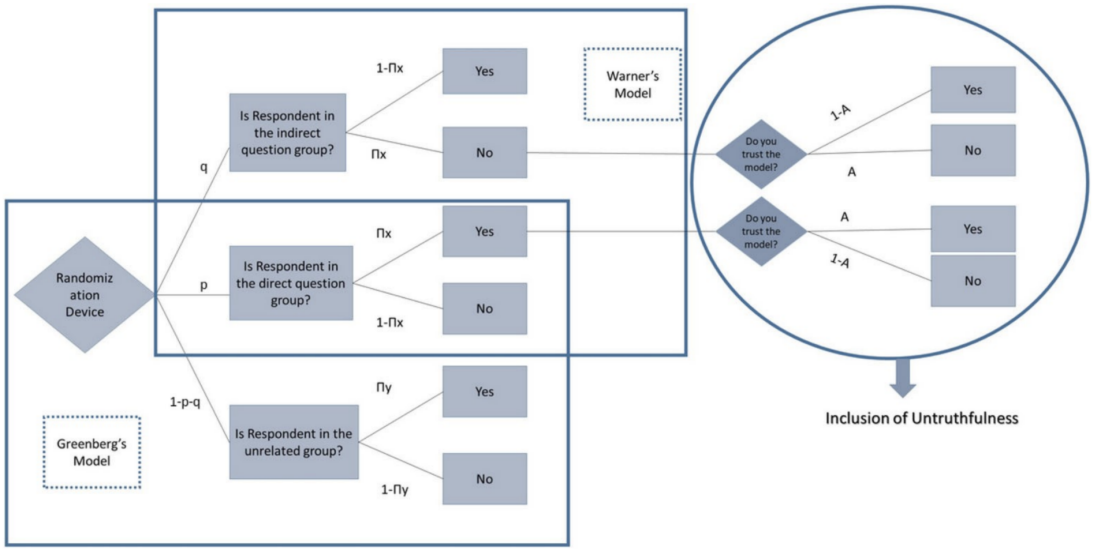

Figure III.2. Mixture RRT Model

the true prevalence of the non-sensitive unrelated trait in the population. Let p and q, respectively, denote the proportion of respondents asked the sensitive question phrased directly and the proportion of respondents asked the sensitive question phrased indirectly. Suppose A is the true proportion of survey participants who would respond to the sensitive question truthfully because they truth the model.

III.2.1 Mixture Binary RRT Model with Unaccounted Untruthfulness

The probability of a "Yes" response P_Y for the model is shown in Figure III.2 is given by

$$P_Y = \pi_X A(p-q) + q + (1-p-q)\pi_Y. \tag{III.5}$$

If a researcher erroneously does not account for the untruthfulness of the respon-

dents then they would end up using the estimator based on

$$P_Y^* = \pi_X(p-q) + q + (1-p-q)\pi_Y. \tag{III.6}$$

This wrong assumption can thus give us the incorrect estimator

$$\widehat{\pi_X}^* = \frac{\widehat{P_Y^*} - q - (1-p-q)\pi_Y}{p-q}, p \neq q. \tag{III.7}$$

Here $\widehat{P_Y*}$ is the proportion of "Yes" responses reported by the sample. This gives us the following expected value and bias for the estimator

$$\begin{aligned} E(\widehat{\pi_X}^*) &= \frac{E(P_Y^*) - q - (1-p-q)\pi_Y}{p-q} \\ &= \frac{E[\pi_X A(p-q) + q + (1-p-q)\pi_Y] - q - (1-p-q)\pi_Y}{p-q} \\ &= \pi_X A \end{aligned} \tag{III.8}$$

and

$$Bias(\widehat{\pi_X}^*) = E(\widehat{\pi_X}^*) - \pi_X = \pi_X(A-1). \tag{III.9}$$

Note, there is no bias when $A = 1$.

III.2.2 Efficiency under Mixture Binary RRT Model with Unaccounted Untruthfulness

For the erroneous estimator given by equation (III.7), the variance and the MSE are respectively given by

$$Var(\widehat{\pi_X}^*) = \frac{P_Y^*(1-P_Y^*)}{(n-1)(p-q)^2}, p \neq q \tag{III.10}$$

and

$$MSE(\widehat{\pi_X}^*) = Var(\widehat{\pi_X}^*) + [Bias(\widehat{\pi_X}^*)]^2. \tag{III.11}$$

III.2.3 Mixture Model Accounting for Untruthfulness

We noted in equation (III.5) for the mixture model that the correct probability of a "Yes" response is given by

$$P_Y = \pi_X A(p-q) + q + (1-p-q)\pi_Y. \tag{III.5}$$

This equation involves two unknown parameters i.e. A and π_X. Therefore, we use a two-question approach to estimate these parameters. The two questions used are as follows.

Question-1: Do you trust the model? (Using Greenberg Binary RRT Model)

Question-2: Do you have the sensitive trait? (Using Mixture Binary RRT Model)

Question-1: (With Greenberg Model) Do you trust the model?

The first question is used to estimate A using the traditional Greenberg et al.(1969)[16] model and the second question is used to estimate the sensitive trait prevalence. Let p_0 be the proportion of respondents who were asked Question-1 directly and π_{Y_0} be the proportion of respondents who were asked the unrelated question instead of Question-1. Then the proportion of respondents who gave a "Yes" response, i.e. P_{Y_0}, for this setup to estimate A is given by

$$P_{Y_0} = p_0 A + (1 - p_0)\pi_{Y_0}, \tag{III.12}$$

The unbiased estimator for A and its expected value are given by

$$\widehat{A} = \frac{\widehat{P_{Y_0}} - (1 - p_0)\pi_{Y_0}}{p_0}, \tag{III.13}$$

and

$$E(\hat{A}) = A; Var(\hat{A}) = \frac{P_{y0}(1 - P_{Y_0})}{np_0^2}. \tag{III.14}$$

Question-2: (With Mixture Model) Do you have the sensitive trait?

The probability of a "Yes" response to Question-2 is given by

$$P_Y = \pi_X A(p - q) + q + (1 - p - q)\pi_Y. \tag{III.15}$$

Then the estimator of the sensitive trait prevalence π_X is given by

$$\widehat{\pi_X} = \frac{\widehat{P_Y} - q - (1 - p - q)\pi_Y}{\hat{A}(p - q)}, p \neq q. \tag{III.16}$$

Note that

$$E(\widehat{P_Y}) = P_Y = \pi_X A(p - q) + q + (1 - p - q)\pi_Y;\ \ Var(\widehat{P_Y}) = \frac{P_Y(1 - P_Y)}{n}. \tag{III.17}$$

We use the first-order Taylor's approximation to rewrite the estimator in (III.16).

Using the result

$$f(x,y) \approx f(a,b) + (x-a)f_x(a,b) + (y-b)f_y(a,b).$$

and taking $x = \hat{P_Y}$, $y = \hat{A}$, $a = P_Y$, $b = A$, $\hat{\pi_X}$ can be approximated as

$$\widehat{\pi_X} \approx \frac{P_Y - q - (1-p-q)\pi_Y}{A(p-q)} - (\hat{A} - A)\left[\frac{P_Y - q - (1-p-q)\pi_Y}{A^2(p-q)}\right] + (\widehat{P_Y} - P_Y)\left[\frac{1}{A(p-q)}\right], p \neq q \quad \text{(III.18)}$$

The expected value of this estimator is given by

$$E(\widehat{\pi_X}) \approx \pi_X, \quad \text{(III.19)}$$

Note that the estimator in equation (III.18) is asymptotically unbiased as the expected values of the second and the third terms in equation (III.18) reduce to zero (using (III.14) and (III.17)).

III.2.4 Efficiency under Mixture Binary RRT Model with Accounted Untruthfulness

The efficiency of a model estimator, as established in Chapter II, is evaluated by the mean square error for the estimator. The MSE for an estimator $\hat{\theta}$ for the parameter θ is given by

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + [Bias(\hat{\theta})]^2$$

Therefore, when an estimator is unbiased, the variance alone can give us the measure of the efficiency of an estimator. Since the proposed estimator given in

equation (III.18) is approximately unbiased, we can compute the approximate variance of this estimator to measure its efficiency. The variance for this approximated $\widehat{\pi_X}$ from equation (III.18) is given by

$$Var(\widehat{\pi_X}) \approx \left[\frac{P_Y - q - (1-p-q)\pi_Y}{A^2(p-q)}\right]^2 Var(\hat{A}) + \left[\frac{1}{A(p-q)}\right]^2 Var(\widehat{P_Y}), p \neq q. \tag{III.20}$$

The values of $Var(\hat{A})$ and $Var(\widehat{P_Y})$ are given in equations (III.14) and (III.17) respectively.

III.2.5 Introduction to Privacy under Binary RRT Models

Although in regular surveys, a researcher may prioritize estimation efficiency, when the survey is on a sensitive topic, respondent privacy must also be ensured. If survey participants are not convinced about their responses staying private, it is possible they might refuse to participate, or worse, they might intentionally report untruthful responses. Therefore, evaluating the privacy level offered under an RRT model is equally important as evaluating the efficiency of the model estimator.

Lanke (1976)[38], proposed a measure for the privacy loss under a binary RRT model as described below.

Let

P(S|Y) be probability someone has the sensitive trait given they reported a "Yes" and

P(S|N) be probability someone has the sensitive trait given they reported a "No"

Then a measure of privacy loss δ is given by

$$\delta = Max(P(S|Y), P(S|N)). \tag{III.21}$$

One can transform this measure of privacy loss under a model such that it reflects the primary protection offered by the same model. This can be achieved by using a measure proposed by Fligner et al. (1977) [14] which is given by

$$PP = \frac{1-\delta}{1-\pi_X}. \tag{III.22}$$

III.2.6 Privacy under Mixture Binary RRT Model with Accounted Untruthfulness

Using the definition of this privacy loss measure given by equation (III.21) for the model shown in Figure III.2 is given by

$$\delta = Max(\eta_1, \eta_2) \text{ where,}$$

$$\eta_1 = P(S|Y) = \frac{pA\pi_X + (1-p-q)\pi_X\pi_Y + q\pi_X(1-A)}{\pi_X A(p-q) + q + (1-p-q)\pi_Y},$$

$$\eta_2 = P(S|N) = \frac{qA\pi_X + (1-p-q)\pi_X(1-\pi_Y) + p\pi_X(1-A)}{\pi_X A(q-p) + p + (1-p-q)(1-\pi_Y)}.$$

Theorem III.1. *The minimum value of δ is π_x.*

Proof. Let $p = q$, then

$$P(S|Y) = \frac{pA\pi_X + (1-p-q)\pi_X\pi_Y + q\pi_X(1-A)}{\pi_X A(p-q) + q + (1-p-q)\pi_Y} = \pi_x$$

$$P(S|N) = \frac{qA\pi_X + (1-p-q)\pi_X(1-\pi_Y) + p\pi_X(1-A)}{\pi_X A(q-p) + p + (1-p-q)(1-\pi_Y)} = \pi_x$$

Hence when $p = q$

$$\eta_1 = \eta_2 = Max(\eta_1, eta_2) = \delta = \pi_X.$$

Now assume $\delta < \pi_x$, which. implies $\exists p$ and $\exists q$ such that

$$P(S|Y) < \pi_x \ \ and \ \ P(S|N) < \pi_x$$

$$\begin{aligned}
&\Rightarrow \frac{P(S \cap Y)}{P(Y)} < \pi_x; \ \ \frac{P(S \cap N)}{P(N)} < \pi_x \\
&\Rightarrow P(S \cap Y) < \pi_x P(Y); \ \ P(S \cap N) < \pi_x P(N) \\
&\Rightarrow P(S \cap Y) + P(S \cap N) < \pi_x P(Y) + \pi_x P(N) \\
&\Rightarrow P(S) < \pi_x
\end{aligned}$$

This creates a contradiction to the assumption, since $P(S) = \pi_x$. Hence $\delta \geq \pi_x$ and the minimum privacy loss under this model is attained when $p = q$. This result implies that the closer p and q are to each other, the more privacy the proposed mixture model would provide to the survey participants. □

It must be noted that since $\delta \geq \pi_X$, the proportion of the maximum primary protection that can be achieved by a model is represented by the primary protection

measure PP (equation III.22).

III.2.7 Unified Measure of Privacy and Efficiency

When one wants to evaluate the overall performance of an RRT model, both the model efficiency and the privacy protection offered by the model. Often the model efficiency, evaluated by the mean squared error of the estimator, worsens when the privacy level is increased for that model. This can complicate the assessment of the overall model performance. Gupta et al.(2018)[21] proposed a combined measure for quantitative RRT models which is given by

$$M = \frac{MSE(\hat{\theta})}{PL}, \tag{III.23}$$

where PL is the privacy level of the model with the estimator $\hat{\theta}$ for estimating the population parameter θ. The higher the value of PL, the more privacy is offered to the survey participants under the model. Moreover, a researcher would also want to minimise the $MSE(\hat{\theta})$.

We modified the combined measure of estimator quality in RRT models proposed by Gupta et al. (2018) [21] and propose a new unified measure of privacy and efficiency for binary RRT models. If π_x is the true sensitive trait prevalence and δ is the measure of privacy loss, then the proportion of the maximum primary protection that can be achieved by the mixture model is given by

$$PP = \frac{1-\delta}{1-\pi_x}. \tag{III.24}$$

The proposed unified measure of privacy and efficiency is then given by

$$M = \frac{PP^a}{MSE^b}, \tag{III.25}$$

where a and b are weights chosen by the researcher to account for the importance of privacy and efficiency respectively. Note that a higher value of M is preferred since that would correspond to a higher level of Privacy Protection, or a lower MSE, or both.

III.3 Simulation Study

In this section, we examine the performance of the proposed Mixture Binary RRT model and compare it to the performance of the Warner's Indirect Question model[62] and the Greenberg's Unrelated Question model[16]. For this purpose, we evaluate the performance of these three models in terms of MSE, the PP and the proposed combined measure M. The combined measure M has been computed for both the scenario where we account for the level of trust A (Scenario-2) and the scenario where we do not account for it (Scenario-1).

We ran a simulation study with 10000 iterations with samples of size $n = 500$ in each iteration. We assume the true sensitive trait prevalence in the population to be $\pi_X = 0.4$. Greenberg et al. (1969)[16] recommended that π_Y should be kept as small as cooperation would allow if $\pi_X < 0.5$. Therefore, the true prevalence of the non-sensitive unrelated trait in the population is assumed to be $\pi_Y = 0.1$ for this simulation study. We evaluate the model performance at varying levels of respondent trust A in the model. $A = 1$ corresponds to the case where all survey participants trust the model and hence all respondents respond truthfully in the RRT survey using the proposed mixture model. Lower values of A correspond to a lower proportion of

respondents trusting the proposed model in the survey.

These simulation results have been summarized in Table III.2. Subscript 1 refers to the cases when untruthfulness is accounted for and subscript 2 refers to the cases when untruthfulness is unaccounted for. In the cases where the proportion of individuals who are asked the indirect question $q = 0$, the proposed model gets reduced to Greenberg's unrelated question model[16]. Similarly, in the cases where the proportion of individuals who are asked the indirect question $q = 1 - p$, the proposed model gets reduced to Warner's indirect question model[62].

Our simulation results for the mixture binary RRT model show that the Greenberg model performs the best in terms of efficiency (i.e. MSE) and the Warner's model performs the best in terms of privacy. However, the mixture model outperforms both models when the values for the proposed unified measure of privacy and efficiency shown in (III.25) are compared. This unified measure allows the researcher to adjust for the weights for the privacy protection level and the MSE simultaneously. Overall, through our work, we established that the mixture model in fact outperforms both the Greenberg's Unrelated Question model and the Warner's Indirect Question model when privacy and efficiency are simultaneously factored in with equal weights.

In Table III.2, we can clearly note the impact of lack of truth and untruthful responses under the model by comparing the columns for $\hat{\pi}_{x1}$ and $\hat{\pi}_{x2}$. We observe that as the level of truth A in the model drops, and more respondents give untruthful responses when one does not account for A, the estimates show considerable differences. For instance, when $A = 1$ with $p = 0.4$ and $q = 0.05$, the two estimates are fairly similar in value. However, the two estimates vary considerably as A drops. The impact of untruthful responses on the efficiency can be noted by looking at the MSE_1 and MSE_2 columns. It can be noted that the MSE values increase as the A value drops

Table III.2. Theoretical (bold) and empirical values based on 10000 iterations $n = 500$, $\pi_X = 0.4$, $\pi_Y = 0.1$, $\pi_{y0} = 0.1$, $p_0 = 0.7$

p	q	$1-p-q$	A	$\hat{\pi}_{x1}$	MSE_1	$\hat{\pi}_{x2}$	MSE_2	PP	M_1	M_2
0.4	0	0.6	1	0.4003	0.0023	0.4001	0.0021	0.2735	119.6437	127.8524
					0.0023		**0.0021**	**0.2727**	**119.7078**	**126.8912**
0.4	0	0.6	0.9	0.4005	0.0027	0.3601	0.0036	0.2938	110.1171	80.5532
					0.0027		**0.0037**	**0.2941**	**109.2467**	**78.446**
0.4	0	0.6	0.8	0.4003	0.0031	0.3198	0.0084	0.3195	101.9188	38.2897
	Greenberg's Model				**0.0032**		**0.0085**	**0.3191**	**98.659**	**37.3304**
0.4	0.05	0.5	1	0.4007	0.0031	0.4002	0.0028	0.4291	137.2452	144.4069
					0.0032		**0.003**	**0.4285**	**135.8376**	**141.6272**
0.4	0.05	0.5	0.9	0.4005	0.0037	0.36	0.0045	0.4536	121.7742	101.0926
					0.0038		**0.0046**	**0.4545**	**120.6014**	**98.2577**
0.4	0.05	0.5	0.8	0.4007	0.0046	0.3202	0.0092	0.4837	108.4597	52.7426
	Mixture Model				**0.0046**		**0.0094**	**0.4839**	**105.4075**	**51.3333**
0.4	0.6	0	1	0.3982	0.0124	0.3977	0.0122	0.8341	67.4374	68.4523
					0.0126		**0.0125**	**0.8333**	**66.9597**	**66.64**
0.4	0.6	0	0.9	0.4006	0.0157	0.3602	0.0141	0.8471	54.0927	59.9959
					0.0156		**0.0141**	**0.8474**	**54.3387**	**60.082**
0.4	0.6	0	0.8	0.4023	0.0197	0.3215	0.0186	0.8621	43.8717	46.2804
	Warner's Model				**0.0197**		**0.0189**	**0.8621**	**43.7237**	**45.6**
0.55	0	0.45	1	0.4001	0.0014	0.3998	0.0013	0.1698	117.6022	130.5046
					0.0014		**0.0013**	**0.1698**	**119.6422**	**131.601**
0.55	0	0.45	0.9	0.4009	0.0017	0.3604	0.0028	0.1862	111.2111	65.4519
					0.0017		**0.0029**	**0.1852**	**109.8513**	**64.0702**
0.55	0	0.45	0.8	0.3998	0.0019	0.3195	0.0076	0.2046	110.5253	26.9
	Greenberg's Model				**0.002**		**0.0077**	**0.2036**	**100.3351**	**26.4773**
0.55	0.1	0.35	1	0.3995	0.0023	0.3993	0.0021	0.4285	189.8	201.6383
					0.0023		**0.0021**	**0.4286**	**189.2684**	**200.7001**
0.55	0.1	0.35	0.9	0.3997	0.0027	0.3594	0.0037	0.4557	170.4568	122.8569
					0.0027		**0.0037**	**0.4545**	**166.3642**	**121.6865**
0.55	0.1	0.35	0.8	0.4014	0.0032	0.3206	0.0083	0.4845	153.7125	58.7283
	Mixture Model				**0.0034**		**0.0085**	**0.4839**	**144.1011**	**56.69**
0.55	0.45	0	1	0.3975	0.0509	0.3971	0.0507	0.9173	18	18.0905
					0.0502		**0.0501**	**0.9184**	**18.2908**	**18.3379**
0.55	0.45	0	0.9	0.3988	0.0628	0.3587	0.0524	0.9252	14.73	17.6645
					0.062		**0.0517**	**0.9259**	**14.9379**	**17.9165**
0.55	0.45	0	0.8	0.3937	0.0788	0.3145	0.0574	0.9322	11.8286	16.2399
	Warner's Model				**0.0784**		**0.0565**	**0.9336**	**11.9041**	**16.5291**
0.7	0	0.3	1	0.4001	0.001	0.3996	0.0008	0.0968	99.1116	114.5333
					0.001		**0.0009**	**0.0968**	**96.4091**	**110.623**
0.7	0	0.3	0.9	0.3999	0.0011	0.3597	0.0025	0.1062	94.2058	43.204
					0.0012		**0.0025**	**0.1064**	**88.3861**	**42.9863**
0.7	0	0.3	0.8	0.4003	0.0013	0.3198	0.0072	0.1178	88.582	16.3157
	Greenberg's Model				**0.0015**		**0.0073**	**0.1181**	**80.9974**	**16.2355**
0.7	0.15	0.15	1	0.4002	0.0017	0.3998	0.0016	0.4291	251.218	272.8059
					0.0016		**0.0016**	**0.4286**	**252.462**	**273.2205**
0.7	0.15	0.15	0.9	0.4006	0.002	0.36	0.0031	0.4551	225.2544	144.796
					0.0021		**0.0032**	**0.4545**	**219.325**	**143.4535**
0.7	0.15	0.15	0.8	0.4006	0.0025	0.3202	0.0079	0.4842	194.8597	61.2223
	Mixture Model				**0.0026**		**0.008**	**0.4839**	**188.0234**	**60.7222**
0.7	0.3	0	1	0.4009	0.0033	0.4003	0.0031	0.6521	199.4005	211.1705
					0.0032		**0.0031**	**0.6522**	**201.276**	**209.62**
0.7	0.3	0	0.9	0.4008	0.0039	0.3604	0.0046	0.6761	174.6029	147.0005
					0.004		**0.0047**	**0.6757**	**168.9802**	**143.4183**
0.7	0.3	0	0.8	0.4008	0.0049	0.3201	0.0094	0.6998	143.7185	74.3188
	Warner's Model				**0.005**		**0.0095**	**0.7009**	**139.1164**	**37.6955**

and they become worse if the level of truthfulness A is not accounted for. We also observe the resultant impact of the untruthful responses on the proposed combined measure M. When respondents are providing untruthful responses, i.e. $A < 1$, we can note that $M_1 > M_2$. Based on the definition of the proposed unified measure as shown in equation III.25, we can see that a higher value of M indicates an overall better model performance. Note that for this study we assume the weight for this unified measure to be $a = 1$ and $b = 1$. Therefore, based on our results, we can infer that when respondents are providing untruthful responses due to a lack of trust in the mixture model, the overall model performance can be considerably improved by accounting for and estimating A. The bold entries in Table III.2 are the theoretical values of the measures while the non-bold figures are the empirical values of the various measures summarized in this table. Comparing the theoretical values with the corresponding empirical values for the MSEs, PPs and the unified measures M_1 and M_2, we can note that our theoretical results match with the corresponding empirical results reasonably well.

In particular, it can be noted that the Greenberg model ($q = 0$) performs the best in terms of efficiency, i.e. has the lowest MSE. The Warner's model ($q = 1 - p$) performs the best in terms of Privacy Protection. However, when we consider the efficiency and privacy protection, simultaneously with equal weights ($a = 1, b = 1$), the proposed Mixture Binary RRT model performs the best, i.e. the proposed model has the highest value for M.

For instance, consider the case with $A = 0.8$ (i.e. 20% untruthful responses) corresponding to $p = 0.7$ and $q = 0$ (i.e. Greenberg's Model). Here, the theoretical $MSE_1 = 0.0015$, theoretical $PP = 0.1181$ and the theoretical $M_1 = 80.9974$. Similarly, when $A = 0.8$, consider the case where $p = 0.7$ and $q = 0.3$ (i.e. Warner's Model).

Here, the theoretical $MSE_1 = 0.0050$, theoretical $PP = 0.7009$ and the theoretical $M_1 = 139.1164$. Lastly, when $A = 0.8$, consider the case when $p = 0.7$ and $q = 0.15$. Here, the theoretical $MSE_1 = 0.0026$, theoretical $PP = 0.4839$ and the theoretical $M_1 = 188.0234$. When we compare these three cases, we observe that clearly the Greenberg's Model (i.e. $p = 0.7$ and $q = 0$) has the least MSE. We also note that the Warner's Model (i.e. $p = 0.7$ and $q = 0.3$) has the highest primary protection. However, when we compare the unified measure of privacy and efficiency (i.e. M_1) for the three models, we note that the mixture model outperforms the other two models (188.0234 vs. 80.9974, 139.1164). We also note that for the stated cases, theoretical M_2 is always lower than the theoretical M_1.

III.3.1 Impact of untruthfulness on the unified measure

We can note from the results summarized in Table III.2 that not estimating A always corresponds to a lower and less favorable value of the unified measure M. For instance, the maximum value of M observed when A is not being estimated is obtained when $p = 0, q = 0.55$. This maximum value of $M = 144.3773$ and corresponds to $MSE = 0.0052$ and primary protection $PP = 0.7542$. However, when A is estimated, the maximum value of M is observed when $p = 0, q = 0.80$. This maximum value of $M = 423.3382$ and corresponds to $MSE = 0.0011$ and primary protection $PP = 0.4838$. Hence we can note that not estimating A can give us a higher bias and hence a higher MSE which results in a lower value for M. Based on these results it is very clear that not estimating A is always the less efficient method when even a small proportion of respondents are providing untruthful responses in the survey.

When we plot the values of the unified measure M along with corresponding values

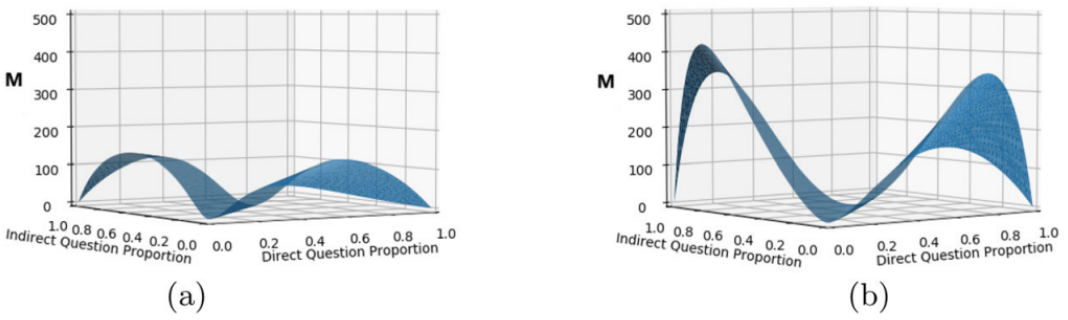

Figure III.3. Unified Measure (M) for different choices of p and q $\pi_X = 0.3$,$\pi_Y = 0.1$, $n = 500$, $A = 0.8$, $a = 1$, $b = 1$ (a) Untruthfulness is not accounted for. (b) Untruthfulness is accounted for.

of p and q (Figure III.3), we note that the most favorable values of M are obtained when p or q get closer to 1, but are not too close to 1, and $p + q \neq 1$. In particular, we can note that when we estimate A, we obtain a higher value for the unified measure M near the tails (i.e. near $p = 1$ or $q = 1$) of the graph (FigureIII.3) (b)). However, when we do not estimate A, the optimal value of the unified measure M is farther away from the tails compared to the case where we estimate A. Thus, it would be fair to infer that respondent cooperation becomes less of an issue when we account for untruthfulness by estimating A. We must also note that the value of μ_Y also impacts the values of M. Therefore we follow the guideline from the work by Greenberg et al. (1969)[16] as stated earlier in this section.

III.4 Concluding Chapter Remarks

In this chapter, the mixture binary RRT model was introduced along with the work done by Young et al. (2019)[67]. The main contribution of this chapter is that we first verify the work done by Young et al. (2019)[67] through a simulation study and establish that most traditional RRT models make the assumption that under an RRT

survey, the respondents have no longer have a reason to lie and since they completely trust the model, there are no untruthful responses. However, it is possible that some respondents may not be convinced and due to concerns for their privacy might still lie. Through the simulation results, we demonstrate the impact of untruthful responses under the traditional Greenberg et al. (1969) model.

Following this idea, and combining elements of Warner's indirect question model[62] and Greenberg's unrelated question model[16] to reap the benefits of both, we introduce the proposed mixture binary RRT model. We assess this proposed model in two scenarios- when the untruthful responses are accounted for and when the untruthful responses are not accounted for. Through both theoretical and empirical results, that match reasonably well, we establish that the proposed mixture binary RRT model has the best overall performance when both efficiency and privacy are factored in with equal importance. Furthermore, the choice from three questions in the mixture model helps improve the respondent participation as compared to when they have a choice of two questions which is offered by the Warner's indirect question model[62] and Greenberg's unrelated question model[16]. In particular, we also infer that research utilizing the proposed mixture binary RRT model must choose a value of p such that it results in a high level of cooperation, and still gives favorable value (i.e. a high value) of the combined measure M.

Chapter IV: Optional Mixture Binary RRT Model with a Unified Measure of Privacy and Efficiency

Gupta et al. (2002)[19] proposed an alternative approach to RRT models known as Optional RRT models. Such models allow respondents to report their unscrambled responses if they do not find the question sensitive, and a scrambled response otherwise. Various researchers have shown that optional models help improve the efficiency of their corresponding non-optional counterparts. Some significant work on Optional RRT models has been done by various researchers including Gupta et al. (2010)[22], Kalucha et al. (2016)[30], Mehta and Aggarwal (2018)[42], Gupta et al. (2018b)[21], Narjis and Shabbir (2020)[46], Narjis and Shabbir (2021)[45] Khalil et al. (2021)[33] and Zhang et al. (2021)[69].

Gupta et al. (2002)[19] showed that forcing all sample participants to provide a scrambled response, irrespective of whether they find the question sensitive or not, adds unnecessary noise to the data collected from the sample surveyed. They proposed an Optional RRT model where respondents who do not find the question sensitive are asked to report their true response to the sensitive question while other respondents who find the question sensitive are asked to follow the scrambling rules of the model. Through their work, they show that giving respondents the option to report their true response, unaltered, if they do not find the question sensitive helps improve model

efficiency. Gupta et al. (2018)[21] established that the inclusion of the optionality element in a model does not impact the privacy offered by the model as only those people who do not have their responses protected and who do not find the question sensitive, to begin with, and report an unscrambled response to the surveyor.

It is well established by Lovig et al. (2021)[41] that lack of trust in an RRT model can be accounted for and that the mixture model successfully combines the benefits of both the Warner (1965)[62] Indirect Question Model and the Greenberg et al. (1969)[16] Unrelated Question model. Moreover, it has also been established in the literature that integration of the optionality component helps improve the model efficiency. In this paper, our main goal is to verify if the integration of optionality helps improve the efficiency of the model proposed by Lovig et al. (2021)[41]. With that in mind, in Chapter IV[1], we present the Sapra et. al. (2022)[54] model where we proposed an Optional Mixture binary RRT model to help mitigate the effect of respondents' lack of trust in binary RRT models while acknowledging that sensitivity to a survey question is subjective.

IV.1 Proposed Optional Mixture Binary RRT Model

In this section, we propose an Optional Mixture Binary RRT Model (Figure IV.1) that integrates the idea of optionality with a mixture of the Warner's Indirect Question Model[62] and the Unrelated Question model[16] to help combine their respective strengths. Incorporating optionality helps us acknowledge that not all survey respondents might find the survey question sensitive[19]. In this model, we also account for

[1]This work was first published in Journal of Statistical Theory and Practice, Volume 16, number 3, pages 51, 14 July 2022 by Springer Nature. The original article is available online at: https://doi.org/10.1007/s42519-022-00279-3 and this chapter has a version of this work.

the lack of trust as suggested by Young et al. (2019)[67] and Lovig et al. (2021)[41].

In Sapra et al.(2022)[54], we integrate the aspect of optionality into the model shown in Figure III.2, to obtain an Optional Mixture Binary RRT model which is shown in Figure IV.1. We have also accounted for the lack of trust in the model.

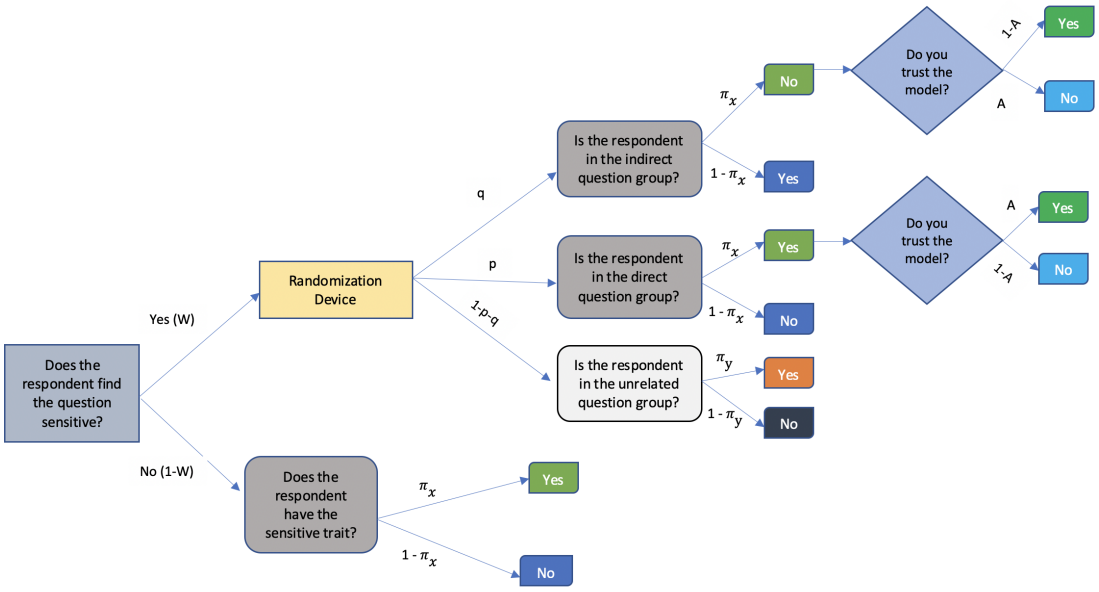

Figure IV.1. Optional Mixture Binary RRT Model

Under this Optional Mixture Binary RRT model, the respondents who do not find the question sensitive can directly respond to the sensitive question by providing unaltered true responses. However, if they find the question sensitive, the randomization device would help them scramble their response using the mixture model. Here, the true proportion of the individuals in the population that find the question sensitive is called the sensitivity level of the question and is denoted by W. Further, p is the proportion of sample respondents that are asked the sensitive question phrased directly and q is the proportion of sample respondents that are asked the sensitive question phrased indirectly. The total size of the sample drawn, using a simple random sample

without replacement, is denoted by n. The proportion of individuals who trust the model and hence provide truthful responses is denoted by A.

Let π_X be the true sensitive trait prevalence in the population and π_Y be the true non-sensitive unrelated trait prevalence in the population. If for the model shown in Figure IV.1, A denotes the true level of trust in the model, then the probability of a "Yes" response is given by

$$P_{oy} = P(Yes) = \pi_x[(1 - W) + WA(p - q)] + W[q + (1 - p - q)\pi_y]. \tag{IV.1}$$

However if one erroneously assumes that all respondents trust the model, i.e. $A = 1$, when in fact they do not (i.e. $A < 1$), one would use the probability of a "Yes" response given by

$$P^*_{oy} = \pi_x[(1 - W) + W(p - q)] + W[q + (1 - p - q)\pi_y], \tag{IV.2}$$

and we would get the wrong estimator is given by

$$\hat{\pi}^*_x = \frac{\hat{P}^*_{oy} - \hat{W}[q + (1 - p - q)\pi_y]}{[\hat{W}(p - q) + (1 - \hat{W})]}, \tag{IV.3}$$

where $\hat{P}^*_{oy}$ is the proportion of "Yes" responses in the sample survey.

Using Taylor's approximation, we can re-write the incorrect estimator under the wrong assumption of $A = 1$, as follows.

$$\widehat{\pi}_x^* \approx \frac{\pi_x[WA(p-q)+(1-W)]}{[W(p-q)+(1-W)]} + (\widehat{P}_{oy}^* - P_{oy}^*)\frac{1}{[W(p-q)+(1-W)]} - (\widehat{W}-W)\frac{[q+(1-p-q)\pi_y + P_{oy}^*(p-q-1)]}{[W(p-q-1)+1]^2}. \quad \text{(IV.4)}$$

Then the approximate expected value of this estimator is given by

$$E(\hat{\pi}_x^*) \approx \frac{\pi_x[WA(p-q)+(1-W)]}{[W(p-q)+(1-W)]} = \frac{\pi_x(A\alpha+\beta)}{(\alpha+\beta)} < \pi_x(since\ A<1), \quad \text{(IV.5)}$$

where $\alpha = W(p-q)$ and $\beta = (1-W)$.

Then the approximate bias for the wrong estimator from equation (IV.3) is given by

$$Bias(\hat{\pi}_x^*) = E(\hat{\pi}_x^*) - \pi_x \approx \frac{\pi_x\alpha(A-1)}{(\alpha+\beta)} < 0 \ \ (since\ A<1) \quad \text{(IV.6)}$$

Hence if one wrongly assumes that there are no untruthful responses when respondents do not trust the model, the estimates would have a considerable negative bias.

In Table (IV.1), for fixed W and other model parameters, we can see that as the level of trust (A) in the model declines, considerable bias is introduced in the estimates computed using the incorrect estimator from equation (IV.3). This shows that if one wrongly assumes no respondent is lying (i.e. $A = 1$) even when some respondents are lying, we get poor estimates of the sensitive trait prevalence. Moreover, we note that for fixed p and q the impact of A on the estimates is less adverse for lower W which

Table IV.1. Simulation Results: Estimator performance when one wrongly assumes $A = 1$ ($N = 10000, n = 500, \pi_x = 0.4$)

p	q	$1-p-q$	W	A	$\widehat{\pi_x}^*$
0.55	0.1	0.35	0.8	1	0.399971
0.55	0.1	0.35	0.8	0.9	0.374264
0.55	0.1	0.35	0.8	0.8	0.348607
0.55	0.1	0.35	0.5	1	0.400160
0.55	0.1	0.35	0.5	0.9	0.387602
0.55	0.1	0.35	0.5	0.8	0.375218
0.55	0.1	0.35	0.3	1	0.399956
0.55	0.1	0.35	0.3	0.9	0.393404
0.55	0.1	0.35	0.3	0.8	0.386981
0.7	0.15	0.15	0.8	1	0.400318
0.7	0.15	0.15	0.8	0.9	0.372558
0.7	0.15	0.15	0.8	0.8	0.345072
0.7	0.15	0.15	0.5	1	0.400083
0.7	0.15	0.15	0.5	0.9	0.385786
0.7	0.15	0.15	0.5	0.8	0.371593
0.7	0.15	0.15	0.3	1	0.399777
0.7	0.15	0.15	0.3	0.9	0.392056
0.7	0.15	0.15	0.3	0.8	0.384447

shows that the inclusion of optionality is helpful in mitigating the lack of trust in the model even when one wrongly assumes $A = 1$.

For instance, when $p = 0.55$ and $q = 0.1$, let us compare the worsening of the negative bias of the value of the estimates $\widehat{\pi_x}^*$ when compared to the true $\pi_x = 0.4$ for different values of W. When $W = 0.3$, at $A = 1, 0.9$ and 0.8, the corresponding values of $\widehat{\pi_x}^*$ are 0.3999, 0.3934 and 0.3870. Although the bias worsens as A drops for $W = 0.3$, it does not have a huge impact on the value of the estimates. When $W = 0.5$, at $A = 1, 0.9$ and 0.8, the corresponding values of $\widehat{\pi_x}^*$ are 0.4002, 0.3876 and 0.3486. Again the bias worsens as A drops for $W = 0.5$, and has a moderately more impact on the value of the estimates compared to what we observed for $W = 0.3$. When $W = 0.8$, at $A = 1, 0.9$ and 0.8, the corresponding values of $\widehat{\pi_x}^*$ are 0.3999, 0.3743 and 0.3486. For $W = 0.8$ the impact of this bias is a lot more than what we

saw for $W = 0.3$ and $W = 0.5$. It is very clear that the impact of the negative bias shown theoretically in equation (IV.6), worsens as W increases. Therefore, the worst possible bias would be observed at the maximum value of W i.e. at $W = 1$. Hence, we can infer that the inclusion of optionality alone helps mitigate the bias introduced by untruthful responses to some extent.

IV.1.1 Accounting for Lack of Trust in Optional Mixture RRT Model

We have established that not accounting for lack of trust can introduce a negative bias into our estimates. Therefore we account for the lack of trust using an estimate of A and an estimate of the sensitivity level of the survey question W.

Suppose one correctly accounts for lack of trust, then using IV.1 leads to

$$\hat{\pi}_x = \frac{\hat{P}_{oy} - \hat{W}[q + (1 - p - q)\pi_y]}{[\hat{W}\hat{A}(p - q) + (1 - \hat{W})]}, \tag{IV.7}$$

where $\hat{P}_{oy}$ is the proportion of "Yes" responses in the survey.

We can see that the estimator given in equation IV.7 relies on three unknown parameters i.e. π_x, A and W. We propose to estimate A and W separately using a Greenberg et al. (1969)[16] model.

These estimates, $\hat{A}$ and $\hat{W}$ can then further be used to estimate the sensitive trait prevalence π_x using the estimator given in equation IV.7. Using Taylor's approximation we can re-write the correct estimator from equation (IV.7) as follows.

$$\widehat{\pi}_x \approx \frac{\lambda}{[A\alpha+\beta]} + (\widehat{P}_{oy} - P_{oy})\frac{1}{[A\alpha+\beta]} - (\widehat{W} - W)\frac{[A\alpha+\beta]\gamma + \lambda[A(p-q)-1]}{[A\alpha+\beta]^2} - (\hat{A} - A)\frac{\alpha\lambda}{[A\alpha+\beta]^2}, \quad \text{(IV.8)}$$

where $\alpha = W(p-q)$, $\beta = (1-W)$, $\gamma = q + (1-p-q)\pi_y$ and $\lambda = P_{oy} - W\gamma$.

Then the approximate expected value of the estimator is given by

$$E(\hat{\pi}_x) \approx \frac{P_{oy} - W[q + (1-p-q)\pi_y]}{[WA(p-q) + (1-W)]} = \pi_x, \quad \text{(IV.9)}$$

Since

$$E(\hat{A}) = A, \quad \text{(IV.10)}$$

and

$$E(\hat{W}) = W. \quad \text{(IV.11)}$$

Therefore, when we account for the lack of trust in the model, the estimator for the proposed Optional Mixture RRT Model is asymptotically unbiased.

IV.1.2 Efficiency of Optional Mixture Binary RRT Model

The approximate variance for the estimator in equation (IV.7) is given by

$$Var(\widehat{\pi}_x) \approx Var(\widehat{P}_{oy})\left[\frac{1}{(A\alpha+\beta)^2}\right] + Var(\widehat{W})\left[\frac{[A\alpha+\beta]\gamma + [P_{oy}-\gamma W][A(p-q)-1]}{(A\alpha+\beta)^2}\right]^2 + Var(\widehat{A})\left[\frac{\alpha^2\lambda^2}{(A\alpha+\beta)^4}\right]. \quad \text{(IV.12)}$$

$$MSE(\hat{\pi}_x) \approx Var(\hat{\pi}_x), \quad \text{(IV.13)}$$

where, $Var(\hat{P}_{oy})$ is given by

$$Var(\hat{P}_{oy}) = \frac{P_{oy}(1-P_{oy})}{n}. \quad \text{(IV.14)}$$

$Var(\hat{A})$ and $Var(\hat{W})$ will be the variance of the Greenberg et al. (1969)[16] estimators for A and W which are given by

$$Var(\hat{A}) = \frac{P_{Ya}(1-P_{Ya})}{np_a^2}, \quad \text{(IV.15)}$$

and

$$Var(\hat{W}) = \frac{P_{Yw}(1-P_{Yw})}{np_w^2}. \quad \text{(IV.16)}$$

Here P_{YA} and P_{YW} are the probability of a "Yes" response for the two Greenberg models used to estimate A and W respectively. Further, p_A and p_W denote the proportions of individuals asked the direct question under the two Greenberg models estimating A and W respectively.

IV.1.3 Privacy of the Optional Mixture RRT Model

Using definition proposed by Lanke (1976)[38], privacy loss is given by

$$\delta = Max(\eta_1, \eta_2), \tag{IV.17}$$

where η_1 and η_2 are given by

$$\eta_1 = P(S|Y) = \frac{W[q\pi_x(1-A) + p\pi_x A + (1-p-q)\pi_y\pi_x] + (1-W)\pi_x}{P_{oy}}, \tag{IV.18}$$

$$\eta_2 = P(S|N) = \frac{W[q\pi_x A + p\pi_x(1-A) + (1-p-q)(1-\pi_y)\pi_x]}{1-P_{oy}}. \tag{IV.19}$$

Lovig et al. (2021)[41] showed that $\delta \geq \pi_x$ and Gupta et al. (2018)[21] showed that optionality does not affect privacy. As optionality should not affect the value of the privacy loss measure δ, it is fair to infer that $\delta \geq \pi_x$ for the proposed Optional Mixture RRT Model as well.

Primary protection offered by the proposed model is given by

$$PP = \frac{1-\delta}{1-\pi_x}. \tag{IV.20}$$

Here δ is the measure for privacy loss [38]. The combined measure of privacy and efficiency as proposed by Lovig et al. (2021)[41] is given by

$$M = \frac{PP^a}{MSE^b}. \tag{IV.21}$$

A model with better overall performance would have a higher value of M as it indicates a higher privacy level, a lower MSE, or both.

IV.2 Simulation Study on the Optional Mixture Binary RRT Model

We compare our theoretical results with empirical results generated by conducting a simulation study with $N = 10000$ iterations each with a simple random sample of size $n = 500$ drawn with replacement. The true sensitive trait prevalence π_x is assumed to be 0.4. The unrelated trait prevalence π_y for the Optional Mixture Model is assumed to be 0.1. The parameters for estimating A are $\pi_{y0a} = 0.15$ and $p_{0a} = 0.75$ and the parameters for estimating W are $\pi_{y0w} = 0.1$ and $p_{0w} = 0.7$.

Both A and W have been estimated using samples of size 500 for each simulation. However, one could use a sample size different from what we use to estimate sensitive trait prevalence. In order to ensure the independence of A and W from π_X, we generate separate independent samples to simulate the use of pre-surveys for this purpose.

The complete simulation results for the Optional Mixture RRT Model have been given in Table (IV.2). The columns with subscript T are the theoretical values of the listed measures for specified values of the model parameters. These have been computed using the formulas introduced in the previous section. The columns with subscript E are the corresponding empirical values.

In Table (IV.2), the cases highlighted in yellow are for the optional Greenberg's [16] Model cases ($q = 0$) and the cases highlighted in green are for the optional Warner's [62] Model cases ($1 - p - q = 0$). The cases that have not been highlighted are for the

proposed Optional Mixture RRT Model. It can be seen that all the theoretical values and their corresponding empirical values match reasonably well. Although we show results only for $p = 0.55$ and $p = 0.7$ with various values of q, we see similar results for other choices of p and q as well.

Table IV.2. Simulation Results Mixture ORRT: N=1000, n=500, $\pi_x = 0.4$, $\pi_y = 0.1$, $\pi_{y0w} = 0.1$, $p_{0w} = 0.7$, $\pi_{y0a} = 0.15$, $p_{0a} = 0.75$

p	q	$1-p-q$	W	A	$\hat{W}_E$	$\hat{A}_E$	$\hat{\pi}_{XE}$	$MSE(\hat{\pi}_X)_E$	$MSE(\hat{\pi}_X)_T$	PP_E	PP_T	M_E	M_T
0.55	0	0.45	1	1	1.0003	1.0001	0.3984	0.0014	0.0015	0.1727	0.1698	126.04	115.92
0.55	0	0.45	1	0.9	1.0003	0.8999	0.4038	0.0016	0.0018	0.1884	0.1852	120.80	105.02
0.55	0	0.45	1	0.8	1.0003	0.8001	0.4049	0.0019	0.0022	0.2078	0.2036	110.75	94.33
0.55	0	0.45	0.9	1	0.9002	1.0010	0.3966	0.0012	0.0013	0.1711	0.1698	147.23	133.56
0.55	0	0.45	0.9	0.9	0.9002	0.9007	0.4002	0.0013	0.0015	0.1862	0.1852	146.29	124.66
0.55	0	0.45	0.9	0.8	0.9002	0.8006	0.4002	0.0015	0.0018	0.2049	0.2036	140.04	116.04
0.55	0	0.45	0.8	1	0.8000	1.0003	0.3972	0.0011	0.0011	0.1706	0.1698	161.93	152.40
0.55	0	0.45	0.8	0.9	0.8000	0.9005	0.4002	0.0011	0.0013	0.1854	0.1852	163.09	145.88
0.55	0	0.45	0.8	0.8	0.8000	0.8003	0.4003	0.0013	0.0015	0.2039	0.2036	159.03	139.76
0.55	0.1	0.35	1	1	1.0003	1.0001	0.3974	0.0022	0.0023	0.4329	0.4286	199.81	187.27
0.55	0.1	0.35	1	0.9	1.0003	0.8999	0.4028	0.0025	0.0028	0.4607	0.4545	181.69	163.58
0.55	0.1	0.35	1	0.8	1.0003	0.8001	0.4039	0.0031	0.0034	0.4922	0.4839	156.71	140.58
0.55	0.1	0.35	0.9	1	0.9002	1.0010	0.3970	0.0017	0.0018	0.4313	0.4286	251.63	234.83
0.55	0.1	0.35	0.9	0.9	0.9002	0.9007	0.4003	0.0019	0.0021	0.4577	0.4545	237.49	213.44
0.55	0.1	0.35	0.9	0.8	0.9002	0.8006	0.4003	0.0023	0.0025	0.4877	0.4839	215.64	192.43
0.55	0.1	0.35	0.8	1	0.8000	1.0003	0.3977	0.0015	0.0015	0.4314	0.4286	297.49	287.57
0.55	0.1	0.35	0.8	0.9	0.8000	0.9005	0.4004	0.0016	0.0017	0.4571	0.4545	289.16	269.60
0.55	0.1	0.35	0.8	0.8	0.8000	0.8003	0.4005	0.0018	0.0019	0.4869	0.4839	271.60	251.89
0.55	0.45	0	1	1	1.0003	1.0001	0.3751	0.0703	0.0508	1.0331	0.9184	14.71	18.09
0.55	0.45	0	1	0.9	1.0003	0.8999	0.4090	8.9664	0.0626	1.1250	0.9259	0.13	14.79
0.55	0.45	0	1	0.8	1.0003	0.8001	0.3777	0.8624	0.0791	0.1383	0.9336	0.16	11.80
0.55	0.45	0	0.9	1	0.9002	1.0010	0.3980	0.0145	0.0141	0.9597	0.9184	66.29	65.30
0.55	0.45	0	0.9	0.9	0.9002	0.9007	0.3986	0.0163	0.0155	0.9976	0.9259	61.24	59.84
0.55	0.45	0	0.9	0.8	0.9002	0.8006	0.3981	0.0184	0.0171	0.9974	0.9336	54.08	54.58
0.55	0.45	0	0.8	1	0.8000	1.0003	0.3997	0.0066	0.0065	0.9351	0.9184	142.04	141.94
0.55	0.45	0	0.8	0.9	0.8000	0.9005	0.4006	0.0069	0.0068	0.9441	0.9259	136.24	135.30
0.55	0.45	0	0.8	0.8	0.8000	0.8003	0.4007	0.0074	0.0072	0.9528	0.9336	128.96	128.78
0.7	0	0.3	1	1	1.0003	1.0001	0.3968	0.0009	0.0010	0.0980	0.0968	103.97	95.34
0.7	0	0.3	1	0.9	1.0003	0.8999	0.4019	0.0011	0.0012	0.1077	0.1064	102.47	86.53
0.7	0	0.3	1	0.8	1.0003	0.8001	0.4028	0.0012	0.0015	0.1198	0.1181	95.90	78.22
0.7	0	0.3	0.9	1	0.9002	1.0010	0.3961	0.0008	0.0009	0.0977	0.0968	115.38	104.61
0.7	0	0.3	0.9	0.9	0.9002	0.9007	0.3999	0.0009	0.0011	0.1071	0.1064	117.07	97.41
0.7	0	0.3	0.9	0.8	0.9002	0.8006	0.3999	0.0011	0.0013	0.1190	0.1181	112.76	90.69
0.7	0	0.3	0.8	1	0.8000	1.0003	0.3969	0.0008	0.0008	0.0972	0.0968	121.22	114.26
0.7	0	0.3	0.8	0.9	0.8000	0.9005	0.4002	0.0009	0.0010	0.1064	0.1064	124.12	108.91
0.7	0	0.3	0.8	0.8	0.8000	0.8003	0.4003	0.0010	0.0011	0.1182	0.1181	120.48	104.06
0.7	0.15	0.15	1	1	1.0003	1.0001	0.3958	0.0016	0.0017	0.4313	0.4286	263.58	**252.88**
0.7	0.15	0.15	1	0.9	1.0003	0.8999	0.4008	0.0019	0.0021	0.4588	0.4545	241.86	**219.29**
0.7	0.15	0.15	1	0.8	1.0003	0.8001	0.4015	0.0023	0.0026	0.4896	0.4839	210.87	**187.35**
0.7	0.15	0.15	0.9	1	0.9002	1.0010	0.3957	0.0014	0.0014	0.4287	0.4286	316.16	**299.86**
0.7	0.15	0.15	0.9	0.9	0.9002	0.9007	0.3992	0.0015	0.0017	0.4549	0.4545	296.96	**269.64**
0.7	0.15	0.15	0.9	0.8	0.9002	0.8006	0.3992	0.0018	0.0020	0.4847	0.4839	267.14	**240.48**
0.7	0.15	0.15	0.8	1	0.8000	1.0003	0.3969	0.0012	0.0012	0.4292	0.4286	360.02	**351.05**
0.7	0.15	0.15	0.8	0.9	0.8000	0.9005	0.3998	0.0013	0.0014	0.4548	0.4545	348.02	**325.44**
0.7	0.15	0.15	0.8	0.8	0.8000	0.8003	0.4000	0.0015	0.0016	0.4844	0.4839	322.06	**300.47**
0.7	0.3	0	1	1	1.0003	1.0001	0.3933	0.0033	0.0033	0.6542	0.6522	200.88	200.56
0.7	0.3	0	1	0.9	1.0003	0.8999	0.3982	0.0039	0.0040	0.6819	0.6757	176.67	168.81
0.7	0.3	0	1	0.8	1.0003	0.8001	0.3985	0.0048	0.0050	0.7098	0.7009	147.43	139.22
0.7	0.3	0	0.9	1	0.9002	1.0010	0.3962	0.0024	0.0024	0.6544	0.6522	276.39	267.57
0.7	0.3	0	0.9	0.9	0.9002	0.9007	0.3993	0.0028	0.0029	0.6800	0.6757	246.84	236.57
0.7	0.3	0	0.9	0.8	0.9002	0.8006	0.3992	0.0032	0.0034	0.7064	0.7009	217.42	206.90
0.7	0.3	0	0.8	1	0.8000	1.0003	0.3974	0.0018	0.0019	0.6545	0.6522	354.98	344.88
0.7	0.3	0	0.8	0.9	0.8000	0.9005	0.4001	0.0020	0.0021	0.6790	0.6757	331.92	316.42
0.7	0.3	0	0.8	0.8	0.8000	0.8003	0.4003	0.0023	0.0024	0.7050	0.7009	302.09	288.67

We can make several critical observations about the performance of the proposed model based on the results shown in Table (IV.2).

1. The impact of untruthfulness due to lack of trust in the model shown by Lovig et al. (2021)[41] can be confirmed with these results for fixed values of A. We observe that the Greenberg et al. (1969)[16] model (yellow) produces the best results in terms of MSE and the Warner (1965)[62] model (green) produces the best results in terms of PP. However, our optional mixture model (white) performs the best in terms of the unified measure M.

2. The comparison of the three colored blocks in the M_T column shows that the overall performance of the model improves (i.e. higher M values) as W decreases. This indicates the usefulness of the optional model and how it is an improvement over the non-optional Lovig et al. (2021) model.

3. Further, for $p = 0.7$ and $q = 0.15$, within each of the three blocks highlighted in the M_T column, we note that for a fixed W, the model performance declines (i.e. lower M values) as A goes down. However, this impact of A on M is less adverse for lower values of W. For instance, for $W = 1$, we note a drop of 25.9% in M as A drops from 1 to 0.8. Whereas the corresponding drop in values of M for $W = 0.9$ and $W = 0.8$ is 19.8% and 14.4% respectively. This is yet another piece of evidence that optionality is useful due to its ability to mitigate the effect of a lack of trust on the overall model performance.

IV.3 Concluding Chapter Remarks

In this chapter, a modified version of the model introduced in Chapter III was introduced. The modification made was the inclusion of an optionality element. In practice, this element allows respondents who find the question sensitive to provide a scrambled response. However, if the respondents do not find the question sensitive, they have the option to respond to the sensitive question directly. When we evaluate the performance of this optional mixture binary RRT model, we note that the inclusion of optionality does not harm the privacy offered by the Lovig et al. (2021)[41] model. In addition to this, it always improves the model's efficiency. This results in an overall improvement in the model performance in terms of the unified measure M which was proposed by Lovig et al. (2021)[41]. Therefore, the proposed optional binary RRT model is an improvement over the non-optional mixture binary RRT model proposed by Lovig et al. (2021)[41].

Chapter V: Hybrid (Encryption + RRT) Model

In Chapter I, we introduced various techniques for sensitive data collection and their background. Methods like the Bogus Pipeline method, SDB scale and the Unmatched Count Technique have their limitations with respect to respondent privacy and the extent to which they help mitigate the respondents' social desirability bias. On the other hand, RRT ensures complete privacy, barring a surveyor's ability to guess a respondent's true status with respect to the sensitive survey question based on their reported response. The measure of how tough or easy it could be for a surveyor to guess a respondent's true status is how we measure the amount of privacy loss that can be expected on average under an RRT model. Although in terms of respondent privacy, RRT does a good job, the improved privacy comes at the cost of estimation accuracy. Asymmetric encryption is a method that allows one to collect and decrypt secure and private data with absolutely no error. However, respondent privacy relies completely on the level of security of the decryption key for the protocol. As long as the key cannot be broken or retrieved through corrupt means, the privacy of all respondents is protected. Hence neither RRT nor encryption is perfect as a sensitive data collection method.

In this Chapter, we introduce a hybrid model that combines the elements of an RRT model as well as the Paillier encryption protocol. The goal of this work was to leverage the strengths of both methods to help improve the overall sensitive data

collection through a sample survey.

V.1 Paillier Encryption Protocol

Paillier encryption technique is an asymmetric encryption technique. This means that, unlike symmetric encryption where both the sender and the receiver of a private message have the same key, one needs separate encryption and decryption keys. The sender only has access to the public encryption key while the receiver has access to the decryption key which allows them to decode the original message. The encrypted information cannot be decoded unless one has access to the decryption key. The general encryption-decryption process for any asymmetric encryption scheme has been shown in Figure V.1.

There are three stages involved in a Paillier encryption protocol like any asymmetric encryption system.

Key Generation > **Encryption** > ...*Data Transfer*.... > **Decryption**

The algorithms used in the Paillier encryption scheme at these three stages have been presented below.

Stage-1 Paillier Key Generation Algorithm The following steps can be followed to generate the public-private key pair for a Paillier encryption protocol (30)[34].

Step-1: Two large primes are selected, say p and q such that their greatest common divisor is 1.

Step-2: Set $n = pq$

Step-3: Calculate $\lambda(n) = lcm(p-1, q-1)$ (Here Carmichael function λ returns the least common multiple of $p-1$ and $q-1$)

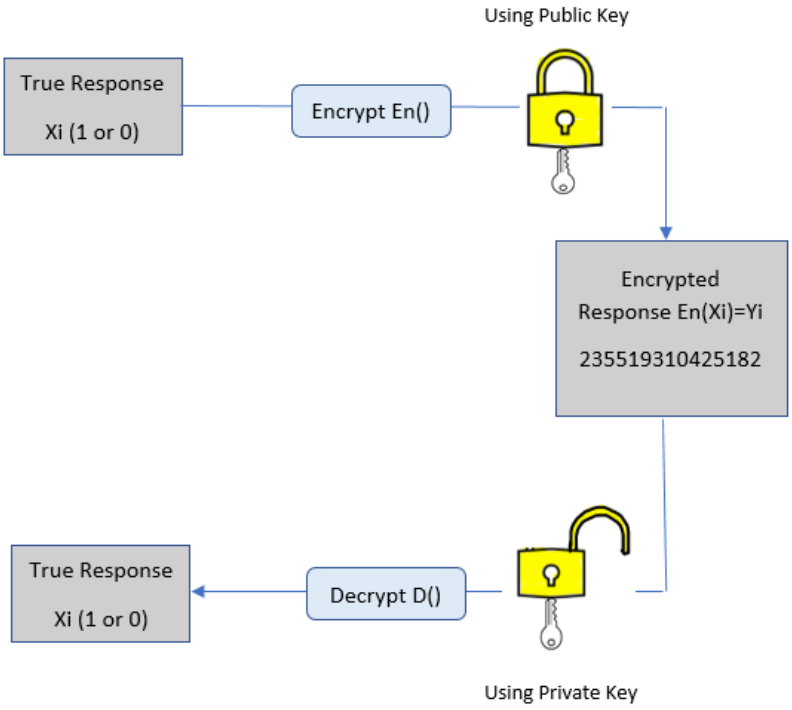

Figure V.1. General Asymmetric Encryption Process

Step-4: Select a random integer g such that $g \in \mathbb{Z}^*_{n^2}$

Step-5: Define $L(x) = \frac{x-1}{n}$

Step-6: Ensure n divides the order of g by confirming the existence of the following modular multiplicative inverse:

$$u = (L(g^{\lambda} \ mod \ n^2))^{-1} \ mod \ (n)$$

Step-7: Public Key $= (n, g)$

Step-8: Private Key $= (\lambda, u)$

Stage-2 Paillier Encryption Algorithm We can encrypt a message, say m, where $m \in \mathbb{Z}_n$ as follows:

Step-1: Generate a random integer r such that $r \in \mathbb{Z}_{n^2}^*$

Step-2: Compute the ciphertext c such that $c = (g^m * r^n) \ mod \ (n^2)$

Stage-3 Evaluation Stage Addition can now be performed on the ciphertext if needed and the sum of ciphertexts is then decrypted using the decryption algorithm in the next stage. Note that since the plaintext is encrypted as an exponent (Step-2 in Stage-2), a multiplication operation needs to be computed on the values of the same base, which in this case is g [34].

If the encrypted sum is y, then it can be decrypted as:

$$d_K(y) = [L(y^\lambda \ mod \ n^2)][L(g^\lambda \ mod \ n^2)]^{-1} \ mod \ n$$

Stage-4 Paillier Decryption Algorithm The encrypted message c, where $c \in \mathbb{Z}_{n^2}^*$, can be decrypted using the following step.

Compute plaintext $m = [L(C^\lambda \ mod \ (n^2)) \times u] \ mod \ n$

This protocol can be used on one respondent at a time in a survey that uses the Paillier encryption scheme for sensitive data collection to help keep every respondent secure from anyone who does not have the decryption key.

V.2 Additive Homomorphism Property of Paillier Encryption Scheme

In Chapter II, we introduced a special property of the Paillier encryption scheme [48]. Suppose that there are two respondents in a survey and their true responses to the survey question are X_1 and X_2 respectively. These original responses, in their unencrypted state, are referred to as the plaintexts. Suppose that the encryption

function used for the Paillier scheme is denoted by $En(.)$ and that when plaintext is supplied o this function, the encrypted value is obtained as an output to this function. Let $En(X_1) = Y_1$ and $En(X_2) = Y_2$ where Y_1 and Y_2 are the encrypted values of the plaintext X_1 and X_2 respectively are also called ciphertexts. Then a special property of Paillier encryption allows us to compute the sum of the original plaintexts X_1 and X_2 even when all we have available is the product of ciphertexts Y_1 and Y_2 and the decryption key. Note that this means, without having access to the plaintext or the ciphertexts of individuals, we can still gauge the sum of responses on an aggregate level.

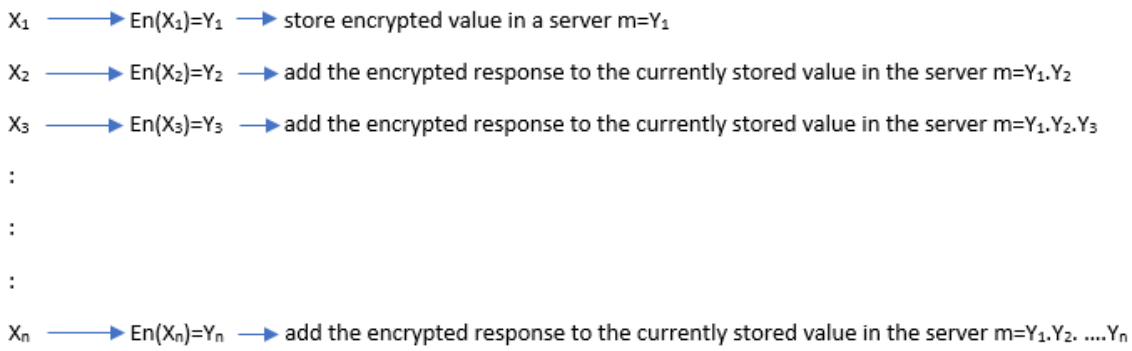

Figure V.2. Simplified Paillier Encryption Process

This special property of Paillier encryption, which allows us to compute the sum of plaintexts, by decrypting only the product of ciphertexts, is referred to as the additive homomorphism of the Paillier encryption scheme. We can generalize and use this property even when we have more than two respondents. The idea to implement the Paillier encryption technique, in a survey with binary response questions, has been depicted in Figure V.2. The respondents can encrypt their responses and once a ciphertext is generated for their response, it is transmitted to a server for storage. After the first encrypted response has been stored on the server, the next respondent's

ciphertext is multiplied by the previously stored cipher text and the new value that is stored now is the product of the first two ciphertexts. Similarly, as more respondents take part in the survey and transmit their responses as ciphertexts, they keep on getting multiplied to the product of the previously transmitted ciphertexts. Once all respondents go through this process, the server would finally just store the product of all the ciphertexts for all the survey respondents. This product can then be decrypted by a third-party organization to obtain the sum of responses from this survey. Note that since we are implementing this method for a binary response question survey (i.e. responses are either 1/"Yes" or 0/"No"), the sum of all responses would essentially be the sum of all 1s or count of "Yes" responses in the survey. This sum can then be utilized to obtain an estimate of a sensitive trait prevalence in the population.

Thus in a binary response survey, we can use the Paillier encryption method to compute

$$D\Big(\big(Y_1 \times Y_2 \times Y_3.... \times Y_n\big)\Big) = Y = X_1 + X_2 + .. + X_n, \tag{V.1}$$

where $D(.)$ denotes the decryption function under the Paillier encryption protocol.

V.3 Example of Paillier Encryption Application

Paillier Encryption Application in E-Voting This technique is used in electronic voting systems used in elections [6]. Suppose Adam, Bill and Chelsea are election candidates and only 6 people can vote in the election. Their votes have been shown in figure(V.3). Note that each candidate is assigned a certain number of bits during the election. Adam was assigned "010000", Bill was assigned "000100" and Chelsea was assigned "000001". These bit scores can then be converted from a binary number

Voter #	Adam	Bill	Chelsea	Bit Score
1			✓	00 00 01 = 1
2		✓		00 01 00 = 4
3		✓		00 01 00 = 4
4			✓	00 00 01 = 1
5	✓			01 00 00 = 16
6			✓	00 00 01 = 1

Figure V.3. Election Results

system to their equivalent in the decimal number system as follows:

- $(010000)_2 = (0 * 2^5) + (1 * 2^4) + (0 * 2^3) + (0 * 2^2) + (0 * 2^1) + (0 * 2^0) = (16)_{10}$.
- $(000100)_2 = (0 * 2^5) + (0 * 2^4) + (0 * 2^3) + (1 * 2^2) + (0 * 2^1) + (0 * 2^0) = (4)_{10}$.
- $(000001)_2 = (0 * 2^5) + (0 * 2^4) + (0 * 2^3) + (0 * 2^2) + (0 * 2^1) + (1 * 2^0) = (1)_{10}$.

Now, for a simple example to see how this method works, we use small prime numbers to generate the keys. Suppose $p = 5$ and $q = 7$ [37]. Then,

$n = p * q = 5 * 7 = 35$ and $n^2 = 35^2 = 1225$

$\lambda = lcm(p - 1, q - 1) = lcm(4, 6) = 12$

Let g, the random element of the public (encryption) key as described in step-4 of the key generation, be chosen as 141.

Suppose, all the votes and the values of random integers denoted by x and r respectively are shown below:

$\mathbf{x}$	$\mathbf{r}$	$\mathbf{e_K(x, r)}$
1	4	359
4	17	173
4	26	486
1	12	1088
16	11	541
1	32	163

Here, the first vote is denoted by $x_1 = 1$ and a random integer r_1 was chosen as 4. Then the ciphertext $e_K(x_1, r_1) = e_K(1, 4) = (141^1 * 4^{35}) \ mod \ 1225$.

To compute the result of the election, the sum of votes, we multiply the encrypted data modulo n^2 and then we decrypt the result as explained in the algorithm. This can be done as shown here:

$(359 * 173 * 486 * 1088 * 541 * 163) \ mod \ 1225 = 983 = y(say)$

This can be decrypted as:

$$L(y^\lambda \ mod \ n^2) = L(983^{12} \ mod \ 1225) = \frac{36-1}{35} = 1$$

$$L(g^\lambda \ mod \ n^2) = L(141^{12} \ mod \ 1225) = \frac{456-1}{35} = 13$$

$$\text{Then decrypted sum } d_K(y) = (1 * 13)^{-1} \ mod \ 35 = 27$$

Thus, a voter's selection can be encrypted such that the ciphertext is incomprehensible to anyone without the private key to decrypt it. The sum of all selections can then be decrypted by the election administrators to evaluate election results without having to decrypt every individual's response thereby providing them with iron-clad privacy with 100% accuracy in evaluating the population proportion, something that can be visualized the same as sensitive trait prevalence.

V.4 Hybrid (Paillier + Warner RRT) Model

One may wonder why one should even consider an alternative such as RRT if we are guaranteed essentially complete privacy and efficiency under the Paillier encryption scheme. However, it is worth noting that such encryption protocols are extremely expensive in terms of computation and their performance is not efficient enough to be of practical use [40], [70], [64]. Another practical concern could be that the true

privacy protection of the survey participants, in addition to the security level of the private key, relies on the honesty of the surveyor, unlike the RRT where the surveyor's intentions cannot expose the respondents as the respondents are not forced to directly disclose nor enter their true response to the sensitive survey question into a system. Moreover, a system based only on (Paillier) encryption requires more resources to deploy a large-scale data collection process.

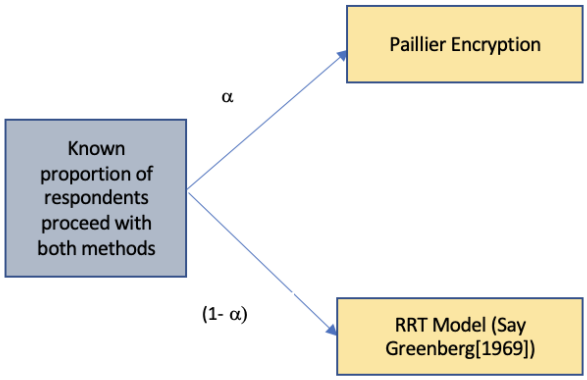

Figure V.4. Hybrid Encryption-RRT Model

However, it is worth noting if researchers have the resources to allow a proportion of their respondents to go through this process (Figure V.4), that would help boost the overall efficiency of the estimation process by balancing out some of the uncertainty brought in by methods (such as RRT) used to collect data from the remaining respondents. Intuitively this should help as a greater portion of truthful responses being collected during a survey should surely boost the overall estimation accuracy. At the moment, there is no such hybrid protocol in the sample survey nor in the encryption literature.

V.4.1 Proposed Hybrid (Paillier + Warner RRT) Model

In this section, we propose a hybrid mixture model which has the elements of both the Paillier encryption (1999)[48] and the Warner's Indirect Question (1965)[62] model. Although one could use any other binary RRT model such as the Greenberg et al. (1969)[16] model, we choose Warner's (1965)[62] model for this study as a special case. The proposed hybrid model has been shown in Figure V.5.

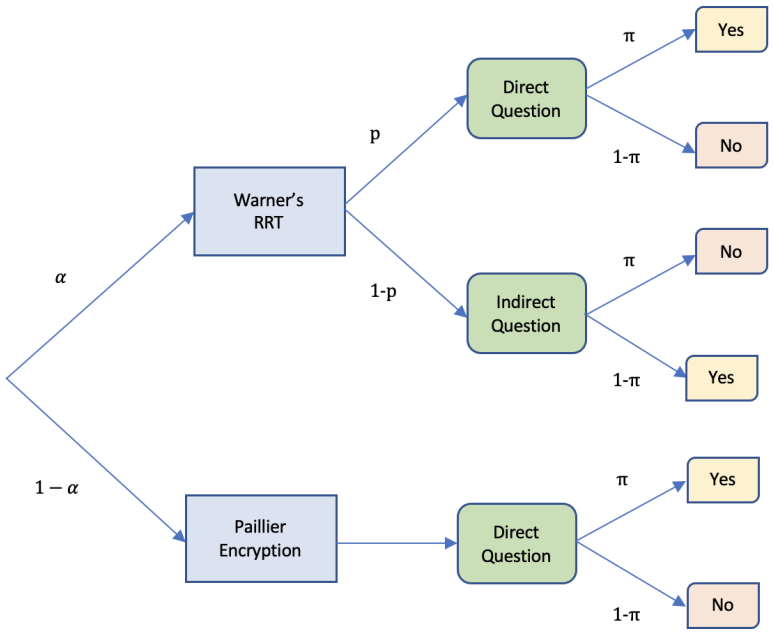

Figure V.5. Hybrid Model using Paillier Encryption & Warner's Indirect Question RRT Model

Here π is the true prevalence of the sensitive trait in the population. Let α be the mixing parameter for the hybrid model which denotes the proportion of respondents that get randomly prompted to go through the Warner's (1965)[62] segment while the remaining $(1-\alpha)$ respondents go through the Paillier encryption scheme. Let p be

the proportion of respondents that go through the Warner's model segment and are asked to respond to the sensitive survey question directly.

First, let us consider just the encryption segment of this model. Let n be the number of respondents using Paillier's Encryption approach and let X_i $(i = 1, 2..., n)$ be the true response of the respondent responding using Paillier's encryption approach. Note that X_i, $(i = 1, 2..., n)$, can take values 0 ("No") or 1 ("Yes"). If $En(.)$ denote the Paillier encryption function and $D(.)$ denotes the Paillier decryption function, then we have

$$X_i \sim Bernoulli(\pi) \tag{V.2}$$

$$\sum_{i=1}^{n} X_i \sim Binomial(n, \pi). \tag{V.3}$$

If $\widehat{P}_{ye}$ denotes the proportion of "Yes" responses in the sample (i.e. the proportion of 1s in the sample) and Y_i denotes the encrypted response of the respondent i.e. $Y_i = En(X_i)$, $(i = 1, 2, .., n)$ then

$$D\bigg(\big(Y_1 \times Y_2 \times Y_3.... \times Y_n\big)\bigg) = Y = \sum_{i=1}^{n} X_i, \tag{V.4}$$

where $Y \sim Binomial(n, \pi)$. Therefore, an estimator for the true sensitive trait prevalence in the population, using a Paillier-only model, would be given by

$$\widehat{\pi}_e = \frac{Y}{n} = \frac{\sum X_i}{n} = \widehat{P}_{ye}. \tag{V.5}$$

The expected value of this estimator for the Paillier encryption segment is given by

$$E(\widehat{\pi}_e) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n}E(y) = \pi. \tag{V.6}$$

Therefore, the binomial estimator for the Paillier encryption segment is unbiased. The variance/MSE for this estimator is given by

$$Var(\widehat{\pi}_e) = Var(\widehat{P}_{ye}) = \frac{\pi(1-\pi)}{n}. \tag{V.7}$$

Now let's shift our focus back to the hybrid model shown in Figure V.5. Suppose that of an overall respondent group of size n, n_1 respondents use the Warner's RRT method while the remaining n_2 respondents use the Paillier encryption method. Here, $n = n_1 + n_2$, $\alpha = \frac{n_1}{n}$ and $(1-\alpha) = \frac{n_2}{n}$.

Then an overall estimator for the proposed Hybrid model can be given by

$$\widehat{\pi}_H = \alpha\widehat{\pi}_w + (1-\alpha)\widehat{\pi}_e, \tag{V.8}$$

where $\widehat{\pi}_w$ is the estimator for the Warner's RRT segment and $\widehat{\pi}_e$ is the estimator for the Paillier encryption segment.

The expected values for the proposed estimator $\widehat{\pi}_H$ is given by

$$E(\widehat{\pi}_H) = \alpha E(\widehat{\pi}_w) + (1-\alpha)E(\widehat{\pi}_e) = \pi. \tag{V.9}$$

The variance for the estimator for the proposed hybrid model is given by

$$\begin{aligned} Var(\widehat{\pi}_H) &= \alpha^2 Var(\widehat{\pi}_w) + (1-\alpha)^2 Var(\widehat{\pi}_e) \\ &= \left(\frac{n_1}{n}\right)^2 \left[\frac{\pi(1-\pi)}{n_1} + \frac{p(1-p)}{n_1(1-2p)^2}\right] + \left(\frac{n_2}{n}\right)^2 \left[\frac{\pi(1-\pi)}{n_2}\right] \\ &= \frac{\pi(1-\pi)}{n} + \left(\frac{n_1}{n}\right) \frac{p(1-p)}{n(1-2p)^2} \\ &\leq \frac{\pi(1-\pi)}{n} + \frac{p(1-p)}{n(1-2p)^2}. \end{aligned} \tag{V.10}$$

Thus,

$$Var(\hat{\pi}_e) \leq Var(\widehat{\pi}_H) \leq Var(\widehat{\pi}_w). \tag{V.11}$$

Hence this hybrid model is more efficient than a purely Warner's Indirect Question model[62].

Paillier encryption, like other encryption techniques, is known to offer a strong security level which can be breached with a very low probability. Various algorithms have been discovered that help in the prime factorization of the public key element n and thus help crack the private decryption key[43]. One such algorithm that helps one use the brute force method requires 2^{k-1} tries where the size of the key used is k-bits. The optimized brute force method requires $\sqrt{n}$ tries where n is an element of the public key and is the composite number for which we are trying to find the two large prime factors p and q. Note that these two numbers of tries needed can be only used as a measure of privacy loss under the Paillier encryption protocol if the researcher is honest and does not misuse the decryption key to retrieve private individual responses.

However, there is no such respondent security if the researcher misuses the key.

No such concern exists for Warner's Indirect Question model (1965)[62]. Given this background, we will only discuss the efficiency of the proposed hybrid model and not its privacy.

V.5 Simulation Study on Hybrid Model

In this section, we examine the performance of the proposed Hybrid mixture of the Paillier encryption[48] and the Warner's Indirect question RRT model[62]. For this purpose, we evaluate the performance of these three models in terms of MSE. Both these measures were computed separately for the RRT and the encryption segments and were then weighted by the mixing parameter α.

Following is an example of how the Paillier segment works in the algorithm for the proposed Hybrid model.

V.5.1 Paillier Encryption Simulation-Example for Binary Responses

Following is a simulated example of what the responses of 10 respondents would look like in one iteration of a binary response survey.

After obtaining the encrypted values $(Y_1, Y_2, \ldots . Y_n)$ for corresponding true responses $(X_1, X_2, ...X_n)$, we compute the value of $(Y_1 \times Y_2 \times ..... \times Y_n)\ mod\ n^2$ i.e. the homomorphic sum of encrypted values (not the regular sum). We then decrypt the above value which yields the sum of true responses i.e.

$$\mathbf{D}[(\mathbf{Y_1} \times \mathbf{Y_2} \times ..... \times \mathbf{Y_n})\ \mathbf{mod}\ \mathbf{n^2}] = \mathbf{X_1} + \mathbf{X_2} + ... + \mathbf{X_n} = \mathbf{D}(\mathbf{f_{PHE}}(\mathbf{Y_1}, \mathbf{Y_2}, ..., \mathbf{Y_n}))$$

Here $f_{PHE}(.)$ takes the product of encrypted responses modulo n^2. Here $n = pq$ which is obtained from the generated key such that p and q are large prime numbers. For the example shown in Table V.1, we have:

Table V.1. Paillier Encryption Example for Binary Responses

Respondent	True Response	X_i	r_i	$En(X_i) = Y_i$ (Encrypted Response)
April	No	0	46	182736476844811857541098659091823052 33
Ben	No	0	42	843479344140819742235228534035653775 86
Claire	No	0	128	449582500106678998021276723788129646 65
Damon	Yes	1	33	106651239789784217596419352015012429788
Elaine	Yes	1	123	990367187956263797444648429478657970 60
Fred	No	0	54	320699662045737292753933496769859672 73
George	No	0	31	406815530308717972218161491885107569 19
Haley	Yes	1	136	987411188594772761552324743473505063 10
Iris	Yes	1	74	817424702478855203476637037652594066 47
Jay	Yes	1	150	840731234876239741387637344154556678 94

$$p = 3014225839 \text{ and } q = 3236597281$$

$$n = 9755835154827343759 \; ; \; n^2 = 951763195681650623736739054675562500 81$$

$$\text{Public (Encryption) Key} = (n, g) \; ; \; \text{Private (Decryption) Key} = (\lambda, \mu)$$

$$\lambda = lcm(p-1, q-1) \; ; \; \mu = \left[L(g^{\lambda} \; mod \; n^2)\right]^{-1} \; mod \; n;$$

Here, $X_1 + X_2 + ... + X_n = 0 + 0 + 0 + 1 + 1 + 0 + 0 + 1 + 1 + 1 = 5$

i.e. the number of "Yes" responses in the sample.

$Y_1 \times Y_2 \times .....Y_n =$ 64799291812010996928111983995289594699808028 5

08145782269507689107742009435906646523237343334809662473579028831 2

57488384412402445696847326949660772804569080232369854765968617482 5

06407082438923076970084176647141269605464160133129308597184570169 40

72727143671979478629650726747433611452178680970873972489464116036720

51125380039759391274926876353984174328749043105220470752661777600 0

Thus, $(Y_1 \times Y_2 \times .....Y_n) \; mod \; n^2 =$ 78559384731536776506688196137760672371

and

$D((Y_1 \times Y_2 \times .....Y_n) \ mod \ n^2) = 5 = X_1 + X_2 + ... + X_n$ i.e. the number of "Yes" responses in the sample.

To study the performance of the proposed hybrid model, we ran a simulation study with 10000 iterations with samples of size $n = 500$ in each iteration. We use various values for α and p (Warner model parameter) to evaluate how it impacts the performance of the proposed model. True sensitive trait prevalence is assumed to be $\pi = 0.3$. The features of the encryption and the decryption keys generated during this study (Stage-1) have been given below:

```
Public Encryption Key:
{
   'g': 14383665101757927738,
   'n': 14383665101757927737,
   'nsquare': 206889821759528897681108067328513941169,
   'max_int': 4794555033919309244
}

Private Decryption Key:
{
   'p': 3733593781,
   'q': 3852498677,
   'psquare': 13939722521521875961,
   'qsquare': 14841746056286750329,
   'p_inverse': 2266736745,
   'hp': 2196775372,
   'hq': 1585761932
```

```
}
```

The simulation results from this study have been summarized in Table V.2. From this table, we can note that when the values of α and p are fixed, the MSE for the proposed hybrid model is always considerably better than that of the traditional Warner's Indirect question model (1965)[62]. For example, in the case when $\alpha = 0.1$ and $p = 0.75$, the theoretical MSE for the estimator proposed by Warner (1965)[62] and that for the proposed estimator are 0.00117 and 0.00057 respectively. However, the MSE for the hybrid model worsens as α increases. For instance, consider the case when $p = 0.75$ for varying values of α. As α varies from 0.1, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 to 0.9, the theoretical MSE of the proposed estimator for the hybrid model introduced in this chapter takes values 0.00057, 0.00087, 0.00102, 0.00117, 0.00132, 0.00147, 0.00162 and 0.00177 respectively. This happens because a higher value of α represents a lower proportion of respondents going through the encryption segment of the proposed hybrid model. This leads to more noise in the aggregate data thus lowering the MSE. Therefore, it is recommended to keep the α values as low as possible. We also observe that for a fixed level of α, as p decreases, the MSE for the proposed estimator for the hybrid model increases. For instance, consider the case when $\alpha = 0.1$. When p changes from 0.85, 0.8, and 0.75, the theoretical values of the MSE_H are 0.000472, 0.000509 and 0.000570 respectively. Hence, it would be advisable to use a higher value of p. This recommendation aligns with the fact that the MSE for the model worsens as p moves closer to 0.5 as stated in Warner (1965)[62].

Table V.2. Simulation results for Hybrid model: $Iterations = 10000$, $n = 500$, $\pi = 0.3$ for various levels of mixture (α) and the Warner's model parameter (p).

α	**p**	$\hat{\pi}$	$\mathbf{MSE_h - E}$	$\mathbf{MSE_h - T}$	$\mathbf{MSE_w - T}$	$\mathbf{MSE_{en} - T}$
0.1	0.85	0.299966	0.000480	0.000472	0.000784	0.000420
0.1	0.8	0.299972	0.000532	0.000509	0.000953	0.000420
0.1	0.75	0.299959	0.000617	0.000570	0.001170	0.000420
0.3	0.85	0.299940	0.000594	0.000576	0.000784	0.000420
0.3	0.8	0.299938	0.000740	0.000687	0.000953	0.000420
0.3	0.75	0.299776	0.000970	0.000870	0.001170	0.000420
0.4	0.85	0.299992	0.000641	0.000628	0.000784	0.000420
0.4	0.8	0.300124	0.000822	0.000776	0.000953	0.000420
0.4	0.75	0.300015	0.001118	0.001020	0.001170	0.000420
0.5	0.85	0.300004	0.000688	0.000680	0.000784	0.000420
0.5	0.8	0.300061	0.000904	0.000864	0.000953	0.000420
0.5	0.75	0.300051	0.001276	0.001170	0.001170	0.000420
0.6	0.85	0.300381	0.000777	0.000732	0.000784	0.000420
0.6	0.8	0.300299	0.001025	0.000953	0.000953	0.000420
0.6	0.75	0.300287	0.001449	0.001320	0.001170	0.000420
0.7	0.85	0.299534	0.000819	0.000784	0.000784	0.000420
0.7	0.8	0.299315	0.001096	0.001042	0.000953	0.000420
0.7	0.75	0.299207	0.001584	0.001470	0.001170	0.000420
0.8	0.85	0.299601	0.000860	0.000836	0.000784	0.000420
0.8	0.8	0.299591	0.001181	0.001131	0.000953	0.000420
0.8	0.75	0.299403	0.001742	0.001620	0.001170	0.000420
0.9	0.85	0.299955	0.000909	0.000888	0.000784	0.000420
0.9	0.8	0.299770	0.001242	0.001220	0.000953	0.000420
0.9	0.75	0.299639	0.001829	0.001770	0.001170	0.000420

V.6 Concluding Chapter Remarks

In this chapter, we proposed a hybrid model that has elements of both the traditional Warner's Indirect Question model (1965)[62] as well as the partially homomorphic Paillier encryption scheme (1999)[48]. We also compare the performance of the proposed model with those of Warner's Indirect Question model (1965)[62] and Paillier encryption scheme (1999)[48]. Based on our results, we make various important observations.

If one were to simply consider efficiency, the Paillier encryption method performs the best. However, the respondent privacy under this method, in addition to the

security level of the key, also relies on the honesty level of the surveyor. It must be noted that the surveyor, who has the decryption key, could also potentially decrypt individual responses if they wanted. Thus, there is no privacy in the case the surveyor is dishonest. Since, under RRT models there is no such concern, a mixture such as the model proposed in this chapter leads to better respondent cooperation.

Through this study, we were able to confirm that such a hybrid model could potentially help give a huge boost to the efficiency of the survey model when the survey question is on a sensitive topic. Prior to this work, no such hybrid mixture model has ever been studied in the area of RRT. One could potentially improve upon the efficiency and primary protection offered by the hybrid model by switching the Warner's Indirect Question model (1965)[62] segment with a method such as the Greenberg's Unrelated Question model (1969)[16] or with the Lovig et al. (2021) Mixture binary RRT model[41]. Furthermore, the primary protection of the encryption model can be modified to account for the damage to the respondents' privacy due to potential dishonesty or corruption of the surveyor.

Chapter VI: Mitigating Lack of Trust in Quantitative RRT Models

In Chapter II, we presented different types of RRT models. In particular, we discussed how the type of responses possible for the sensitive question necessitates a different scrambling mechanism to help protect respondent privacy. If the survey question could result in binary responses such as "Yes" or "No", then we need a binary RRT model. Binary RRT models can be helpful when the survey question might be "Have you smoked marijuana in the last two weeks?". However, when the survey question requires a quantitative response, we need to use quantitative RRT models. Such models may be helpful when the survey question might be "How many times have you smoked marijuana in the last two weeks?". In Section II.3, we introduced a few traditional quantitative RRT models.

One of the first quantitative RRT models was also proposed by S L Warner, the same researcher who presented the first ever RRT model, i.e. the Indirect Question Binary model[62]. Warner (1971)[63] also proposed an RRT model for surveys where the sensitive question has a quantitative response. Under this model, each respondent is asked to scramble his/her true response with a random additive or multiplicative noise before they report it. Gupta et al. (2002)[19] proposed an alternative RRT model known as an Optional RRT model that allows respondents to report their unscrambled

response if they do not find the question sensitive, and a scrambled response otherwise. They have shown empirically and theoretically, that optional models tend to be more efficient than their corresponding non-optional counterparts. Several other significant RRT methods have been proposed over the past few decades by various researchers including Diana and Perri (2011)[10], Kalucha et al. (2016)[30], Mehta and Aggarwal (2018)[42], Narjis and Shabbbir (2021)[45], Khalil et al. (2021)[33] and Zhang et al. (2021)[69].

Typically, one would think that RRT models can be trusted and hence the respondent would have no reason to lie. However, in the area of binary RRT models, Young et al. (2019)[67] first introduced the idea of respondents' lack of respondent in RRT models and showed how even a small proportion of respondents not trusting the RRT model can lead to unreliable data and subsequently biased estimates. Lovig et al. (2021)[41] also addressed the lack of respondent trust and accounted for untruthfulness by introducing a Mixture Binary RRT Model. While accounting for lack of trust has been addressed under binary RRT models, none of the quantitative RRT models currently accounted for respondents' lack of trust in the quantitative RRT models prior to Zhang et al. (2022)[25].

In this Chapter[1], we propose an Optional Enhanced Trust model that helps mitigate the effect of respondents' lack of trust in the quantitative RRT models. This model mixes the elements of the Warner Additive Model (1971)[63] and the Diana and Perri linear combination model (2011)[10] to mitigate respondent lack of trust.

[1]A portion of this chapter is based on an Accepted Manuscript of an article published by Taylor & Francis in Journal of Communications in Statistics - Simulation and Computation on 3 June 2022, available online: https://www.tandfonline.com/doi/full/10.1080/03610918.2022.2082477

VI.1 Background for Optional Enhanced Trust Model

VI.1.1 Efficiency of RRT Models

Model efficiency is considered a measure of the quality of the model estimator, usually measured through mean squared error (MSE). The MSE accounts for both the variance and the bias associated with the estimator. This measure is used as a measure of estimator efficiency for both binary and quantitative RRT models.

VI.1.2 Privacy Level in Quantitative RRT Models

Respondent privacy is crucial for the anonymity of respondents as well as the mitigation of non-response and untruthfulness. According to Yan et al. (2008)[66], privacy level for quantitative RRT models is given by

$$\nabla = E[Z - Y]^2. \tag{VI.1}$$

A higher value of ∇ is preferred as it indicates a higher average deviation from the true response i.e. a higher privacy level.

VI.1.3 Combined Measure for Efficiency and Privacy

In order to choose an appropriate model, we need to consider both efficiency and privacy levels simultaneously. Since there is usually a trade-off between these two characteristics, Gupta et al. (2018)[21] proposed a combined measure of estimator quality (δ) to simultaneously evaluate a model based on its efficiency as well as respondent privacy. This combined measure for efficiency and privacy is given by

$$\delta = \frac{MSE(\hat{\mu}_Y)}{\bigtriangledown}. \tag{VI.2}$$

Here $MSE(\hat{\mu}_Y)$ is the MSE for the estimator $\hat{\mu}_Y$ and ∇ is the privacy level for the quantitative RRT model used in the survey. A lower value of δ is preferred as a lower value of δ indicates a lower variance (i.e. greater efficiency), or a higher level of privacy, or both.

VI.1.4 Warner Additive Model (1971)

In Chapter II, we introduced the following additive RRT model proposed by Warner (1971)[63]. Let Y be the sensitive variable and Z be the reported response. Let S be the additive scrambling variable with mean μ_S and variance σ_S^2. Then the reported response Z is given by

$$Z = Y + S, \tag{VI.3}$$

If we assume that $E(S) = 0$, then

$$E(Z) = E(Y) + E(S) = \mu_Y. \tag{VI.4}$$

Then, an unbiased estimator for the mean of the sensitive variable can be given by

$$\widehat{\mu_Y} = \bar{Z}. \tag{VI.5}$$

For this estimator,

$$Var(\hat{\mu}_Y) = Var(\bar{Z}) = \frac{\sigma_Z^2}{n} = \frac{1}{n}(\sigma_Y^2 + \sigma_S^2), \tag{VI.6}$$

where n is the sample size. Here, $\frac{\sigma_S^2}{n}$ is the penalty for adding the scrambling noise through the RRT model shown by equation (VI.3). The privacy level offered by this model is given by

$$\nabla = E[(Y + S) - Y]^2 = E[S]^2 = \sigma_S^2. \tag{VI.7}$$

Therefore, the combined measure for privacy and efficiency, as defined by Gupta et al. (2018)[21] is given by

$$\delta = \frac{1}{n}\left(\frac{\sigma_Y^2 + \sigma_S^2}{\sigma_S^2}\right) = \frac{1}{n}\left(\frac{\sigma_Y^2}{\sigma_S^2} + 1\right). \tag{VI.8}$$

Under the model given by equation (VI.3), all subjects are required to scramble their responses by adding the random noise they obtain as the outcome of a randomization device that generates noise values from a distribution with known mean and variance. From equations (VI.6, VI.7, VI.8) we note that using an additive scrambling variable with larger variance worsens the model efficiency but it also simultaneously improves the privacy level. The overall combined measure δ improves (i.e. becomes lower) indicating that the researchers should not hesitate to use some extra noise.

VI.1.5 A Linear Combination Model (2011)

Diana and Perri (2011)[10] proposed a linear combination model which allows the survey respondents to use a multiplicative noise as well as an additive noise to scramble their response. Their goal was to optimize the model efficiency and privacy level by leveraging the benefits of both additive and multiplicative scrambling techniques. Since the respondents scramble their true responses before they report them, surveyors cannot know the true answers of individual respondents.

Let Y be the sensitive variable and Z be the reported response. Let S and T be the scrambling variables with means μ_S, μ_T and variances σ_S^2 and σ_T^2 respectively. Here, Y, T and S are mutually independent. Then the linear combination model proposed by Diana and Perri (2011)[10] is given by

$$Z = TY + S. \tag{VI.9}$$

If it is assumed that $E[T] = \mu_T = 1$ and $E[S] = \mu_S = 0$, then the expected value and the variance of the reported response would be given by

$$E(Z) = \mu_Y, \tag{VI.10}$$

and

$$Var(Z) = \sigma_T^2[\mu_Y^2 + \sigma_Y^2] + \sigma_Y^2 + \sigma_S^2. \tag{VI.11}$$

Therefore, an unbiased estimator of μ_Y given by

$$\hat{\mu}_Y = \frac{\bar{Z} - \mu_S}{\mu_T} = \bar{Z}. \tag{VI.12}$$

The variance of the estimator shown in equation (VI.12) is given by

$$Var(\hat{\mu}_Y) = \frac{1}{n}[\sigma_T^2(\mu_Y^2 + \sigma_Y^2) + \sigma_Y^2 + \sigma_S^2], \tag{VI.13}$$

The privacy level for the model shown in equation (VI.9) and the combined measure of privacy and efficiency δ[21] are given by

$$\bigtriangledown = \sigma_T^2\sigma_Y^2 + \sigma_T^2\mu_Y^2 + \sigma_S^2, \tag{VI.14}$$

and

$$\delta = \frac{1}{n}\left(\frac{\sigma_T^2\sigma_Y^2 + \sigma_T^2\mu_Y^2 + \sigma_S^2 + \sigma_Y^2}{\sigma_T^2\sigma_Y^2 + \sigma_T^2\mu_Y^2 + \sigma_S^2}\right) = \frac{1}{n}\left(1 + \frac{\sigma_Y^2}{\sigma_T^2\sigma_Y^2 + \sigma_T^2\mu_Y^2 + \sigma_S^2}\right). \tag{VI.15}$$

From equations (VI.13, VI.14, VI.15) we can see that introducing a multiplicative noise with larger variance, σ_T^2, worsens the efficiency of the model but it also simultaneously improves the privacy level with an overall gain in model quality.

VI.2 Estimation of the Mean and Sensitivity Level using Optional Enhanced Trust (OET) Model

In this section, we propose an Optional Enhanced Trust model (Figure VI.1), which combines the elements of both the Warner's additive model (1971)[63] and the Diana and Perri (2011)[10] Linear Combination model. This model integrates the strengths of optionality with an enhanced scrambling technique for respondents who do not trust the Warner's Additive model (1971)[63].

Let Y be the sensitive study variable and Z be the reported response. Let S and T be the scrambling variables with means μ_S, μ_T and variances σ_S^2 and σ_T^2 respectively. Moreover, let W represent the sensitivity level of the survey question meaning a proportion $(1-W)$ of the respondents do not consider the question sensitive and hence will provide an unscrambled response. Let A represent the proportion of respondents that trust the Warner's Additive model(1971)[63] and hence do not need additional noise. Here, Y, T and S are assumed to be mutually independent.

The proposed Optional Enhanced Trust model mitigates the effect of respondents'

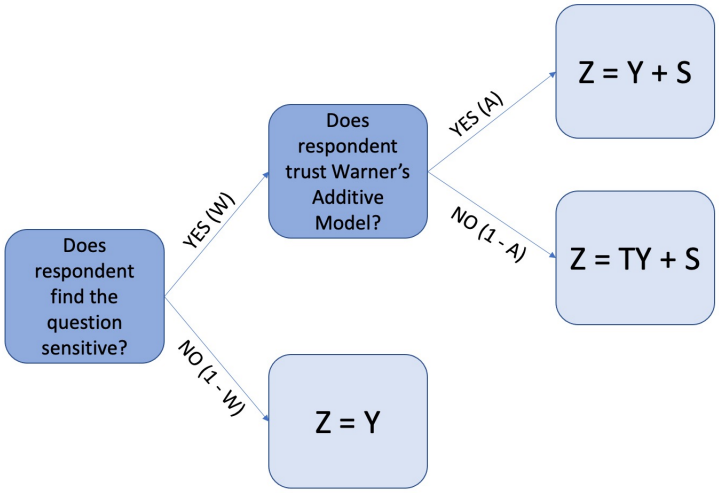

Figure VI.1. Optional Enhanced Trust Model

lack of trust by allowing more noise to respondents who do not trust the Warner's Additive model (1971) [63]. The proposed model is an optional RRT model since it allows respondents to simply report their true responses if they do not find the survey question sensitive. However, if they do find the question sensitive they have the option to scramble their response using either of the two scrambling techniques available to them based on whether or not they trust the additive model. Under this model, the reported response is given as shown in equation (VI.16).

$$Z = \begin{cases} Y & with\ probability\ 1-W \\ Y+S & with\ probability\ WA \\ TY+S & with\ probability\ W(1-A). \end{cases} \tag{VI.16}$$

Respondents who trust Warner's Additive model (1971)[63] simply use a random additive noise to scramble their response before reporting it to the surveyor. The respondents who do not trust the Warner's Additive model (1971)[63] alone are allowed

to use first a multiplicative noise and then an additive noise to scramble their true response before they report it. This alternative scrambling technique is based on the Diana and Perri (2011)[10] model. It helps in improving the respondent privacy level which can help in lowering the level of untruthfulness. The surveyor would still just have information about the reported response (Z) but has no idea about whether the respondent reported their true response (Y), or a scrambled response based on one of the two scrambling schemes available to them as shown in equation (VI.16).

To estimate the sensitivity level W in addition to the sensitive mean μ_Y, we assume that $E[S] = \theta$ and $E[T] = 1$. Then the expected value of the reported response Z would be given by

$$E[Z] = (1 - W)E[Y] + (WA)E[Y + S] + W(1 - A)E[TY + S] = \mu_Y + W\theta. \quad \text{(VI.17)}$$

Note that equation (VI.17) is based on two unknown parameters, the unknown mean μ_Y and the unknown sensitivity level W. However, one could get around this problem by using a Split-Sample approach as used in Gupta et al. (2010)[22]. In order to implement the split-sample technique with the OET model, we split the complete sample of size n into two sub-samples of size n_1 and n_2 such that $n_1 = n_2 = \frac{n}{2}$. Here the survey respondents in the i^{th} sub-sample use scrambling variables T and S_i $(i = 1, 2)$ with means μ_T and θ_i and variances σ_T^2 and $\sigma_{S_i}^2$ $(i = 1, 2)$ respectively. We may point out that an equal sample split is not necessary. We do so only for convenience.

Under this model, the expected value and the variance of the reported response in the i^{th} sub-sample $(i = 1, 2)$ is given by

$$E[Z_i] = \mu_Y + W\theta_i, \tag{VI.18}$$

and

$$\sigma^2_{Z_i} = Var(Z_i) = \sigma^2_Y + W(1-A)(\sigma^2_T\sigma^2_Y + \sigma^2_T\mu^2_Y) + W\sigma^2_S + W\theta^2_i - W^2\theta^2_i, \ . \tag{VI.19}$$

Estimating $E(Z_i)$ by $\overline{Z_i}$, we get

$$\overline{Z_1} = \hat{\mu}_Y + \hat{W}\theta_1, \tag{VI.20}$$

and

$$\overline{Z_2} = \hat{\mu}_Y + \hat{W}\theta_2. \tag{VI.21}$$

Using equations (VI.20) and (VI.21), the estimators for the sensitive mean μ_Y and the sensitivity level W are given by

$$\hat{\mu}_Y = \frac{\theta_1\overline{Z_2} - \theta_2\overline{Z_1}}{\theta_1 - \theta_2}, \tag{VI.22}$$

and

$$\hat{W} = \frac{\overline{Z_1} - \overline{Z_2}}{\theta_1 - \theta_2}. \tag{VI.23}$$

The variances for the estimators in Equations (VI.22) and (VI.23) are given by

$$Var(\hat{\mu}_Y) = \frac{(\theta_1^2+\theta_2^2)(\sigma_Y^2 + W(1-A)(\sigma_T^2\sigma_Y^2+\sigma_T^2\mu_Y^2)+W\sigma_S^2)+W(1-W)2\theta_1^2\theta_2^2}{n_s(\theta_1-\theta_2)^2}, \tag{VI.24}$$

$$Var(\hat{W}) = \frac{2(\sigma_Y^2 + W(1-A)(\sigma_T^2\sigma_Y^2+\sigma_T^2\mu_Y^2)+W\sigma_S^2)+(W-W^2)(\theta_1^2+\theta_2^2)}{n_s(\theta_1-\theta_2)^2}. \tag{VI.25}$$

Note that both (VI.24) and (VI.25) suggest using θ_1 and θ_2 that are not too close to each other.

The privacy level and the combined measure of privacy and efficiency for the model given in equation (VI.16) are given by

$$\bigtriangledown = (1-A)(\sigma_T^2\sigma_Y^2+\sigma_T^2\mu_Y^2)+\sigma_S^2+\frac{\theta_1^2+\theta_2^2}{2}, \tag{VI.26}$$

and

$$\delta = \frac{(\theta_1^2+\theta_2^2)(\sigma_Y^2 + W(1-A)(\sigma_T^2\sigma_Y^2+\sigma_T^2\mu_Y^2)+W\sigma_S^2)+W(1-W)2\theta_1^2\theta_2^2}{n_s(\theta_1-\theta_2)^2((1-A)(\sigma_T^2\sigma_Y^2+\sigma_T^2\mu_Y^2)+\sigma_S^2+\frac{\theta_1^2+\theta_2^2}{2})}. \tag{VI.27}$$

VI.3 Ratio Estimator of the Mean for the OET Model

In Subsection II.5.2, we presented a discussion on various types of sensitive mean estimators that can be used in the presence of one non-sensitive auxiliary variable which has a strong positive correlation with the sensitive study variable. The types of estimators introduced were ratio, regression and generalized estimators in the presence

of one non-sensitive auxiliary variable. Most of these estimators were either proposed under the non-optional or optional versions of the Warner's Additive model (1971)[63] or the Pollock and Beck model (1976)[50] where the reported response Z is given by

$$Z = Y + S, \tag{VI.28}$$

where S is random scrambling variable, usually with mean $\mu_S = 0$ and variance σ_S^2. In this section, we will introduce a ratio estimator for the mean of the sensitive variable Y when complete information is available for the non-sensitive auxiliary variable X which has a strong positive correlation with Y.

In Subsection II.5.2, we introduced the additive ratio estimator proposed by Kalucha et al. (2015)[29] under the optional version of Warner's additive model (1971)[63]. Following this work, we propose an additive ratio estimator for the sensitive mean μ_Y under the Optional Enhanced Trust (OET) model introduced in Section VI.2.

Let Y be the sensitive study variable and X be the non-sensitive auxiliary variable which has a strong positive correlation with Y. Let T and S_i $(i = 1, 2)$ be the scrambling variables that are all assumed to be independent of both Y and X. Let μ_T and σ_T^2 be the mean and the variance of T, and θ_i and $\sigma_{S_i}^2$ $(i = 1, 2)$ be the mean and the variance of the scrambling variables S_i $(i = 1, 2)$. Let $\mu_X = E(X)$, $\mu_Y = E(Y)$, $\mu_{Z_1} = E(Z_1)$ and $\mu_{Z_2} = E(Z_2)$ be the population means for X, Y, Z_1 and Z_2 respectively where Z_i $(i = 1, 2)$ denote the reported responses from the two sub-samples and are given by,

$$Z_i = \begin{cases} Y & \textit{with probability } 1 - W \\ Y + S_i & \textit{with probability } WA \\ TY + S_i & \textit{with probability } W(1 - A), \end{cases} \tag{VI.29}$$

where $i = 1, 2$. If a simple random sample of size n is drawn, without replacement, then the additive ratio estimator of the sensitive mean μ_Y is given by

$$\widehat{\mu_{ar}} = \left[\frac{\bar{z_1}\theta_2 - \bar{z_2}\theta_1}{\theta_2 - \theta_1}\right]\left[\frac{\mu_X}{\bar{x_1}} + \frac{\mu_X}{\bar{x_2}}\right]\left(\frac{1}{2}\right). \tag{VI.30}$$

Here n_1 is the size of the first sub-sample and n_2 is the size of the second sub-sample such that $n_1 + n_2 = n$. For this study, although it is not required, we use $n_1 = n_2$ for easier interpretation. Further, $\bar{z}_i$ and $\bar{x}_i$ $(i = 1, 2)$ are the sub-sample means for the reported response Z and the auxiliary variable X respectively.

Assume that a large sample is drawn so that $|\delta_{Z_i}| < 1$ and $|\delta_{X_i}| < 1$ $(i = 1, 2)$, where $\delta_{Z_1} = \frac{\bar{z_1} - \mu_{Z_1}}{\mu_{Z_1}}$, $\delta_{Z_2} = \frac{\bar{z_2} - \mu_{Z_2}}{\mu_{Z_2}}$, $\delta_{X_1} = \frac{\bar{x_1} - \mu_{X_1}}{\mu_{X_1}}$ and $\delta_{X_2} = \frac{\bar{x_2} - \mu_{X_2}}{\mu_{X_2}}$. Then the additive ratio estimator from equation (VI.30) can be re-written as

$$\widehat{\mu_{ar}} = \left(\frac{1}{2}\right)\left[\mu_Y + \frac{\theta_1}{\theta_1 - \theta_2}\delta_{Z_2}\mu_{Z_2} - \frac{\theta_2}{\theta_1 - \theta_2}\delta_{Z_1}\mu_{Z_1}\right]\left[(1+\delta_{X_1})^{-1} + (1+\delta_{X_2})^{-1}\right] \tag{VI.31}$$

Under the assumptions of bivariate normality (Sukhatme et al. (1970))[59]: $E(\delta_{X_1}) = 0$; $E(\delta_{X_2}) = 0$; $E(\delta_{Z_1}) = 0$; $E(\delta_{Z_2}) = 0$; $E(\delta_{X_1}^2) = (\frac{1-f_1}{n_1})C_{X_1}^2$; $E(\delta_{X_2}^2) = (\frac{1-f_2}{n_2})C_{X_2}^2$; $E(\delta_{Z_1}^2) = (\frac{1-f_1}{n_1})C_{Z_1}^2$; $E(\delta_{Z_2}^2) = (\frac{1-f_2}{n_2})C_{Z_2}^2$; $E(\delta_{X_1}\delta_{X_2}) = 0$, $E(\delta_{Z_1}\delta_{Z_2}) = 0$; $E(\delta_{Z_1}\delta_{X_2}) = 0$; $E(\delta_{Z_2}\delta_{X_1}) = 0$; $E(\delta_{Z_1}\delta_{X_1}) = (\frac{1-f_1}{n_1})C_{Z_1X_1}$; $E(\delta_{Z_2}\delta_{X_2}) = (\frac{1-f_2}{n_2})C_{Z_2X_2}$, where $f_1 = \frac{n_1}{N}$, $f_2 = \frac{n_2}{N}$, $C_{Z_1X_1} = \rho_{Z_1X_1}C_{Z_1}C_{X_1}$ and $C_{Z_2X_2} = \rho_{Z_2X_2}C_{Z_2}C_{X_2}$. Since the two sub-samples come from the same population, the coefficients of variation $C_{X_1} = C_{X_2} = C_X$, and the correlations $\rho_{Z_1X_1} = \rho_{Z_1X}$ and $\rho_{Z_2X_2} = \rho_{Z_2X}$.

Thus using the first-order Taylor's approximation, the bias of the additive ratio estimator under the OET model is given by

$$Bias(\widehat{\mu_{ar}}) \approx \frac{\mu_Y}{2}\left[\left(\frac{1-f_1}{n_1}\right)C_{X_1}^2 + \left(\frac{1-f_2}{n_2}\right)C_{X_2}^2\right] + \frac{1}{2}\rho_{YX}\sigma_Y C_X\left[\left(\frac{\theta_1}{\theta_1-\theta_2}\right)\left(\frac{1-f_2}{n_2}\right) - \left(\frac{\theta_2}{\theta_1-\theta_2}\right)\left(\frac{1-f_1}{n_1}\right)\right] \quad \text{(VI.32)}$$

The expression for the MSE of the additive ratio estimator under the OET model, correct up to the first order of approximation, is given by

$$\begin{aligned} MSE(\widehat{\mu_{ar}}) &= E(\widehat{\mu_{ar}} - \mu_Y)^2 \\ &\approx \frac{\mu_Y^2}{4}C_X^2\left[\left(\frac{1-f_1}{n_1}\right) + \left(\frac{1-f_2}{n_2}\right)\right] + \left(\frac{\theta_1}{\theta_1-\theta_2}\right)^2\left(\frac{1-f_2}{n_2}\right)\sigma_{Z_2}^2 \\ &+ \left(\frac{\theta_2}{\theta_1-\theta_2}\right)^2\left(\frac{1-f_1}{n_1}\right)\sigma_{Z_1}^2 + \frac{1}{\theta_1-\theta_2}\mu_Y\rho_{YX}\sigma_Y C_X\left[\theta_2\left(\frac{1-f_1}{n_1}\right) - \theta_1\left(\frac{1-f_2}{n_2}\right)\right] \end{aligned} \quad \text{(VI.33)}$$

The privacy level offered under this scenario is given by

$$\nabla = E[Z-Y]^2 = (1-A)(\sigma_T^2\sigma_Y^2 + \sigma_T^2\mu_Y^2) + \sigma_S^2 + \frac{\theta_1^2+\theta_2^2}{2}. \quad \text{(VI.34)}$$

Note that, since the model used to compute the privacy level for the ordinary RRT estimator under the OET model, given by equation (VI.22), has not changed, the privacy level offered to the respondents would be the same whether auxiliary variable X is available or not.

The combined measure for privacy and efficiency is given by

$$\delta = \frac{MSE(\widehat{\mu}_{ar})}{\nabla}, \tag{VI.35}$$

where ∇ is given by equation (VI.34).

VI.4 Regression Estimator of the Mean for the OET Model

In Subsection II.5.2, we introduced the regression estimator proposed by Gupta et al. (2012)[20] which estimates, the mean μ_Y of the sensitive study variable Y and the sensitivity level W, when the information about the non-sensitive and highly correlated auxiliary variable X is available for each individual in the population. This work was done under the optional version of Warner's additive model (1971)[63] and the Pollock and Beck (1976)[50] model.

Following this work, we propose a regression estimator for the sensitive mean while simultaneously estimating the sensitivity level W, under the Optional Enhanced Trust (OET) model introduced in Section VI.2.

Let Y be the sensitive study variable with mean μ_Y and variance σ_Y^2, and X be the non-sensitive auxiliary variable, which has a strong positive correlation with Y, and has a mean μ_X and variance σ_X^2. Let T and S_i $(i = 1, 2)$ be the scrambling variables that are all assumed to be independent of both Y and X. Let the μ_T and σ_T^2 be the mean and the variance of T and θ_i and $\sigma_{S_i}^2$ $(i = 1, 2)$ be the mean and the variance of the scrambling variables S_i $(i = 1, 2)$. Let $\mu_{Z_1} = E(Z_1)$ and $\mu_{Z_2} = E(Z_2)$ be the population means for Z_1 and Z_2 respectively where Z_i $(i = 1, 2)$ denote the reported responses from the two sub-samples are given by,

$$Z_i = \begin{cases} Y & with\ probability\ 1-W \\ Y+S_i & with\ probability\ WA \\ TY+S_i & with\ probability\ W(1-A), \end{cases} \tag{VI.36}$$

where $i = 1, 2$. If a simple random sample of size n is drawn, without replacement, then a regression estimator of the sensitive mean μ_Y under the OET model is given by

$$\widehat{\mu_{AReg}} = \left[\frac{\bar{z_1}\theta_2 - \bar{z_2}\theta_1}{\theta_2 - \theta_1}\right] + \left[\widehat{\beta_{Z_1X_1}}(\mu_X - \bar{x_1}) + \widehat{\beta_{Z_2X_2}}(\mu_X - \bar{x_2})\right]\left(\frac{1}{2}\right), \tag{VI.37}$$

where $\widehat{\beta_{Z_iX_i}} = \frac{s_{z_ix_i}}{\sigma^2_{X_i}}$ $(i = 1, 2)$ are the sample regression coefficients between Z_i and X_i respectively. Further, $\bar{z_i}$ and $\bar{x_i}$ $(i = 1, 2)$ are the sub-sample means for the reported response Z and the auxiliary variable X respectively. Again we define the following error terms:

$$\delta_{z_1} = \frac{\bar{z_1}-\mu_{Z_1}}{\mu_{Z_1}};\ \delta_{z_2} = \frac{\bar{z_2}-\mu_{Z_2}}{\mu_{Z_2}};\ \delta_{z_1} = \frac{\bar{x_1}-\mu_{X_1}}{\mu_{X_1}};\ \delta_{x_2} = \frac{\bar{x_2}-\mu_{X_2}}{\mu_{X_2}};$$
$$\delta_{S_{x_1}} = \frac{s^2_{x_1}-\sigma^2_X}{\sigma^2_X};\ \delta_{S_{x_2}} = \frac{s^2_{x_2}-\sigma^2_X}{\sigma^2_X};\ \delta_{S_{z_1x}} = \frac{s^2_{z_1x_1}-\sigma_{Z_1X}}{\sigma_{Z_1X}};\ \delta_{S_{z_2x}} = \frac{s^2_{z_2x_2}-\sigma_{Z_2X}}{\sigma_{Z_2X}}$$

Here s_i^2 is the sample variance for the auxiliary variable in the i^{th} sub-sample and $s_{z_ix_i}$ is the sample covariance between the reported response and the auxiliary variable response in the i^{th} sub-sample $(i = 1, 2)$. Since the two sub-samples were ultimately drawn from the same population, $\mu_X = E(X) = E(X_1) = E(X_2)$ and $\sigma^2_X = \sigma^2_{X_1} = \sigma^2_{X_2}$. Let σ_{Z_iX} $(i = 1, 2)$ be the population co-variances between Z_i $(i = 1, 2)$ and auxiliary variable X. Then substituting the error terms defined above, the sample regression coefficients can be re-written as

$$\widehat{\beta_{Z_1X_1}} = \frac{s_{z_1x_1}}{\sigma^2_{X_1}} = \frac{s_{z_1x_1}}{\sigma^2_X} = \frac{\sigma_{Z_1X}(1+\delta_{S_{z_1x}})}{\sigma^2_X} = \beta_{Z_1X}(1+\delta_{S_{z_1x}}), \tag{VI.38}$$

and

$$\widehat{\beta_{Z_2X_2}} = \frac{s_{z_2x_2}}{\sigma^2_{X_2}} = \frac{s_{z_2x_2}}{\sigma^2_X} = \frac{\sigma_{Z_2X}(1+\delta_{S_{z_2x}})}{\sigma^2_X} = \beta_{Z_2X}(1+\delta_{S_{z_2x}}). \tag{VI.39}$$

Substituting the values of the regression coefficients as obtained in equations (VI.38,VI.39) along with error terms defined earlier, the estimator given by equation (VI.37) can be re-written as

$$\begin{aligned}\widehat{\mu_{AReg}} = &\left[\frac{\theta_2(\delta_{z_1}\mu_{Z_1}+\mu_{Z_1}) - \theta_1(\delta_{z_2}\mu_{Z_2}+\mu_{Z_2})}{\theta_2-\theta_1}\right] \\ &+ \frac{1}{2}\left[\beta_{Z_1X}(1+\delta_{z_1x})(\mu_X - \bar{x_1}) + \beta_{Z_2X}(1+\delta_{z_2x})(\mu_X - \bar{x_2})\right]\end{aligned} \tag{VI.40}$$

When we simplify the above expression, the estimator from equation (VI.37) is given by

$$\begin{aligned}\widehat{\mu_{AReg}} = &\left[\frac{\theta_2\mu_{Z_1} - \theta_1\mu_{Z_2}}{\theta_2-\theta_1}\right] + \left(\frac{\theta_2}{\theta_2-\theta_1}\right)\mu_{Z_1}\delta_{z_1} - \left(\frac{\theta_1}{\theta_2-\theta_1}\right)\mu_{Z_2}\delta_{z_2} \\ &- \frac{\mu_X}{2}\left[\beta_{Z_1X}\delta_{x_1}(1+\delta_{S_{z_1x}}) + \beta_{Z_2X}\delta_{x_2}(1+\delta_{S_{z_2x}})\right]\end{aligned} \tag{VI.41}$$

Substituting for μ_{Z_1} and μ_{Z_2} in the first term, on the right-hand-side, in equation (VI.41) we get

$$\widehat{\mu_{AReg}}-\mu_Y = \left(\frac{\theta_2}{\theta_2-\theta_1}\right)\mu_{Z_1}\delta_{z_1}-\left(\frac{\theta_1}{\theta_2-\theta_1}\right)\mu_{Z_2}\delta_{z_2}-\frac{\mu_X}{2}\left[\beta_{Z_1X}\delta_{x_1}(1+\delta_{S_{z_1x}})+\beta_{Z_2X}\delta_{x_2}(1+\delta_{S_{z_2x}})\right] \tag{VI.42}$$

Under the assumptions of bivariate normality (Sukhatme et al. (1970))[59]: $E(\delta_{X_1}) = 0$; $E(\delta_{X_2}) = 0$; $E(\delta_{Z_1}) = 0$; $E(\delta_{Z_2}) = 0$; $E(\delta^2_{X_1}) = (\frac{1-f_1}{n_1})C^2_{X_1}$; $E(\delta^2_{X_2}) = (\frac{1-f_2}{n_2})C^2_{X_2}$; $E(\delta^2_{Z_1}) = (\frac{1-f_1}{n_1})C^2_{Z_1}$; $E(\delta^2_{Z_2}) = (\frac{1-f_2}{n_2})C^2_{Z_2}$; $E(\delta_{X_1}\delta_{X_2}) = 0$, $E(\delta_{Z_1}\delta_{Z_2}) = 0$; $E(\delta_{Z_1}\delta_{X_2}) = 0$; $E(\delta_{Z_2}\delta_{X_1}) = 0$; $E(\delta_{Z_1}\delta_{X_1}) = (\frac{1-f_1}{n_1})C_{Z_1X_1}$; $E(\delta_{Z_2}\delta_{X_2}) = (\frac{1-f_2}{n_2})C_{Z_2X_2}$; $E(\delta_{x_1}\delta_{S_{x_1}}) = (\frac{1-f_1}{n_1})\frac{\mu_{03}}{\mu_{02}}\frac{1}{\mu_X}$; $E(\delta_{x_2}\delta_{S_{x_2}}) = (\frac{1-f_2}{n_2})\frac{\mu_{03}}{\mu_{02}}\frac{1}{\mu_X}$; $E(\delta_{x_1}\delta_{S_{z_1x}}) = (\frac{1-f_1}{n_1})\frac{\mu_{12}}{\mu_{11}}\frac{1}{\mu_X}$; $E(\delta_{x_2}\delta_{S_{z_2x}}) = (\frac{1-f_2}{n_2})\frac{\mu_{12}}{\mu_{11}}\frac{1}{\mu_X}$, where $f_1 = \frac{n_1}{N}$, $f_2 = \frac{n_2}{N}$, $C_{Z_1X_1} = \rho_{Z_1X_1}C_{Z_1}C_{X_1}$ and $C_{Z_2X_2} = \rho_{Z_2X_2}C_{Z_2}C_{X_2}$. Since the two sub-samples come from the same population, the coefficients of variation $C_{X_1} = C_{X_2} = C_X$, and the correlations $\rho_{Z_1X_1} = \rho_{Z_1X}$ and $\rho_{Z_2X_2} = \rho_{Z_2X}$.

Then the bias for the proposed regression estimator from equation (VI.37) is given by

$$\begin{aligned} Bias(\widehat{\mu_{AReg}}) &= E[\widehat{\mu_{AReg}} - \mu_Y] \\ &= -\frac{1}{2}\left[\beta_{Z_1X}\left(\frac{1-f_1}{n_1}\right)\frac{\mu_{12}}{\mu_{11}} + \beta_{Z_2X}\left(\frac{1-f_2}{n_2}\right)\frac{\mu_{12}}{\mu_{11}}\right], \end{aligned} \tag{VI.43}$$

where $\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}\left(Z_i-\mu_{Z_i}\right)^r\left(X_i-\mu_X\right)^s$ $(r,s = 0,1,2,3)$.

Squaring and taking the expectation of both sides of the equation (VI.42) and retaining terms up to second order, the MSE of the proposed regression estimator can be given by

$$MSE(\widehat{\mu_{AReg}}) = E[\widehat{\mu_{AReg}} - \mu_Y]^2$$
$$\approx \left(\frac{1-f_1}{n_1}\right)\left[\left(\frac{\theta_2}{\theta_2-\theta_1}\right)^2\sigma^2_{Z_1} + \frac{1}{4}\beta^2_{Z_1X}\sigma^2_{X_1} - \left(\frac{\theta_2}{\theta_2-\theta_1}\right)\beta_{Z_1X}\sigma_{Z_1X_1}\right]$$
$$+ \left(\frac{1-f_2}{n_2}\right)\left[\left(\frac{\theta_1}{\theta_2-\theta_1}\right)^2\sigma^2_{Z_2} + \frac{1}{4}\beta^2_{Z_2X}\sigma^2_{X_2} + \left(\frac{\theta_1}{\theta_2-\theta_1}\right)\beta_{Z_2X}\sigma_{Z_2X_2}\right], \tag{VI.44}$$

where $\theta_1 \neq \theta_2$.

Here we use the following expressions to evaluate the approximate MSE of the proposed regression estimator for the OET model.

$$\sigma^2_{Z_i} = \sigma^2_Y + W(1-A)[\sigma^2_T\sigma^2_Y + \sigma^2_T\mu^2_Y] + W\sigma^2_{S_i} + W\theta^2_i - W^2\theta^2_i, (i = 1, 2). \tag{VI.45}$$

Further, the population regression coefficients are given by

$$\beta_{Z_iX} = \frac{\sigma_{Z_iX}}{\sigma^2_x} = \frac{\sigma_{YX}}{\sigma^2_X} = \frac{\rho_{YX}\sigma_Y\sigma_X}{\sigma^2_X} = \frac{\rho_{YX}\sigma_Y}{\sigma_X}, i = 1, 2. \tag{VI.46}$$

Further, the following equations hold true for the population covariances, correlations and variances:

$$\sigma_{Z_iX_i} = \sigma_{YX_i} = \sigma_{YX} = \rho_{YX}\sigma_Y\sigma_X, \tag{VI.47}$$

$$\sigma^2_{X_i} = \sigma^2_X, \tag{VI.48}$$

and

$$\rho_{Z_iX} = \frac{\rho_{YX}\sigma_Y}{\sigma_{Z_i}} (i = 1, 2). \tag{VI.49}$$

Using equations (VI.47,VI.48,VI.49), we can re-write the MSE of the proposed regression estimator under the OET model as

$$\begin{aligned} MSE(\widehat{\mu_{AReg}}) \approx & \left(\frac{1-f_1}{n_1}\right)\left[\left(\frac{\theta_2}{\theta_2-\theta_1}\right)^2\sigma_{Z_1}^2 + \rho_{YX}^2\sigma_Y^2\left[\frac{1}{4} - \left(\frac{\theta_2}{\theta_2-\theta_1}\right)\right]\right] \\ & + \left(\frac{1-f_2}{n_2}\right)\left[\left(\frac{\theta_1}{\theta_2-\theta_1}\right)^2\sigma_{Z_21}^2 + \rho_{YX}^2\sigma_Y^2\left[\frac{1}{4} - \left(\frac{\theta_1}{\theta_2-\theta_1}\right)\right]\right], \end{aligned} \tag{VI.50}$$

where $\theta_1 \neq \theta_2$.

VI.5 Simulation Study

In this Chapter, we introduced an optional enhanced trust model (Figure VI.1) along with the ordinary estimator (equation VI.22), an additive ratio estimator (equation VI.30) and a regression estimator (equation VI.37) for the sensitive mean while simultaneously estimating the sensitivity level W. In this section, the results of a simulation study to evaluate and compare the performance of the ordinary RRT mean estimator $t_0 = \hat{\mu}_Y$, the additive ratio estimator $t_1 = \widehat{\mu_{ar}}$ and the linear regression estimator $t_2 = \widehat{\mu_{AReg}}$. We also evaluate the performance of an unbiased estimator of the sensitivity level denoted by $\hat{W}$.

We conducted a simulation study with 10000 iterations with samples of size $n = 500$ each from a finite population of size $N = 5000$. We generate this finite population

under two scenarios (i.e. a high and a moderate correlation between Y and X).

Scenario-1

We first generate the finite population from a bivariate normal distribution with means and covariances of (X, Y) given as follows:

$$\mu = \begin{bmatrix} 6 \\ 10 \end{bmatrix}, \Sigma = \begin{bmatrix} 8 & 10.1823 \\ 10.1823 & 16 \end{bmatrix}, \ \rho_{YX} = 0.9 \tag{VI.51}$$

i.e. we first generate a sample of 5000 units by using

$$\mu_X = 6, \mu_Y = 10, \sigma_X^2 = 8, \sigma_Y^2 = 16, \rho_{YX} = 0.9. \tag{VI.52}$$

We treat this sample as our finite population of size $N = 5000$. Then from this finite population, we repeatedly draw samples of size $n = 500$ using simple random sampling without replacement (SRSWOR). Each of these samples is split into two sub-samples of equal size $n_1 = n_2 = 250$ in every iteration. It must be noted that the real parameters for the 5000 units in the generated finite population are quite close to the assumed parameters, but are not exactly the same, and are given by

$$\mu_X = 6.002015, \mu_Y = 9.995325, \sigma_X^2 = 7.95262, \sigma_Y^2 = 15.80488, \rho_{YX} = 0.8986312. \tag{VI.53}$$

For this simulation study, we used the parameters for the finite population from equation (VI.53) and not the ones assumed to generate the finite population in equation (VI.52).

We assume the scrambling variables T and S_i $(i = 1, 2)$ to be normally distributed with known means and variances $(\mu_T = 1 = E(T), \mu_{S_1} = \theta_1 = 2, \mu_{S_2} = \theta_2 = 1, \sigma_T^2 = 0.5, \sigma_{S_i}^2 = Var(S_i) = 0.5\sigma_X^2; i = 1, 2)$. The empirical results have been averaged over 10000 iterations. The empirical MSE of the estimators $\hat{\mu}_j = t_j$ were computed by

$$MSE(t_j) = \frac{1}{10000} \sum_{i=1}^{10000} \left(\hat{\mu}_i - \mu_Y \right)^2. \tag{VI.54}$$

Here $\hat{\mu}_j = t_j$ can be $\hat{\mu}_Y = t_0$, $\widehat{\mu_{ar}} = t_1$ and $\widehat{\widehat{\mu_{AReg}}} = t_2$. Privacy level ∇ was computed for the OET model by taking the mean squared difference between the reported response Z and actual status for the sensitive study variable Y for only those respondents that consider the question sensitive and choose to scramble their responses. This mean squared difference in the report and the actual response is given by

$$\nabla_{oet} = E[Z - Y]^2. \tag{VI.55}$$

Note that although we propose various estimators, they were all proposed under the same model. Hence the value of ∇_{oet} corresponding to all three estimators would be the same since the privacy level is measured for the model being used and not the estimators proposed under that model. Following this, we also compute the unified measure of privacy and efficiency δ which was proposed by Gupta et al. (2018)[21] and is given by

$$\delta = \frac{MSE(t_j)}{\nabla_{oet}}, j = 0, 1, 2. \tag{VI.56}$$

Here we use MSE instead of the variances of the estimators to effectively evaluate

the efficiency of the biased estimators. The simulation results for Scenario-1 have been summarized in Table VI.1. The theoretical values have been listed in bold while the regular figures represent the empirical values of the various measures listed in the table.

From Table VI.1, we can make a number of observations. For any fixed value of A, we note that as W increases, MSE worsens for all estimators while the privacy level remains the same. For example, consider the cases where $A = 0.95$. When W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the ordinary RRT mean estimator, $t_0 = \hat{\mu}_Y$, takes values 0.3456, 0.3728, 0.4091, 0.4199, and 0.4302 respectively. Further, as W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the additive ratio estimator, $t_1 = \widehat{\mu_{ar}}$, takes values 0.3238, 0.3510, 0.3872, 0.3981, and 0.4083 respectively. Also, as W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the linear regression estimator, $t_2 = \widehat{\mu_{AReg}}$, takes values 0.3214, 0.3486, 0.3848, 0.3957, and 0.4059 respectively. Therefore, we can infer that the efficiency of all three estimators is the worst at $W = 1$. Hence optionality element of the proposed model leads to better efficiency for all three estimators introduced and used under this model. Furthermore, based on this example, we can also note that t_2 has the best efficiency followed closely by t_1 and both these estimators that utilize auxiliary variable information X are more efficient than the ordinary RRT mean estimator t_0 that does not utilize the auxiliary information. The theoretical privacy level is the same corresponding to all three estimators since the model is the same for all three estimators. Therefore, the lowest δ values are corresponding to the linear regression estimator t_2 which is slightly lower than the values of δ_{t1}.

Another observation that can be made using Table VI.1 is that for fixed values of

Table VI.1. Simulation results (for estimating sensitive mean μ_Y and sensitivity level W (Theoretical (**bold**) and Empirical Measures): $Iterations = 10000$, $N = 5000, n = 500$, $\mu_Y = 10$, $\rho_{YX} = 0.9$ for various levels of trust (A) and the sensitivity level (W).

A	W	$\hat{W}$	$Var(\hat{W})$	$MSE(t_0)$	$MSE(t_1)$	$MSE(t_2)$	∇_{oet}	$\delta(t_0)$	$\delta(t_1)$	$\delta(t_2)$
1	1	1.0007	0.1540	0.3814	0.3596	0.3581	6.4376	0.0592	0.0559	0.0556
			0.1584	**0.3751**	**0.3533**	**0.3509**	**6.4924**	**0.0578**	**0.0544**	**0.0540**
1	0.9	0.8957	0.1588	0.3916	0.3709	0.3696	6.4339	0.0609	0.0576	0.0574
			0.1570	**0.3704**	**0.3485**	**0.3461**	**6.4924**	**0.0570**	**0.0537**	**0.0533**
1	0.8	0.7966	0.1578	0.3878	0.3673	0.3657	6.4339	0.0603	0.0571	0.0568
			0.1552	**0.3650**	**0.3432**	**0.3408**	**6.4924**	**0.0562**	**0.0529**	**0.0525**
1	0.5	0.4994	0.1503	0.3697	0.3465	0.3454	6.4397	0.0574	0.0538	0.0536
			0.1474	**0.3453**	**0.3235**	**0.3211**	**6.4924**	**0.0532**	**0.0498**	**0.0495**
1	0.3	0.3005	0.1419	0.3512	0.3276	0.3262	6.4367	0.0546	0.0509	0.0507
			0.1402	**0.3291**	**0.3073**	**0.3049**	**6.4924**	**0.0507**	**0.0473**	**0.0470**
0.95	1	0.9935	0.1833	0.4529	0.4321	0.4307	9.3115	0.0486	0.0464	0.0462
			0.1816	**0.4302**	**0.4083**	**0.4059**	**9.3905**	**0.0458**	**0.0435**	**0.0432**
0.95	0.9	0.8905	0.1730	0.4299	0.4117	0.4090	9.3103	0.0462	0.0442	0.0439
			0.1779	**0.4199**	**0.3981**	**0.3957**	**9.3905**	**0.0447**	**0.0424**	**0.0421**
0.95	0.8	0.7920	0.1689	0.4180	0.3993	0.3966	9.3110	0.0449	0.0429	0.0426
			0.1737	**0.4091**	**0.3872**	**0.3848**	**9.3905**	**0.0436**	**0.0412**	**0.0410**
0.95	0.5	0.4938	0.1556	0.3848	0.3651	0.3624	9.3065	0.0413	0.0392	0.0389
			0.1590	**0.3728**	**0.3510**	**0.3486**	**9.3905**	**0.0397**	**0.0374**	**0.0371**
0.95	0.3	0.2954	0.1445	0.3582	0.3388	0.3358	9.3142	0.0385	0.0364	0.0361
			0.1472	**0.3456**	**0.3238**	**0.3214**	**9.3905**	**0.0368**	**0.0345**	**0.0342**
0.9	1	0.9955	0.2071	0.5154	0.4923	0.4913	12.1819	0.0423	0.0404	0.0403
			0.2047	**0.4852**	**0.4634**	**0.4610**	**12.2886**	**0.0395**	**0.0377**	**0.0375**
0.9	0.9	0.8900	0.1932	0.4790	0.4612	0.4582	12.2000	0.0393	0.0378	0.0376
			0.1987	**0.4695**	**0.4476**	**0.4452**	**12.2886**	**0.0382**	**0.0364**	**0.0362**
0.9	0.8	0.7908	0.1867	0.4608	0.4428	0.4397	12.2027	0.0378	0.0363	0.0360
			0.1923	**0.4531**	**0.4313**	**0.4289**	**12.2886**	**0.0369**	**0.0351**	**0.0349**
0.9	0.5	0.4928	0.1674	0.4134	0.3939	0.3912	12.1821	0.0339	0.0323	0.0321
			0.1706	**0.4004**	**0.3785**	**0.3761**	**12.2886**	**0.0326**	**0.0308**	**0.0306**
0.9	0.3	0.2961	0.1518	0.3764	0.3570	0.3540	12.1861	0.0309	0.0293	0.0291
			0.1541	**0.3622**	**0.3403**	**0.3379**	**12.2886**	**0.0295**	**0.0277**	**0.0275**
0.85	1	0.9938	0.2324	0.5777	0.5560	0.5545	15.0673	0.0383	0.0369	0.0368
			0.2279	**0.5403**	**0.5185**	**0.5161**	**15.1866**	**0.0356**	**0.0341**	**0.0340**
0.85	0.9	0.8904	0.2161	0.5335	0.5165	0.5135	15.0874	0.0354	0.0342	0.0340
			0.2196	**0.5190**	**0.4972**	**0.4948**	**15.1866**	**0.0342**	**0.0327**	**0.0326**
0.85	0.8	0.7911	0.2073	0.5097	0.4923	0.4892	15.0931	0.0338	0.0326	0.0324
			0.2108	**0.4972**	**0.4753**	**0.4729**	**15.1866**	**0.0327**	**0.0313**	**0.0311**
0.85	0.5	0.4941	0.1796	0.4433	0.4239	0.4213	15.0608	0.0294	0.0281	0.0280
			0.1822	**0.4279**	**0.4060**	**0.4037**	**15.1866**	**0.0282**	**0.0267**	**0.0266**
0.85	0.3	0.2958	0.1592	0.3940	0.3749	0.3718	15.0707	0.0261	0.0249	0.0247
			0.1611	**0.3787**	**0.3568**	**0.3544**	**15.1866**	**0.0249**	**0.0235**	**0.0233**
0.8	1	0.9956	0.2544	0.6316	0.6101	0.6083	17.9504	0.0352	0.0340	0.0339
			0.2511	**0.5954**	**0.5735**	**0.5711**	**18.0847**	**0.0329**	**0.0317**	**0.0316**
0.8	0.9	0.8913	0.2365	0.5863	0.5675	0.5651	17.9817	0.0326	0.0316	0.0314
			0.2404	**0.5686**	**0.5467**	**0.5443**	**18.0847**	**0.0314**	**0.0302**	**0.0301**
0.8	0.8	0.7916	0.2250	0.5559	0.5372	0.5346	17.9859	0.0309	0.0299	0.0297
			0.2294	**0.5412**	**0.5194**	**0.5170**	**18.0847**	**0.0299**	**0.0287**	**0.0286**
0.8	0.5	0.4935	0.1898	0.4704	0.4503	0.4480	17.9704	0.0262	0.0251	0.0249
			0.1938	**0.4554**	**0.4336**	**0.4312**	**18.0847**	**0.0252**	**0.0240**	**0.0238**
0.8	0.3	0.2950	0.1656	0.4108	0.3912	0.3882	17.9641	0.0229	0.0218	0.0216
			0.1680	**0.3952**	**0.3733**	**0.3710**	**18.0847**	**0.0219**	**0.0206**	**0.0205**

W, say $W = 0.9$, The efficiency of all estimators worsen as the level of trust A drops. This is expected as fewer people trust the model, they would choose to go with a more enhanced scrambling option. However, the gain in privacy due to enhanced scrambling of responses more than compensates for the drop in efficiency when we consider the unified measure of privacy and efficiency to assess the overall estimator performance. For example, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_0} takes values 0.3704, 0.4199, 0.4695, 0.5190, and 0.5686 respectively. Similarly, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_1} takes values 0.3485, 0.3981, 0.4476, 0.4972, and 0.5467 respectively. Also, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_2} takes values 0.3461, 0.3957, 0.4452, 0.4948, and 0.5443 respectively. Although t_1 and t_2 have reasonably similar overall performances, t_2 appears to be slightly better.

Scenario-2

As shown in Scenario-1, we now generate the finite population of size $N = 5000$ from a bivariate normal distribution with means and covariances of (X, Y) given as follows:

$$\mu = \begin{bmatrix} 6 \\ 10 \end{bmatrix}, \Sigma = \begin{bmatrix} 8 & 6.788225 \\ 6.788225 & 16 \end{bmatrix}, \ \rho_{YX} = 0.6 \tag{VI.57}$$

We have

$$\mu_X = 6, \mu_Y = 10, \sigma_X^2 = 8, \sigma_Y^2 = 16, \rho_{YX} = 0.6. \tag{VI.58}$$

Then from this finite population, we repeatedly draw samples of size $n = 500$ using

simple random sampling without replacement (SRSWOR). It must be noted that the real parameters for the 5000 units in the generated finite population are quite close to these assumed parameters, but are not exactly the same, and are given by

$$\mu_X = 6.006096, \mu_Y = 9.993591, \sigma_X^2 = 8.010827, \sigma_Y^2 = 15.79687, \rho_{YX} = 0.5967508. \tag{VI.59}$$

For this simulation study, we used the parameters for the finite population (VI.59) and not the ones assumed to generate the finite population (VI.58). The only difference between Scenario-1 and Scenario-2 is that we now use a considerably lower value for ρ_{YX} to study how the performance of the three estimators gets affected. We again assume the scrambling variables T and S_i $(i = 1, 2)$ to be normally distributed with known means and variances $(\mu_T = 1 = E(T), \mu_{S_1} = \theta_1 = 2, \mu_{S_2} = \theta_2 = 1, \sigma_T^2 = 0.5, \sigma_{S_i}^2 = Var(S_i) = 0.5\sigma_X^2; i = 1, 2)$. The empirical results have again been averaged over 10000 iterations. The simulation results for Scenario-2 have been summarized in Table VI.2. The theoretical values have been listed in bold while the regular figures represent the empirical values of the various measures listed in the table.

From Table VI.2, we can make a number of observations. For any fixed value of A, we note that as W increases, MSE worsens for all estimators while the privacy level remains the same. For example, consider the cases where $A = 0.95$. When W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the ordinary RRT mean estimator, $t_0 = \hat{\mu}_Y$, takes values 0.3456, 0.3729, 0.4093, 0.4203, and 0.4306 respectively. Further, as W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the additive ratio estimator, $t_1 = \widehat{\mu_{ar}}$, takes values 0.3453, 0.3726, 0.4090, 0.4199, and 0.4302 respectively. Also, as W varies between 0.3, 0.5, 0.8, 0.9 and 1, the

Table VI.2. Simulation results (for estimating sensitive mean μ_Y and sensitivity level W: $Iterations = 10000$, $N = 5000, n = 500$, $\mu_Y = 9$, $\rho_{YX} = 0.6$ for various levels of trust (A) and the sensitivity level (W).

A	W	$\hat{W}$	$Var(\hat{W})$	$MSE(t_0)$	$MSE(t_1)$	$MSE(t_2)$	∇_{oet}	$\delta(t_0)$	$\delta(t_1)$	$\delta(t_2)$
1	1	1.0007	0.1543	0.3818	0.3801	0.3714	6.4666	0.0590	0.0588	0.0574
			0.1585	**0.3755**	**0.3752**	**0.3648**	**6.5216**	**0.0576**	**0.0575**	**0.0559**
1	0.9	0.8955	0.1593	0.3936	0.3948	0.3851	6.4628	0.0609	0.0611	0.0596
			0.1571	**0.3707**	**0.3704**	**0.3600**	**6.5216**	**0.0568**	**0.0568**	**0.0552**
1	0.8	0.7964	0.1583	0.3898	0.3912	0.3813	6.4629	0.0603	0.0605	0.0590
			0.1553	**0.3653**	**0.3650**	**0.3546**	**6.5216**	**0.0560**	**0.0560**	**0.0544**
1	0.5	0.4992	0.1506	0.3708	0.3695	0.3607	6.4687	0.0573	0.0571	0.0558
			0.1475	**0.3454**	**0.3451**	**0.3347**	**6.5216**	**0.0530**	**0.0529**	**0.0513**
1	0.3	0.3003	0.1422	0.3521	0.3505	0.3416	6.4657	0.0545	0.0542	0.0528
			0.1402	**0.3291**	**0.3288**	**0.3184**	**6.5216**	**0.0505**	**0.0504**	**0.0488**
0.95	1	0.9931	0.1839	0.4549	0.4555	0.4460	9.3376	0.0487	0.0488	0.0478
			0.1817	**0.4306**	**0.4302**	**0.4199**	**9.4186**	**0.0457**	**0.0457**	**0.0446**
0.95	0.9	0.8901	0.1734	0.4313	0.4355	0.4239	9.3364	0.0462	0.0466	0.0454
			0.1780	**0.4203**	**0.4199**	**0.4096**	**9.4186**	**0.0446**	**0.0446**	**0.0435**
0.95	0.8	0.7916	0.1691	0.4190	0.4227	0.4112	9.3369	0.0449	0.0453	0.0440
			0.1739	**0.4093**	**0.4090**	**0.3986**	**9.4186**	**0.0435**	**0.0434**	**0.0423**
0.95	0.5	0.4934	0.1559	0.3861	0.3895	0.3781	9.3333	0.0414	0.0417	0.0405
			0.1590	**0.3729**	**0.3726**	**0.3623**	**9.4186**	**0.0396**	**0.0396**	**0.0385**
0.95	0.3	0.2949	0.1448	0.3595	0.3630	0.3513	9.3421	0.0385	0.0389	0.0376
			0.1472	**0.3456**	**0.3453**	**0.3350**	**9.4186**	**0.0367**	**0.0367**	**0.0356**
0.9	1	0.9952	0.2076	0.5172	0.5161	0.5074	12.2061	0.0424	0.0423	0.0416
			0.2049	**0.4856**	**0.4853**	**0.4749**	**12.3156**	**0.0394**	**0.0394**	**0.0386**
0.9	0.9	0.8896	0.1937	0.4807	0.4852	0.4733	12.2247	0.0393	0.0397	0.0387
			0.1988	**0.4698**	**0.4695**	**0.4591**	**12.3156**	**0.0381**	**0.0381**	**0.0373**
0.9	0.8	0.7904	0.1871	0.4621	0.4661	0.4543	12.2274	0.0378	0.0381	0.0372
			0.1924	**0.4534**	**0.4531**	**0.4427**	**12.3156**	**0.0368**	**0.0368**	**0.0359**
0.9	0.5	0.4923	0.1677	0.4148	0.4183	0.4069	12.2084	0.0340	0.0343	0.0333
			0.1706	**0.4005**	**0.4002**	**0.3898**	**12.3156**	**0.0325**	**0.0325**	**0.0316**
0.9	0.3	0.2956	0.1521	0.3777	0.3809	0.3694	12.2127	0.0309	0.0312	0.0302
			0.1541	**0.3622**	**0.3619**	**0.3515**	**12.3156**	**0.0294**	**0.0294**	**0.0285**
0.85	1	0.9935	0.2330	0.5795	0.5797	0.5702	15.0883	0.0384	0.0384	0.0378
			0.2281	**0.5406**	**0.5403**	**0.5300**	**15.2126**	**0.0355**	**0.0355**	**0.0348**
0.85	0.9	0.8901	0.2166	0.5353	0.5404	0.5283	15.1098	0.0354	0.0358	0.0350
			0.2197	**0.5193**	**0.5190**	**0.5086**	**15.2126**	**0.0341**	**0.0341**	**0.0334**
0.85	0.8	0.7908	0.2078	0.5113	0.5159	0.5039	15.1155	0.0338	0.0341	0.0333
			0.2109	**0.4974**	**0.4971**	**0.4867**	**15.2126**	**0.0327**	**0.0327**	**0.0320**
0.85	0.5	0.4935	0.1798	0.4446	0.4481	0.4369	15.0858	0.0295	0.0297	0.0290
			0.1822	**0.4280**	**0.4277**	**0.4173**	**15.2126**	**0.0281**	**0.0281**	**0.0274**
0.85	0.3	0.2952	0.1594	0.3951	0.3984	0.3869	15.0940	0.0262	0.0264	0.0256
			0.1611	**0.3787**	**0.3784**	**0.3680**	**15.2126**	**0.0249**	**0.0249**	**0.0242**
0.8	1	0.9953	0.2548	0.6331	0.6324	0.6232	17.9700	0.0352	0.0352	0.0347
			0.2513	**0.5957**	**0.5954**	**0.5850**	**18.1096**	**0.0329**	**0.0329**	**0.0323**
0.8	0.9	0.8909	0.2372	0.5881	0.5918	0.5806	18.0005	0.0327	0.0329	0.0323
			0.2406	**0.5689**	**0.5686**	**0.5582**	**18.1096**	**0.0314**	**0.0314**	**0.0308**
0.8	0.8	0.7912	0.2256	0.5575	0.5610	0.5497	18.0043	0.0310	0.0312	0.0305
			0.2295	**0.5414**	**0.5411**	**0.5308**	**18.1096**	**0.0299**	**0.0299**	**0.0293**
0.8	0.5	0.4929	0.1902	0.4718	0.4750	0.4640	17.9913	0.0262	0.0264	0.0258
			0.1938	**0.4555**	**0.4552**	**0.4448**	**18.1096**	**0.0252**	**0.0251**	**0.0246**
0.8	0.3	0.2944	0.1659	0.4120	0.4152	0.4038	17.9821	0.0229	0.0231	0.0225
			0.1680	**0.3952**	**0.3949**	**0.3845**	**18.1096**	**0.0218**	**0.0218**	**0.0212**

theoretical MSE for the regression estimator, $t_2 = \widehat{\mu_{AReg}}$, takes values 0.3350, 0.3623, 0.3986, 0.4096, and 0.4199 respectively. Therefore, we can infer that the efficiency of all three estimators is the worst at $W = 1$. Hence optionality element of the proposed model ensures optimal efficiency of any of the three estimators introduced used under this model. Furthermore, based on this example, we can also note that t_2 has the best efficiency followed by t_1. Both these estimators which utilize auxiliary variable information X are more efficient than the ordinary RRT mean estimator t_0 that does not utilize the auxiliary information. The theoretical privacy level is the same corresponding to all three estimators since the model is the same for all three estimators. Therefore, the lowest δ values are corresponding to the regression estimator t_2 which is slightly lower than the values of δ_{t1}.

Another observation that can be made using Table VI.2 is that for fixed values of W, say $W = 0.9$, The efficiency of all estimators worsen as the level of trust A drops. This is expected as fewer people trust the model, and the use of the choice to go with the enhanced scrambling option will rise. However, the gain in privacy due to enhanced scrambling of responses more than compensates for the drop in efficiency when we consider the unified measure of privacy and efficiency to assess the overall estimator performance. For example, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_0} takes values 0.0568, 0.0446, 0.0381, 0.0341, and 0.0314 respectively. Similarly, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_1} takes values 0.0568, 0.0446, 0.0381, 0.0341, and 0.0314 respectively. Also, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_2} takes values 0.0552, 0.0435, 0.0373, 0.0334, and 0.0308 respectively. Hence the best value of δ can be obtained even with considerably lower values A using any of the three estimators. Considering the values of the unified measure δ, t_2 has the best overall performance, followed by t_1 and then t_0.

On comparing the results from Table VI.1 and Table VI.2 we can note the MSE for three estimators worsen when ρ_{YX} lowers. However, the MSE for t_1 and t_2 worsen considerably more than that of t_0 as the ordinary RRT mean estimator t_0 does not rely on the auxiliary variable X. Moreover, we also note that the performance of t_1 worsens so much that there is hardly any difference between the MSEs of t_0 and t_1. However, t_2 still performs the best of the three estimators. For instance, let us consider the case when $A = 0.9$ and $W = 0.8$. When $\rho_{YX} = 0.9$, the theoretical MSE for t_0, t_1 and t_2 are 0.4531, 0.4313 and 0.4289 respectively. However, when $\rho_{YX} = 0.6$, the theoretical MSE for t_0, t_1 and t_2 are 0.4532, 0.4531, and 0.44277 respectively. Further, when $\rho_{YX} = 0.9$, the theoretical δ for t_0, t_1 and t_2 are 0.0369, 0.0351, and 0.0349. However, when $\rho_{YX} = 0.6$, the theoretical δ for t_0, t_1 and t_2 are 0.0368, 0.0368, and 0.0359 respectively. Hence, we can see that the overall performance of the proposed additive ratio estimator (t_1) may not offer anything additional when compared to the overall performance of ordinary RRT mean estimator (t_0) when the correlation between the sensitive study variable Y and the non-sensitive auxiliary variable is weak. Moreover, we note that despite the weaker correlation, the proposed regression estimator (t_2) still outperforms the other two proposed estimators even when the correlation between the sensitive study variable Y and the non-sensitive auxiliary variable is moderate or weak.

From both Table VI.1 and Table VI.2, we can also note that the estimator W performs reasonably well in estimating the value of the sensitivity level W. Furthermore, the empirical and the theoretical values for all the measures listed in both tables are a reasonable match which bolsters our theoretical findings.

VI.6 Concluding Chapter Remarks

In this chapter, we proposed an Optional Enhanced Trust model that offers the option for additional scrambling variables for further response protection of those respondents who are not comfortable using the additive noise only. This method allows us to account for the respondents' lack of trust in the quantitative RRT model being used in a sensitive question survey. Although Lovig et al. (2021)[41] proposed a method to account for the lack of trust in the binary RRT models, such work had not been accomplished in the area of quantitative RRT. We introduce estimators for the sensitive mean that can be used in the absence as well in the presence of a non-sensitive auxiliary variable. The primary finding of this chapter was that when the auxiliary information is available for every population unit, using the linear regression estimator introduced in this chapter, would have the best performance in terms of efficiency as well as in terms of the unified measure of privacy and efficiency. When the correlation between the sensitive study variable Y and the non-sensitive auxiliary variable is strong, the performance of the proposed additive ratio estimator is only slightly worse than the regression estimator and would be a more appropriate choice when the relationship of sensitive study variable Y and auxiliary variable X goes through the origin. However, if the correlation between the sensitive study variable Y and the non-sensitive auxiliary variable is moderate or weak, the performance of the proposed additive ratio estimator can be as bad as that of the ordinary RRT mean estimator and the overall performance might even get slightly worse than that of the ordinary RRT mean estimator.

Therefore, utilizing non-sensitive auxiliary information, whenever possible, can considerably improve the sensitive mean estimation while simultaneously estimating

the sensitivity level of the survey question. Various RRT researchers have tried to propose generalizations of the estimators based on auxiliary variables. We attempt to do something similar in the next Chapter. We will discuss a special case of the work discussed in this chapter in Chapter VII.

Chapter VII: Generalized Mean Estimator under the Optional Enhanced Trust Model

In Chapter VI, we introduced an optional enhanced trust model FigureVI.1 which helps account for respondents' lack of trust in Warner's additive model (1971)[63]. Using the method discussed in Chapter VI, one can simultaneously estimate the mean of the sensitive variable (μ_Y) and the sensitivity level W. However, if one simply wishes to estimate the sensitive mean μ_Y, in the presence of unknown W, then the work can be simplified significantly.

In this chapter[1], we will discuss a special case of the optional enhanced trust model Figure VI.1. It should be noted that a split-sample technique was used for the work discussed in Chapter VI because we were simultaneously estimating both μ_Y and W. However, if we only need to estimate μ_Y, we do not need the Split-Sample approach.

[1]A portion of this chapter is based on an Accepted Manuscript of an article published by Taylor & Francis in Journal of Communications in Statistics - Simulation and Computation on 3 June 2022, available online: https://www.tandfonline.com/doi/abs/10.1080/03610918.2022.2082477

VII.1 Mean Estimation using the Optional Enhanced Trust (OET) Model

Let Y be the sensitive study variable and Z be the reported response. Let S and T be the scrambling variables with means μ_S, μ_T and variances σ_S^2 and σ_T^2 respectively. Moreover, let W represent the sensitivity level of the survey question meaning a proportion $(1-W)$ of the respondents do not consider the question sensitive and hence will provide an unscrambled response. Let A represent the proportion of respondents that trust the Warner's Additive model(1971)[63] and hence do not need additional noise. Here, Y, T and S are assumed to be mutually independent.

As stated in the previous chapter, the Optional Enhanced Trust model mitigates the effect of respondents' lack of trust by allowing respondents to simply report their true responses if they do not find the survey question sensitive. However, if they do find the question sensitive they have the option to scramble their response using either of the two scrambling techniques available to them (i.e. additive noise or an additive and a multiplicative noise) based on whether or not they trust the additive model. Under this model, the reported response is given as shown in equation (VI.16).

$$Z = \begin{cases} Y & with\ probability\ 1-W \\ Y+S & with\ probability\ WA \\ TY+S & with\ probability\ W(1-A). \end{cases} \tag{VII.1}$$

The alternative scrambling technique is based on the Diana and Perri (2011)[10] model. It helps in improving the respondent privacy level which can help in lowering the level of untruthfulness. The surveyor would still just have information about

the reported response (Z) but has no idea about whether the respondent reported their true response (Y), or a scrambled response based on one of the two scrambling schemes available to them as shown in equation (VII.1).

To estimate the sensitive mean μ_Y, we assume that $E[S] = 0$ and $E[T] = 1$. Then the expected value of the reported response Z would be given by

$$E[Z] = (1 - W)E[Y] + (WA)E[Y + S] + W(1 - A)E[TY + S] = \mu_Y. \qquad \text{(VII.2)}$$

Note that $E(Z)$ is independent of W and A. Therefore, the sensitive mean μ_Y can be estimated by

$$\hat{\mu}_Y = \overline{Z}, \qquad \text{(VII.3)}$$

in the presence of A and W without requiring the estimation of A and W.

The variance of this unbiased estimator in equation (VII.3) is given by

$$Var(\hat{\mu}_Y) = \frac{1}{n}[W(1 - A)(\sigma_T^2\sigma_Y^2 + \sigma_T^2\mu_Y^2) + \sigma_Y^2 + W\sigma_S^2]. \qquad \text{(VII.4)}$$

When W and A are unknown, this variance may be estimated by

$$\widehat{Var}(\hat{\mu}_Y) = \frac{\widehat{Var}(Z)}{n} = \frac{s_Z^2}{n}, \qquad \text{(VII.5)}$$

where s_Z^2 is the sample variance of the reported responses.

Based on equations (VII.3) and (VII.5), both μ_Y and $Var(\hat{\mu}_Y)$ can be estimated in the presence of A and W without needing to estimate A or W.

We mention again that Gupta et al. (2018)[21] established that optionality does not

compromise the privacy level since privacy is not a consideration for the respondents who do not find the survey question sensitive. Thus the privacy level of the Optional Enhanced Trust (OET) model is given by

$$\nabla_{oet} = AE[S^2] + (1-A)E[TY + S - Y]^2 = (1-A)(\sigma_T^2\sigma_Y^2 + \sigma_T^2\mu_Y^2) + \sigma_S^2. \quad \text{(VII.6)}$$

The unified measure of privacy and efficiency, for the ordinary RRT mean estimator $\hat{\mu}_Y$, as defined by Gupta et al. (2018)[21] is given by

$$\delta = \frac{Var(\hat{\mu}_Y)}{\nabla_{oet}}, \quad \text{(VII.7)}$$

where $Var(\hat{\mu}_Y)$ and ∇_{oet} can be obtained from equations (VII.5,VII.6).

We can make two important observations about the effects of A and W on the OET model. We can see from equation (VII.7) that as A decreases, i.e. as fewer respondents trust Warner's Additive Model (1971)[63], the unified measure for the model improves. Although the enhanced scrambling option ($TY + S$) lowers model efficiency, it still helps in improving the respondent privacy level. Further, we note that when fewer respondents find the question sensitive (i.e. W decreases), the unified measure improves as more individuals might be likely to report responses Y unscrambled, thereby improving the model efficiency while the privacy level remains constant as established by Gupta et al. (2018)[21].

In the following sections, we introduce a ratio and a regression estimator to improve the efficiency of the estimator given by equation VII.3. We also introduce a generalized estimator which utilizes information obtained from one auxiliary variable. We do so without estimating W, and hence without using the Split-Sample approach.

VII.2 Ratio Estimator of the Mean for the OET Model

In Subsection II.5.2, we presented a discussion on various types of sensitive mean estimators that can be used in the presence of one non-sensitive auxiliary variable which is highly correlated with the sensitive study variable. The types of estimators introduced were ratio, regression and generalized estimators in the presence of one non-sensitive auxiliary variable under the non-optional or optional versions of the Warner's additive model (1971)[63] or the Pollock and Beck model (1976)[50] where the reported response Z is given by

$$Z = Y + S. \tag{VII.8}$$

S is a random scrambling variable, with mean $\mu_S = 0$ and variance σ_S^2.

In this section, we will introduce a ratio estimator for the mean of the sensitive variable Y when complete information is available for the non-sensitive auxiliary variable X which is highly correlated with Y.

In Subsection II.5.2, we introduced the ratio estimator proposed by Sousa et al. (2010)[58] under the non-optional version of Warner's additive model (1971)[63]. Following this work, we propose a ratio estimator for the sensitive mean μ_Y under the Optional Enhanced Trust (OET) model (Figure VI.1).

Let, Y be the sensitive study variable that can not be observed directly and X be the non-sensitive auxiliary variable (positively correlated with Y). Also, let S be scrambling variable (independent of both X and Y) with mean $\mu_S = 0$ and variance σ_S^2. Let μ_X be the known true population mean and σ_X^2 be the known variance of the

non-sensitive auxiliary variable X. Let μ_Y be the unknown true population mean and σ_Y^2 be the unknown variance of the sensitive study variable Y. Then the proposed ratio estimator for the optional enhanced trust model can be given by

$$\widehat{\mu_r} = \bar{z}\frac{\mu_X}{\bar{x}} = \hat{\mu}_Y\frac{\mu_X}{\bar{x}}, \tag{VII.9}$$

where $\bar{z}$ is the sample mean of reported responses and $\bar{x}$ is the sample mean of an auxiliary variable and $\hat{\mu}_Y$ is the ordinary RRT mean estimator given by equation (VII.3). Note that explicit knowledge of W is not needed from the estimator $\hat{\mu}_Y$ (equation VII.3).

A large sample size is assumed such that $|\delta_z| < 1$ and $|\delta_x| < 1$ where these error terms are defined by

$$\delta_x = \tfrac{\bar{x}-\mu_X}{\mu_X} \text{ and } \delta_z = \tfrac{\bar{z}-\mu_Z}{\mu_Z}$$

Using the error terms defined above, the re-written proposed ratio estimator is given by

$$\widehat{\mu_r} = \mu_Z(1+\delta_z)(1+\delta_x)^{-1} = \mu_Y(1+\delta_z)(1+\delta_x)^{-1} \tag{VII.10}$$

Then subtracting μ_Y from both sides in equation (VII.10), and using second order Taylor's approximation we get

$$\widehat{\mu_r} - \mu_Y \approx \mu_Y[\delta_z - \delta_x + \delta_x^2 - \delta_z\delta_x]. \tag{VII.11}$$

Under the assumption of bivariate normality (Sukhatme et al. (1970)[59]):

$$E(\delta_x) = 0; \quad E(\delta_z) = 0; \quad E(\delta_x^2) = \tfrac{1-f}{n}C_X^2; \quad E(\delta_z^2) = \tfrac{1-f}{n}C_Z^2; \quad E(\delta_z\delta_x) = \tfrac{1-f}{n}C_{ZX},$$

where $f = \frac{n}{N}$, $C_{ZX} = \rho_{ZX} C_Z C_X$ and C_Z and C_X are the coefficients of variation of Z and X respectively.

Recognizing that $\mu_Z = \mu_Y$, the bias for the ratio estimator $\widehat{\mu_r}$, correct up to the first order of approximation, is given by

$$Bias(\widehat{\mu_r}) = E[\widehat{\mu_r} - \mu_Y] \approx \mu_Y \left(\frac{1-f}{n}\right)\left[C_X^2 - \rho_{ZX} C_Z C_X\right], \tag{VII.12}$$

where $\rho_{ZX} = \rho_{YX}\frac{\sigma_Y}{\sigma_Z}$.

Using equation (VII.10), the mean squared error (MSE) of the ratio estimator, correct up to the first order of approximation, is given by

$$MSE(\widehat{\mu_r}) = E[\widehat{\mu_r} - \mu_Y]^2 \approx \mu_Y^2 \left(\frac{1-f}{n}\right)\left[C_Z^2 + C_X^2 - 2\rho_{ZX} C_Z C_X\right], \tag{VII.13}$$

where $\rho_{ZX} = \rho_{YX}\frac{\sigma_Y}{\sigma_Z}$.

VII.3 Regression Estimator of the Mean for the OET Model

In Subsection II.5.2, we introduced the regression estimator proposed by Gupta et al. (2012)[23] which is used to estimate the mean of the sensitive study variable Y when information is available on a non-sensitive but highly correlated auxiliary variable, for every unit in the population. They presented this work under the non-optional Pollock and Beck (1976)[50] model. Following this work, in this section, we will present

a regression estimator, under the optional enhanced trust (OET) model given by equation (VII.1).

Let Y be the sensitive study variable and X be the non-sensitive auxiliary variable with has a strong positive correlation with Y. Also, let S be scrambling variable (independent of both X and Y) with mean $\mu_S = 0$ and variance σ_S^2. Let μ_X be the known true population mean and σ_X^2 be the known variance of the non-sensitive auxiliary variable X. Let μ_Y be the unknown true population mean and σ_Y^2 be the unknown variance of the sensitive study variable Y. Then the regression estimator for the sensitive mean μ_Y is given by

$$\widehat{\mu_{reg}} = \bar{z} + \widehat{\beta_{ZX}}(\mu_X - \bar{x}) = \hat{\mu}_Y + \widehat{\beta_{ZX}}(\mu_X - \bar{x}), \tag{VII.14}$$

where $\bar{z}$ is the sample mean of reported responses and $\bar{x}$ is the sample mean of an auxiliary variable and $\hat{\mu}_Y$ is the ordinary RRT mean estimator given by equation (VII.3). $\widehat{\beta_{ZX}}$ is the sample estimate of the regression coefficient between the reported response Z and the auxiliary variable X. The true value of this regression coefficient is given by

$$\beta_{ZX} = \frac{\sigma_{ZX}}{\sigma_X^2} = \rho_{YX}\frac{\sigma_Y}{\sigma_X}, \tag{VII.15}$$

where $\rho_Y X$ is the correlation coefficient between the sensitive study variable Y and the non-sensitive auxiliary variable X. Further, a large sample size is assumed such that $|\delta_z| < 1$ and $|\delta_x| < 1$ where we define the following error terms:

$$\delta_x = \frac{\bar{x}-\mu_X}{\mu_X}; \qquad \delta_z = \frac{\bar{z}-\mu_Z}{\mu_Z}$$
$$\delta_{S_x^2} = \frac{s_x^2-\sigma_X^2}{\sigma_X^2}; \qquad \delta_{S_{zx}} = \frac{s_{zx}-\sigma_{ZX}}{\sigma_{ZX}}$$

Using the error terms defined above, the re-written proposed regression estimator, from equation (VII.14), is given by

$$\widehat{\mu_{reg}} = (\mu_Y + \delta_z\mu_Y) - \left[\beta_{ZX}\mu_X(\delta_x + \delta_x\delta_{S_{zx}})\right] \tag{VII.16}$$

Then subtracting μ_Y from both sides in equation (VII.16) we get

$$\widehat{\mu_{reg}} - \mu_Y = \delta_z\mu_Y - \beta_{ZX}\mu_X\left[(\delta_x + \delta_x\delta_{S_{zx}})\right] \tag{VII.17}$$

Under the assumption of bivariate normality (Sukhatme et al. (1970)[59]):

$$E(\delta_x) = 0; \quad E(\delta_z) = 0; \quad E(\delta_x^2) = \tfrac{1-f}{n}C_X^2; \quad E(\delta_z^2) = \tfrac{1-f}{n}C_Z^2; \quad E(\delta_z\delta_x) = \tfrac{1-f}{n}C_{ZX}$$
$$E(\delta_x\delta_{S_x^2}) = \tfrac{1-f}{n}\tfrac{1}{\mu_X}\tfrac{\mu_{03}}{\mu_{02}}; \quad E(\delta_x\delta_{S_{zx}}) = \tfrac{1-f}{n}\tfrac{1}{\mu_X}\tfrac{\mu_{12}}{\mu_{11}},$$

where $f = \frac{n}{N}$, $C_{ZX} = \rho_{ZX}C_ZC_X$ and C_Z and C_X are the coefficients of variation of Z and X respectively. Also $\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}\left(Z_i - \mu_{Z_i}\right)^r\left(X_i - \mu_X\right)^s$ $(r, s = 0, 1, 2, 3)$. Recognizing that $\mu_Z = \mu_Y$, the bias for the proposed regression estimator $\widehat{\mu_{reg}}$, correct up to the first order of approximation, is given by

$$Bias(\widehat{\mu_{reg}}) = E[\widehat{\mu}_r - \mu_Y] \approx \beta_{ZX}\left(\frac{1-f}{n}\right)\frac{\mu_{12}}{\mu_{11}}. \tag{VII.18}$$

Using equation (VII.16), the mean squared error (MSE) of the proposed regression estimator $\widehat{\mu_{reg}}$, correct up to first order approximation, is given by

$$MSE(\widehat{\mu_{reg}}) = E[\widehat{\mu_{reg}} - \mu_Y]^2 \approx \left(\frac{1-f}{n}\right)\left[\mu_Y^2C_Z^2 + \beta_{ZX}^2\sigma_X^2 - 2\beta_{ZX}\sigma_{ZX}\right], \tag{VII.19}$$

where σ_{ZX} is the population covariance between Z and X.

VII.4 A Generalized Estimator of the Mean for the OET Model

Several RRT researchers have proposed generalized estimators that combine elements of both the ratio and the regression estimators. A couple of examples of such works were discussed in Section II.5.

Khalil et al. (2018)[32] also proposed a generalized RRT estimator for the sensitive study variable Y. Although they did their work both in the presence and in the absence of measurement errors, under the non-optional Pollock and Beck (1976)[50] model, we consider the case where there is no measurement error. Following from their work, we propose a generalized RRT estimator for the sensitive mean μ_Y, in the presence of one non-sensitive but highly correlated auxiliary variable X.

Let Y be the sensitive study variable and X be the non-sensitive auxiliary variable with has a strong positive correlation with Y. Also, let S be scrambling variable (independent of both X and Y) with mean $\mu_S = 0$ and variance σ_S^2. Let μ_X be the known true population mean and σ_X^2 be the known variance of the non-sensitive auxiliary variable X. Let μ_Y be the unknown true population mean and σ_Y^2 be the unknown variance of the sensitive study variable Y. Then the proposed generalized estimator for the sensitive mean μ_Y is given by

$$\widehat{\mu_G} = \Big[\bar{z} + k(\mu_X - \bar{x})\Big]\left(\frac{\mu_D}{\bar{d}}\right)^g = \Big[\hat{\mu}_Y + k(\mu_X - \bar{x})\Big]\left(\frac{\mu_D}{\bar{d}}\right)^g, \tag{VII.20}$$

where $\bar{d} = \lambda(\alpha\bar{x} + \beta) + (1 - \lambda)(\alpha\mu_X + \beta)$ and $\mu_D = \alpha\mu_X + \beta$. Here k and g are

suitable constants. λ is an unknown constant which is determined from the optimality conditions. Further, α and β are known parameters of auxiliary variable X such as C_X, kurtosis, ρ_{ZX}, etc. Various series of estimators can be obtained by using different values of g, k, λ, α and β. Using $g = 1$ will generate a series of ratio estimators while using $g = -1$ will generate a series of product estimators.

We assume a large sample size such that $|\delta_z| > 1$ and $|\delta_x| > 1$ where we define the following error terms:

$$\delta_x = \tfrac{\bar{x}-\mu_X}{\mu_X}; \qquad \delta_z = \tfrac{\bar{z}-\mu_Z}{\mu_Z}$$

Using the error terms defined above, and using the second-order Taylor's approximation, the re-written proposed regression estimator is given by

$$\widehat{\mu_G} = \Big[(\mu_Y + \delta_z\mu_Z) - k\mu_X\delta_x\Big]\left[1 - g\frac{\lambda\alpha\delta_x\mu_X}{\alpha\mu_X+\beta} + \frac{g(g+1)}{2}\left(\frac{\lambda\alpha\delta_x\mu_X}{\alpha\mu_X+\beta}\right)^2\right] \quad \text{(VII.21)}$$

Then subtracting μ_Y from both sides in equation (VII.21) we get

$$\begin{aligned}\widehat{\mu_G} - \mu_Y \approx & \left[-\frac{g\lambda\alpha\mu_Z\mu_X}{\alpha\mu_X+\beta}\delta_x + \frac{g(g+1)}{2}\mu_Z\left(\frac{\lambda\alpha\mu_X}{\alpha\mu_X+\beta}\right)^2\delta_x^2\right] \\ & + \left[\mu_Z\delta_z - \frac{g\lambda\alpha\mu_z\mu_x}{\alpha\mu_X+\beta}\delta_z\delta_x + \frac{g(g+1)}{2}\mu_Z\left(\frac{\lambda\alpha\mu_X}{\alpha\mu_X+\beta}\right)^2\delta_z\delta_x^2\right] \\ & + \left[-k\mu_X\delta_x + \frac{gk\lambda\alpha\mu_X^2}{\alpha\mu_X+\beta}\delta_x^2 - \frac{g(g+1)}{2}k\mu_X^3\left(\frac{\lambda\alpha}{\alpha\mu_X+\beta}\right)^2\delta_x^3\right] \quad \text{(VII.22)}\end{aligned}$$

Under the assumption of bivariate normality (Sukhatme et al. (1970)[59]):

$$E(\delta_x) = 0; \quad E(\delta_z) = 0; \quad E(\delta_x^2) = \tfrac{1-f}{n}C_X^2; \quad E(\delta_z^2) = \tfrac{1-f}{n}C_Z^2; \quad E(\delta_z\delta_x) = \tfrac{1-f}{n}C_{ZX},$$

where $f = \frac{n}{N}$, $C_{ZX} = \rho_{ZX} C_Z C_X$ and C_Z and C_X are the coefficients of variation of Z and X respectively.

Recognizing that $\mu_Z = \mu_Y$, if we consider the scenario with no measurement errors on Z and X, then the Bias of this generalized estimator(VII.20), correct up to first order, is given by

$$\begin{aligned} &Bias(\widehat{\mu_G}) = E[\widehat{\mu_G} - \mu_Y] \\ &\approx \left(\frac{1-f}{n}\right)\left[\left[\frac{g(g+1)}{2}\left(\frac{\lambda\alpha\mu_X}{\alpha\mu_X+\beta}\right)^2\mu_Y + \frac{gk\lambda\alpha\mu_X^2}{\alpha\mu_X+\beta}\right]C_X^2 - \frac{g\lambda\alpha}{\alpha\mu_X+\beta}\sigma_{YX}\right]. \end{aligned} \tag{VII.23}$$

If we consider the scenario with no measurement errors on Z and X, then the MSE of this generalized estimator(VII.20) is given by

$$\begin{aligned} &MSE(\widehat{\mu_G}) = E[\widehat{\mu_G} - \mu_Y]^2 \\ &\approx \Omega\Big[\sigma_Z^2 + g^2\lambda^2R^2\sigma_X^2 + k^2\sigma_X^2 - 2g\lambda R\sigma_{ZX} - 2k\sigma_{ZX} + 2g\lambda kR\sigma_X^2\Big], \end{aligned} \tag{VII.24}$$

where $\Omega = \frac{1-f}{n}$ is the finite population correction factor and $R = \frac{\alpha\mu_Y}{\alpha\mu_X+\beta}$ and $\mu_Z = \mu_Y$. As mentioned earlier, λ is an unknown constant which is determined by optimality conditions. We can obtain the optimal value of λ, i.e. λ_{opt} as follows:

$$\frac{d\,MSE(\widehat{\mu_G})}{d\,\lambda} \approx \Omega\Big[2\lambda g^2R^2\sigma_X^2 - 2gR\sigma_{ZX} + 2gkR\sigma_X^2\Big] = 0. \tag{VII.25}$$

Using equation (VII.25), the optimal values of λ can be given by

$$\lambda_{opt} = \frac{\left(\sigma_{ZX} - k\sigma_X^2\right)}{gR\sigma_X^2}. \tag{VII.26}$$

Hence the minimum MSE for the proposed generalized estimator, using the value of λ_{opt} from equation (VII.26), can be given by

$$MSE(\widehat{\mu_G})_{min} \approx \Omega\sigma_Z^2(1-\rho_{ZX}^2) = \left(\frac{1-f}{n}\right)\sigma_Z^2(1-\rho_{ZX}^2). \tag{VII.27}$$

Here $\rho_{ZX} = \rho_{YX}\frac{\sigma_Y}{\sigma_Z}$.

VII.5 Simulation Study

In this Chapter, we look at a special case of the optional enhanced trust model (Figure VI.1) which was introduced in Chapter VI. Our goal is to simply estimate the mean of the sensitive variable with A and W in the background but when we do not wish to compute their estimates. In order to accomplish it, we consider that the mean of the additive scrambling variable S is $\mu_Y = 0$ instead of $\mu_Y = \theta$ as we had assumed in Chapter VI. As we only have one unknown we wish to estimate, i.e. the sensitive mean μ_Y, we do not use a split-sample technique for the work discussed in this Chapter. Along with an ordinary estimator (equation VII.3), a ratio estimator (equation VII.9) and a regression estimator (equation VII.14) and a generalized estimator (equation VII.20) for the sensitive mean was proposed in this chapter. In this section, the results of a simulation study to evaluate and compare the performance of the ordinary RRT mean estimator $t_0 = \hat{\mu}_Y$, the ratio estimator $t_1 = \widehat{\mu_r}$, the regression estimator $t_2 = \widehat{\mu_{reg}}$ and the generalized estimator $t_3 = \widehat{\mu_G}$.

We conducted a simulation study with 10000 iterations with samples of size $n = 500$ each from a finite population of size $N = 5000$. We generate this finite population under two scenarios (i.e. high and moderate correlation between Y and X).

Scenario-1

We first generate the finite population from a bivariate normal distribution with means and covariances of (X, Y) given as follows:

$$\mu = \begin{bmatrix} 6 \\ 10 \end{bmatrix}, \Sigma = \begin{bmatrix} 8 & 10.1823 \\ 10.1823 & 16 \end{bmatrix}, \ \rho_{YX} = 0.9 \tag{VII.28}$$

i.e. we first generate a population of $N = 5000$ units by using

$$\mu_X = 6, \mu_Y = 10, \sigma_X^2 = 8, \sigma_Y^2 = 16, \rho_{YX} = 0.9. \tag{VII.29}$$

Then from this finite population, we repeatedly draw samples of size $n = 500$ using simple random sampling without replacement (SRSWOR). It must be noted that the real parameters for the 5000 units in the generated finite population are quite close to these assumed parameters, but are not exactly the same, and are given by

$$\mu_X = 6.002015, \mu_Y = 9.995325, \sigma_X^2 = 7.95262, \sigma_Y^2 = 15.80488, \rho_{YX} = 0.8986312. \tag{VII.30}$$

For this simulation study, we used the parameters for the finite population (VII.30) and not the ones assumed to generate the finite population (VII.29). We assume the scrambling variables T and S to be normally distributed with known means and variances $(\mu_T = 1 = E(T), \mu_S = 0; \sigma_T^2 = 0.5, \sigma_S^2 = Var(S) = 0.5\sigma_X^2)$. The empirical results have been averaged over 10000 iterations. The empirical MSE of the estimators $\hat{\mu}_j = t_j \ j = 0, 1, 2, 3$ were computed by

$$MSE(t_j) = \frac{1}{10000} \sum_{i=1}^{10000} \left(\hat{\mu}_i - \mu_Y \right)^2. \tag{VII.31}$$

Here $\hat{\mu}_j = t_j$ can be $\hat{\mu}_Y = t_0$, $\widehat{\mu_r} = t_1$, $\widehat{\mu_{reg}} = t_2$ and $\widehat{\mu_G} = t_3$. Privacy level ∇ was computed for the OET model by taking the mean squared difference between the reported response Z and actual status for the sensitive study variable Y for only those respondents that consider the question sensitive and choose to scramble their responses. This mean squared difference in the report and the actual response is given by

$$\nabla_{oet} = E[Z - Y]^2. \tag{VII.32}$$

Note that although we propose various estimators, they were all proposed under the same model. Hence the value of ∇_{oet} corresponding to all three estimators would be the same as the privacy level is measured for the model being used and not the estimators proposed under that model. Following this, we also compute the unified measure of privacy and efficiency δ which was proposed by Gupta et al. (2018)[21] and is given by

$$\delta = \frac{MSE(t_j)}{\nabla_{oet}}, j = 0, 1, 2. \tag{VII.33}$$

Here we use MSE instead of the variances of the estimators to effectively evaluate the efficiency of the biased estimators. The simulation results for Scenario-1 have been summarized in Table VII.1. The theoretical values have been listed in bold while the regular figures represent the empirical values of the various measures listed in the table.

Table VII.1. Simulation results for estimating sensitive mean μ_Y only (Theoretical (**bold**) and Empirical Measures): $Iterations = 10000$, $N = 5000, n = 500$, $\mu_Y = 10$, $\rho_{YX} = 0.9$, $k = 1$, $g = 1$, $a = 1$ and $b = 0$ for various levels of trust (A) and the sensitivity level (W).

A	W	$MSE(t_0)$	$MSE(t_1)_e$	$MSE(t_2)_e$	$MSE(t_3)_e$	∇_e	$\delta(t_0)_e$	$\delta(t_1)_e$	$\delta(t_2)_e$	$\delta(t_3)_e$
1.00	1.00	0.0379	0.0176	0.0154	0.0158	3.9951	0.0095	0.0044	0.0038	0.0040
		0.0396	**0.0149**	**0.0129**	**0.0127**	**3.9924**	**0.0099**	**0.0037**	**0.0032**	**0.0032**
1.00	0.90	0.0362	0.0159	0.0136	0.0141	4.0149	0.0090	0.0039	0.0034	0.0035
		0.0388	**0.0142**	**0.0122**	**0.0119**	**3.9924**	**0.0097**	**0.0036**	**0.0031**	**0.0030**
1.00	0.80	0.0354	0.0149	0.0126	0.0131	4.0213	0.0088	0.0037	0.0031	0.0033
		0.0380	**0.0135**	**0.0115**	**0.0112**	**3.9924**	**0.0095**	**0.0034**	**0.0029**	**0.0028**
1.00	0.50	0.0330	0.0125	0.0102	0.0107	4.0263	0.0082	0.0031	0.0025	0.0027
		0.0356	**0.0113**	**0.0093**	**0.0091**	**3.9924**	**0.0089**	**0.0028**	**0.0023**	**0.0023**
1.00	0.30	0.0308	0.0103	0.0080	0.0085	3.9276	0.0078	0.0026	0.0020	0.0022
		0.0340	**0.0099**	**0.0076**	**0.0076**	**3.9924**	**0.0085**	**0.0025**	**0.0019**	**0.0019**
0.95	1.00	0.0406	0.0200	0.0177	0.0182	6.4267	0.0063	0.0031	0.0028	0.0028
		0.0454	**0.0202**	**0.0179**	**0.0179**	**6.8905**	**0.0066**	**0.0029**	**0.0026**	**0.0026**
0.95	0.90	0.0396	0.0189	0.0167	0.0171	6.5085	0.0061	0.0029	0.0026	0.0026
		0.0440	**0.0189**	**0.0167**	**0.0166**	**6.8905**	**0.0064**	**0.0027**	**0.0024**	**0.0024**
0.95	0.80	0.0396	0.0190	0.0168	0.0173	6.6166	0.0060	0.0029	0.0025	0.0026
		0.0426	**0.0177**	**0.0154**	**0.0154**	**6.8905**	**0.0062**	**0.0026**	**0.0022**	**0.0022**
0.95	0.50	0.0340	0.0129	0.0107	0.0112	5.9952	0.0057	0.0022	0.0018	0.0019
		0.0385	**0.0139**	**0.0115**	**0.0117**	**6.8905**	**0.0056**	**0.0020**	**0.0017**	**0.0017**
0.95	0.30	0.0325	0.0115	0.0092	0.0096	6.2025	0.0052	0.0019	0.0015	0.0015
		0.0357	**0.0115**	**0.0090**	**0.0092**	**6.8905**	**0.0052**	**0.0017**	**0.0013**	**0.0013**
0.90	1.00	0.0453	0.0246	0.0224	0.0229	8.5674	0.0053	0.0029	0.0026	0.0027
		0.0512	**0.0254**	**0.0231**	**0.0231**	**9.7886**	**0.0052**	**0.0026**	**0.0024**	**0.0024**
0.90	0.90	0.0433	0.0227	0.0205	0.0210	8.2855	0.0052	0.0027	0.0025	0.0025
		0.0492	**0.0236**	**0.0213**	**0.0213**	**9.7886**	**0.0050**	**0.0024**	**0.0022**	**0.0022**
0.90	0.80	0.0424	0.0220	0.0198	0.0203	8.4144	0.0050	0.0026	0.0024	0.0024
		0.0473	**0.0218**	**0.0196**	**0.0196**	**9.7886**	**0.0048**	**0.0022**	**0.0020**	**0.0020**
0.90	0.50	0.0372	0.0158	0.0137	0.0141	8.2773	0.0045	0.0019	0.0017	0.0017
		0.0414	**0.0166**	**0.0141**	**0.0143**	**9.7886**	**0.0042**	**0.0017**	**0.0014**	**0.0015**
0.90	0.30	0.0343	0.0130	0.0108	0.0111	9.0899	0.0038	0.0014	0.0012	0.0012
		0.0375	**0.0130**	**0.0105**	**0.0108**	**9.7886**	**0.0038**	**0.0013**	**0.0011**	**0.0011**
0.85	1.00	0.0499	0.0288	0.0266	0.0271	11.2978	0.0044	0.0025	0.0024	0.0024
		0.0570	**0.0306**	**0.0283**	**0.0283**	**12.6866**	**0.0045**	**0.0024**	**0.0022**	**0.0022**
0.85	0.90	0.0476	0.0266	0.0245	0.0249	10.9846	0.0043	0.0024	0.0022	0.0023
		0.0544	**0.0283**	**0.0260**	**0.0260**	**12.6866**	**0.0043**	**0.0022**	**0.0021**	**0.0021**
0.85	0.80	0.0464	0.0255	0.0234	0.0238	11.0987	0.0042	0.0023	0.0021	0.0021
		0.0519	**0.0260**	**0.0237**	**0.0237**	**12.6866**	**0.0041**	**0.0021**	**0.0019**	**0.0019**
0.85	0.50	0.0399	0.0182	0.0162	0.0165	11.1783	0.0036	0.0016	0.0014	0.0015
		0.0443	**0.0192**	**0.0167**	**0.0169**	**12.6866**	**0.0035**	**0.0015**	**0.0013**	**0.0013**
0.85	0.30	0.0361	0.0147	0.0125	0.0129	11.8427	0.0030	0.0012	0.0011	0.0011
		0.0392	**0.0146**	**0.0121**	**0.0123**	**12.6866**	**0.0031**	**0.0012**	**0.0010**	**0.0010**
0.80	1.00	0.0547	0.0333	0.0312	0.0316	14.1661	0.0039	0.0023	0.0022	0.0022
		0.0628	**0.0358**	**0.0333**	**0.0335**	**15.5847**	**0.0040**	**0.0023**	**0.0021**	**0.0022**
0.80	0.90	0.0515	0.0303	0.0282	0.0286	13.6974	0.0038	0.0022	0.0021	0.0021
		0.0597	**0.0330**	**0.0306**	**0.0307**	**15.5847**	**0.0038**	**0.0021**	**0.0020**	**0.0020**
0.80	0.80	0.0494	0.0284	0.0262	0.0267	14.0054	0.0035	0.0020	0.0019	0.0019
		0.0565	**0.0302**	**0.0278**	**0.0279**	**15.5847**	**0.0036**	**0.0019**	**0.0018**	**0.0018**
0.80	0.50	0.0417	0.0200	0.0180	0.0183	13.4392	0.0031	0.0015	0.0013	0.0014
		0.0472	**0.0218**	**0.0192**	**0.0195**	**15.5847**	**0.0030**	**0.0014**	**0.0012**	**0.0013**
0.80	0.30	0.0372	0.0156	0.0135	0.0138	13.9685	0.0027	0.0011	0.0010	0.0010
		0.0410	**0.0162**	**0.0136**	**0.0139**	**15.5847**	**0.0026**	**0.0010**	**0.0009**	**0.0009**

From Table VII.1, we can make a number of observations. For any fixed value of A, we note that as W increases, MSE worsens for all estimators while the privacy level remains the same. For example, consider the cases where $A = 0.95$. When W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the ordinary RRT mean estimator, $t_0 = \hat{\mu}_Y$, takes values 0.0357, 0.0385, 0.0426, 0.0440, and 0.0454 respectively. Further, as W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the ratio estimator, $t_1 = \widehat{\mu_r}$, takes values 0.0115, 0.0139, 0.0177, 0.0189, and 0.0202 respectively. Also, as W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the regression estimator, $t_2 = \widehat{\mu_{reg}}$, takes values 0.0090, 0.0115, 0.0154, 0.0167, and 0.0179 respectively. When W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the generalized estimator, $t_3 = \widehat{\mu_G}$, takes values 0.0092, 0.0117, 0.0154, 0.0166, and 0.0179 respectively. Therefore, we can infer that the efficiency of all four estimators is the worst possible at their maximum value, i.e. at $W = 1$. Hence optionality element of the proposed model ensures optimal efficiency of any of the three estimators introduced used under this model. Furthermore, based on this example, we can also note that $t3$ has the best efficiency followed closely by t_2 and and then t_1. These three estimators that utilize auxiliary variable information X are considerably more efficient than the ordinary RRT mean estimator t_0 that does not utilize the auxiliary information. The theoretical privacy level is the same corresponding to all four estimators since the model is the same for all four estimators. Therefore, the lowest δ values are corresponding to the generalized estimator t_3 which is fairly similar to the δ values corresponding to the regression estimator t_2.

Another observation that can be made using Table VII.1 is that for fixed values of W, say $W = 0.9$, The efficiency of all estimators worsens as the level of trust A drops. This is expected as fewer people trust the model, they would choose to go with

a more enhanced scrambling option. However, the gain in privacy due to enhanced scrambling of responses more than compensates for the drop in efficiency when we consider the unified measure of privacy and efficiency to assess the overall estimator performance. For example, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_0} takes values 0.0097, 0.0064, 0.0050, 0.0043, and 0.0038 respectively. Similarly, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_1} takes values 0.0036, 0.0027, 0.0024, 0.0022, and 0.0021 respectively. Also, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_2} takes values 0.0031, 0.0024, 0.0022, 0.0021, and 0.0020 respectively. Finally, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_3} takes values 0.0030, 0.0024, 0.0022, 0.0021, and 0.0020 respectively. Hence the best value of δ can be obtained even with considerably lower values A using any of the four estimators. Although t_3 and t_2 have fairly similar overall performances, t_1 is slightly behind t_3 and t_2 with respect to overall performance evaluated by δ values. However, it is noteworthy that t_1, t_2, and t_3 perform considerably better than t_0.

Scenario-2

As shown in Scenario-1, we now generate the finite population from a bivariate normal distribution with means and covariances of (X, Y) given as follows:

$$\mu = \begin{bmatrix} 6 \\ 10 \end{bmatrix}, \Sigma = \begin{bmatrix} 8 & 6.788225 \\ 6.788225 & 16 \end{bmatrix}, \ \rho_{YX} = 0.6 \tag{VII.34}$$

i.e. we first generate a population containing 5000 units by using

$$\mu_X = 6, \mu_Y = 10, \sigma_X^2 = 8, \sigma_Y^2 = 16, \rho_{YX} = 0.6. \tag{VII.35}$$

Then from this finite population, we repeatedly draw samples of size $n = 500$ using simple random sampling without replacement (SRSWOR). It must be noted that the real parameters for the 5000 units in the generated finite population are quite close to these assumed parameters, but are not exactly the same, and are given by

$$\mu_X = 6.006096, \mu_Y = 9.993591, \sigma_X^2 = 8.010827, \sigma_Y^2 15.79687, \rho_{YX} = 0.5967508. \tag{VII.36}$$

For this simulation study, we used the parameters for the finite population (VII.36) and not the ones assumed to generate the finite population (VII.35). The only difference between Scenario-1 and Scenario-2 is that we now use a considerably lower value for ρ_{YX} to study how the performance of the three estimators gets affected. We assume the scrambling variables T and S to be normally distributed with known means and variances ($\mu_T = 1 = E(T), \mu_S = 0; \sigma_T^2 = 0.5, \sigma_S^2 = Var(S) = 0.5\sigma_X^2$). The empirical results have again been averaged over 10000 iterations. The simulation results for Scenario-2 have been summarized in Table VII.2. The theoretical values have been listed in bold while the regular figures represent the empirical values of the various measures listed in the table.

From Table VII.2, we can make a number of observations. For any fixed value of A, we note that as W increases, MSE worsens for all estimators while the privacy level remains the same. For example, consider the cases where $A = 0.95$. When W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the ordinary RRT mean estimator, $t_0 = \hat{\mu}_Y$, takes values 0.0357, 0.0385, 0.0427, 0.0440, and 0.0454 respectively. Further, as W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the ratio estimator, $t_1 = \widehat{\mu_r}$, takes values 0.0319, 0.0344, 0.0381, 0.0394, and

Table VII.2. Simulation results for estimating sensitive mean μ_Y only (Theoretical (**bold**) and Empirical Measures): $Iterations = 10000$, $N = 5000, n = 500$, $\mu_Y = 10$, $\rho_{YX} = 0.6$, $k = 1$, $g = 1$, $a = 1$ and $b = 0$ for various levels of trust (A) and the sensitivity level (W).

A	W	$MSE(t_0)$	$MSE(t_1)$	$MSE(t_2)$	$MSE(t_3)$	∇_{oet}	$\delta(t_0)$	$\delta(t_1)$	$\delta(t_2)$	$\delta(t_3)$
1	1	0.0379	0.0378	0.0280	0.0281	4.0243	0.0094	0.0094	0.0070	0.0070
		0.0396	**0.0354**	**0.0257**	**0.0255**	**4.0216**	**0.0099**	**0.0088**	**0.0064**	**0.0064**
1	0.9	0.0362	0.0358	0.0261	0.0263	4.0443	0.0090	0.0089	0.0065	0.0065
		0.0388	**0.0347**	**0.0250**	**0.0248**	**4.0216**	**0.0097**	**0.0086**	**0.0062**	**0.0062**
1	0.8	0.0354	0.0348	0.0252	0.0253	4.0507	0.0087	0.0086	0.0062	0.0063
		0.0380	**0.0339**	**0.0243**	**0.0241**	**4.0216**	**0.0095**	**0.0084**	**0.0060**	**0.0060**
1	0.5	0.0329	0.0327	0.0229	0.0230	4.0557	0.0081	0.0081	0.0056	0.0057
		0.0356	**0.0318**	**0.0221**	**0.0219**	**4.0216**	**0.0089**	**0.0079**	**0.0055**	**0.0055**
1	0.3	0.0308	0.0306	0.0207	0.0209	3.9564	0.0078	0.0077	0.0052	0.0053
		0.0340	**0.0303**	**0.0205**	**0.0205**	**4.0216**	**0.0085**	**0.0075**	**0.0051**	**0.0051**
0.95	1	0.0406	0.0399	0.0303	0.0304	6.4142	0.0063	0.0062	0.0047	0.0047
		0.0454	**0.0406**	**0.0307**	**0.0308**	**6.9186**	**0.0066**	**0.0059**	**0.0044**	**0.0044**
0.95	0.9	0.0395	0.0388	0.0292	0.0294	6.4913	0.0061	0.0060	0.0045	0.0045
		0.0440	**0.0394**	**0.0296**	**0.0295**	**6.9186**	**0.0064**	**0.0057**	**0.0043**	**0.0043**
0.95	0.8	0.0396	0.0390	0.0294	0.0295	6.5960	0.0060	0.0059	0.0045	0.0045
		0.0427	**0.0381**	**0.0283**	**0.0283**	**6.9186**	**0.0062**	**0.0055**	**0.0041**	**0.0041**
0.95	0.5	0.0339	0.0331	0.0234	0.0236	6.0104	0.0056	0.0055	0.0039	0.0039
		0.0385	**0.0344**	**0.0244**	**0.0245**	**6.9186**	**0.0056**	**0.0050**	**0.0035**	**0.0035**
0.95	0.3	0.0324	0.0317	0.0218	0.0221	6.2178	0.0052	0.0051	0.0035	0.0035
		0.0357	**0.0319**	**0.0219**	**0.0220**	**6.9186**	**0.0052**	**0.0046**	**0.0032**	**0.0032**
0.9	1	0.0453	0.0447	0.0351	0.0352	8.5772	0.0053	0.0052	0.0041	0.0041
		0.0512	**0.0458**	**0.0360**	**0.0360**	**9.8156**	**0.0052**	**0.0047**	**0.0037**	**0.0037**
0.9	0.9	0.0433	0.0427	0.0331	0.0333	8.2966	0.0052	0.0051	0.0040	0.0040
		0.0493	**0.0440**	**0.0342**	**0.0342**	**9.8156**	**0.0050**	**0.0045**	**0.0035**	**0.0035**
0.9	0.8	0.0425	0.0420	0.0324	0.0325	8.4292	0.0050	0.0050	0.0038	0.0039
		0.0473	**0.0423**	**0.0324**	**0.0324**	**9.8156**	**0.0048**	**0.0043**	**0.0033**	**0.0033**
0.9	0.5	0.0371	0.0361	0.0266	0.0268	8.2764	0.0045	0.0044	0.0032	0.0032
		0.0414	**0.0370**	**0.0271**	**0.0271**	**9.8156**	**0.0042**	**0.0038**	**0.0028**	**0.0028**
0.9	0.3	0.0342	0.0335	0.0237	0.0239	9.0907	0.0038	0.0037	0.0026	0.0026
		0.0375	**0.0334**	**0.0235**	**0.0236**	**9.8156**	**0.0038**	**0.0034**	**0.0024**	**0.0024**
0.85	1	0.0499	0.0489	0.0394	0.0396	11.3080	0.0044	0.0043	0.0035	0.0035
		0.0570	**0.0510**	**0.0411**	**0.0412**	**12.7126**	**0.0045**	**0.0040**	**0.0032**	**0.0032**
0.85	0.9	0.0475	0.0465	0.0371	0.0373	10.9951	0.0043	0.0042	0.0034	0.0034
		0.0545	**0.0487**	**0.0389**	**0.0389**	**12.7126**	**0.0043**	**0.0038**	**0.0031**	**0.0031**
0.85	0.8	0.0464	0.0455	0.0360	0.0362	11.1038	0.0042	0.0041	0.0032	0.0033
		0.0519	**0.0465**	**0.0366**	**0.0366**	**12.7126**	**0.0041**	**0.0037**	**0.0029**	**0.0029**
0.85	0.5	0.0398	0.0385	0.0292	0.0294	11.1707	0.0036	0.0034	0.0026	0.0026
		0.0443	**0.0396**	**0.0296**	**0.0297**	**12.7126**	**0.0035**	**0.0031**	**0.0023**	**0.0023**
0.85	0.3	0.0360	0.0351	0.0254	0.0256	11.8520	0.0030	0.0030	0.0021	0.0022
		0.0392	**0.0350**	**0.0250**	**0.0252**	**12.7126**	**0.0031**	**0.0028**	**0.0020**	**0.0020**
0.8	1	0.0546	0.0533	0.0440	0.0442	14.1964	0.0038	0.0038	0.0031	0.0031
		0.0628	**0.0562**	**0.0462**	**0.0464**	**15.6096**	**0.0040**	**0.0036**	**0.0030**	**0.0030**
0.8	0.9	0.0515	0.0501	0.0409	0.0411	13.7220	0.0038	0.0037	0.0030	0.0030
		0.0597	**0.0534**	**0.0434**	**0.0436**	**15.6096**	**0.0038**	**0.0034**	**0.0028**	**0.0028**
0.8	0.8	0.0494	0.0483	0.0389	0.0391	14.0336	0.0035	0.0034	0.0028	0.0028
		0.0566	**0.0506**	**0.0406**	**0.0408**	**15.6096**	**0.0036**	**0.0032**	**0.0026**	**0.0026**
0.8	0.5	0.0417	0.0403	0.0310	0.0312	13.4767	0.0031	0.0030	0.0023	0.0023
		0.0472	**0.0422**	**0.0322**	**0.0324**	**15.6096**	**0.0030**	**0.0027**	**0.0021**	**0.0021**
0.8	0.3	0.0371	0.0360	0.0264	0.0266	14.0071	0.0026	0.0026	0.0019	0.0019
		0.0410	**0.0366**	**0.0265**	**0.0267**	**15.6096**	**0.0026**	**0.0023**	**0.0017**	**0.0017**

0.0406 respectively. Also, as W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the regression estimator, $t_2 = \widehat{\mu_{reg}}$, takes values 0.0219, 0.0244, 0.0283, 0.0296, and 0.0307 respectively. When W varies between 0.3, 0.5, 0.8, 0.9 and 1, the theoretical MSE for the generalized estimator, $t_3 = \widehat{\mu_G}$, takes values 0.0220, 0.0245, 0.0283, 0.0295, and 0.0308 respectively. Therefore, we can infer that the efficiency of all four estimators is the worst possible at their maximum value, i.e. at $W = 1$. Hence optionality element of the proposed model ensures optimal efficiency of any of the three estimators introduced used under this model. Furthermore, based on this example, we can also note that t_3 has the best efficiency followed closely by t_2 and and then t_1. These three estimators that utilize auxiliary variable information X are considerably more efficient than the ordinary RRT mean estimator t_0 that does not utilize the auxiliary information. The theoretical privacy level is the same corresponding to all four estimators since the model is the same for all four estimators. Therefore, the lowest δ values are corresponding to the generalized estimator t_3 which is fairly similar to the δ values corresponding to the regression estimator t_2.

Another observation that can be made using Table VII.2 is that for fixed values of W, say $W = 0.9$, The efficiency of all estimators worsens as the level of trust A drops. This is expected as fewer people trust the model, they would choose to go with a more enhanced scrambling option. However, the gain in privacy due to enhanced scrambling of responses more than compensates for the drop in efficiency when we consider the unified measure of privacy and efficiency to assess the overall estimator performance. For example, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_0} takes values 0.0097, 0.0064, 0.0050, 0.0043, and 0.0038 respectively. Similarly, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_1} takes values 0.0086, 0.0057, 0.0045, 0.0038, and 0.0034 respectively. Also, as A varies between 1, 0.95, 0.9,

0.85, and 0.8, theoretical δ_{t_2} takes values 0.0062, 0.0043, 0.0035, 0.0031, and 0.0028 respectively. Finally, as A varies between 1, 0.95, 0.9, 0.85, and 0.8, theoretical δ_{t_3} takes values 0.0062, 0.0043, 0.0035, 0.0031, and 0.0028 respectively. Hence the best value of δ can be obtained even with considerably lower values A using any of the four estimators. Although t_3 and t_2 have fairly similar overall performances, t_1 is slightly behind t_3 and t_2 with respect to overall performance evaluated by δ values. However, it is noteworthy that t_1, t_2, and t_3 perform considerably better than t_0.

On comparing the results from Table VII.1 and Table VII.2 we can note the MSE for t_1, t_2 and t_3 worsens when ρ_{YX} lowers. However, the MSE for t_1 worsens considerably more than that of t_2 and t_3. The performance of t_2 and t_3 are practically the same. However, t_3 and t_2 still perform the best of the four proposed estimators by a reasonable margin. For instance, let's consider the case when $A = 0.9$ and $W = 0.8$. When $\rho_{YX} = 0.9$, the theoretical MSE for t_0, t_1 t_2 and t_3 are 0.0473, 0.0218, 0.0196, and 0.0196 respectively. However, when $\rho_{YX} = 0.6$, the theoretical MSE for t_0, t_1 and t_2 are 0.0473, 0.0423, 0.0324, and 0.0324 respectively. Further, when $\rho_{YX} = 0.9$, the theoretical δ for t_0, t_1, t_2 and t_3 are 0.0048, 0.0022, 0.0020, and 0.0020 respectively. However, when $\rho_{YX} = 0.6$, the theoretical δ for t_0, t_1, t_2 and t_3 are 0.0048, 0.0043, 0.0033, and 0.0033 respectively. Hence, we can see that the overall performance of the proposed additive ratio estimator (t_1) worsens considerably when the correlation between the sensitive study variable Y and the non-sensitive auxiliary variable is weak. Moreover, we note that despite the weaker correlation, the proposed regression estimator (t_2) and the proposed generalized estimator still outperform the other two proposed estimators, by a reasonable margin, even when the correlation between the sensitive study variable Y and the non-sensitive auxiliary variable is moderate or weak.

Further, comparing the tables in Chapters VI and Chapter VII, we also note that for the case when ρ_{YX} along with all other parameters are the same, the MSE values obtained in Chapter VI tables were drastically higher compared to the corresponding MSE values obtained in Chapter VII tables. For instance, consider the case when $A = 0.9$ and $W = 0.8$. When $\rho_{YX} = 0.9$, the theoretical MSE values for t_0, t_1, and t_2 are 0.4531, 0.4313, and 0.4289 respectively. However, when $\rho_{YX} = 0.6$, the theoretical MSE values for t_0, t_1, and t_2 are 0.0473, 0.0218, and 0.0196 respectively. This drastic difference in MSE values can be attributed to whether or not the split-sample technique was used in the study. Note that the results summarized in Tables (VI.1VI.2) have been computed based on the split sample technique which uses two sub-samples of size 250 each instead of a single sample of size 500 which was used to compute the results summarized in Tables (VII.1, VII.2).

VII.6 Concluding Chapter Remarks

In this chapter, we looked at a special case of the Optional Enhanced Trust model proposed in Chapter VI. We assume that the random additive noise available to the respondents is from a population with a mean of 0 which means on an aggregate level no random noise is being added in a way that would alter the statistical properties of the reported response i.e. on an average the reported response is what the true response to the sensitive question would be for each individual in the population ($\mu_Z = \mu_Y$). We introduce estimators for the sensitive mean that can be used in the absence as well in the presence of a non-sensitive auxiliary variable. The primary finding of this chapter was that when the auxiliary information is available for every population unit, using the generalized estimator or the regression estimator introduced in this

chapter, would have the best performance in terms of efficiency as well as in terms of the unified measure of privacy and efficiency. The performance of the ratio estimator chapter is only slightly worse than the regression estimator and would be a more appropriate choice when the relationship of sensitive study variable Y and auxiliary variable X goes through the origin. These results hold even when the correlation between the sensitive study variable Y and the non-sensitive auxiliary variable X is moderate. Therefore, utilizing non-sensitive auxiliary information, whenever possible, can considerably improve the sensitive mean estimation.

Chapter VIII: Concluding Remarks and Future Directions

VIII.1 General Discussion of Work and Remarks

In this dissertation, a discussion has been presented on accounting for respondents' lack of trust in a survey using the randomized response technique (RRT) for sensitive and private data collection. This work was done for cases when the survey question has both binary and quantitative responses. A class of mixture models in both binary and quantitative RRT areas were proposed in order to mitigate the lack of trust in the traditional RRT models. The performance of the proposed models was evaluated and compared with the performances of traditional models in the respective areas with respect to estimation efficiency (i.e. using MSE), privacy loss or privacy level of the model as appropriate for the type of model and using a unified measure for privacy and efficiency to gauge the overall model performance.

In Chapter I, a background was presented for social desirability bias (SBD) and various methods to help mitigate its effect in a sensitive question survey. In Chapter II, various RRT models from the literature were introduced along with an introduction to the concept of partially homomorphic encryption techniques and their potential utility in the estimation of the sensitive trait prevalence in a population.

In Chapter III, Chapter IV and Chapter V we propose mixture models and methods for the case when the sensitive survey question only has a binary response ("Yes"/"No"). In Chapters VI and VII, we present an optional quantitative RRT model and various estimators under the model both in the presence and in the absence of a non-sensitive auxiliary variable which is highly correlated with the sensitive study variable.

Through all the work summarized in this dissertation, we were able to make some critical observations. We note that accounting for the respondents' lack of trust in the survey model is critical as not doing so could introduce a significant negative bias in our estimates. This would happen as those respondents who do not trust the survey method but have the sensitive trait would lie and report that they do not have the sensitive trait. This would result in a less-than-accurate level of reporting of the sensitive behavior prevalence which would invariably introduce a negative bias. We observe a similar behavior when the goal is with the sensitive mean estimation scenarios. Moreover, using optional models over non-optional models is highly recommended as they improve the estimator efficiency by allowing respondents who do not find the survey question sensitive to not add any unnecessary noise to their response and thus capture more "truth". We also note that the optionality element of an RRT model helps even when a researcher might erroneously ignore the accounting of the respondent lack of trust in the survey method. A prominent finding in the area of sensitive mean estimation was that the proposed ratio, regression and generalized estimators perform better than the ordinary RRT estimators. This can be attributed to the fact the proposed ratio, regression and generalized estimators leverage the high correlation between the unknown sensitive study variable and the non-sensitive auxiliary information which is available on every population unit.

VIII.2 Future Directions

For future studies, one could extend the work done in the binary RRT area using the Hybrid (Paillier + Warner RRT) model to define a more cohesive measure of overall privacy protection under the proposed model that accounts for the threat to respondent privacy due to surveyor dishonesty. In the quantitative area, one could extend the presented work to account for various non-sampling errors such as measurement errors and non-responses. In the sensitive mean estimation area, one could also explore methods to make use of an auxiliary variable when it is no longer assumed to be non-sensitive and/or when only partial information is available on the auxiliary variable.

References

[1] Henry Assael and John Keon. Nonsampling vs. sampling errors in survey research. *Journal of Marketing*, 46(2):114–123, 2023/05/06/ 1982. Full publication date: Spring, 1982.

[2] Graeme Blair, Kosuke Imai, and Yang-Yang Zhou. Design and analysis of the randomized response technique. *Journal of the American Statistical Association*, 110(511):1304–1319, 2015.

[3] Milani Chaloupka. Application of the randomized response technique to marine park management: an assessment of permit compliance. *Environmental Management*, 9:393–398, 09 1985.

[4] Anu Chhabra, B Dass, and Sat Gupta. Estimating prevalence of sexual abuse by an acquaintance with an optional unrelated question rrt model. *North Carolina Journal of Mathematics and Statistics*, 2:1–9, 01 2016.

[5] L P Chow, W Gruhn, and W P Chang. Feasibility of the randomized response technique in rural ethiopia. *American Journal of Public Health*, 69(3):273–276, 1979. PMID: 420374.

[6] A. Chumbley, C. Williams, A. Duna, V. Pandya, and Jimin Khim. Homomorphic encryption. retrieved from Brilliant.org, July 18, 2021.

[7] Janne Chung and Gary S. Monroe. Exploring social desirability bias. *Journal of Business Ethics*, 44(4):291–302, Jun 2003.

[8] Douglas Crowne and David Marlowe. A new scale of social desirability independent of psychopathology. *Journal of consulting psychology*, 24:349–54, 09 1960.

[9] Dan Dalton, JAMES WIMBUSH, and CATHERINE DAILY. Using the unmatched count technique (uct) to estimate base-rates for sensitive behavior. *Personnel Psychology*, 47:817 – 829, 12 2006.

[10] Giancarlo Diana and Pier Francesco Perri. A class of estimators for quantitative sensitive data. *Statistical Papers*, 52:633–650, 08 2011.

[11] Benjamin H. Eichhorn and Lakhbir S. Hayre. Scrambled randomized response methods for obtaining sensitive quantitative data. *Journal of Statistical Planning and Inference*, 7(4):307–316, 1983.

[12] Mohamed Elfil and Ahmed Negida. Sampling methods in clinical research; an educational review. *Emerg. (Tehran)*, 5(1):e52, Jan 2017. PMC5325924.

[13] T. Elgamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, 31(4):469–472, 1985.

[14] Michael A. Fligner, George E. Policello II, and Jagbir Singh. A comparison of two randomized response survey methods with consideration for the level of respondent protection. *Communications in Statistics - Theory and Methods*, 6(15):1511–1524, 1977.

[15] Caroline Fontaine and Fabien Galand. A survey of homomorphic encryption for nonspecialists. *EURASIP Journal on Information Security*, 2007(1):013801, Dec 2007.

[16] Bernard G. Greenberg, Abdel-Latif A. Abul-Ela, Walt R. Simmons, and Daniel G. Horvitz. The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association*, 64(326):520–539, 1969.

[17] Bernard G. Greenberg, Roy R. Kuebler Jr., James R. Abernathy, and Daniel G. Horvitz. Application of the randomized response technique in obtaining quantitative data. *Journal of the American Statistical Association*, 66(334):243–250, 1971.

[18] S. Gupta, G. Kalucha, J. Shabbir, and B. K. Dass. Estimation of finite population mean using optional rrt models in the presence of nonsensitive auxiliary information. *American Journal of Mathematical and Management Sciences*, 33(2):147–159, 2014.

[19] Sat Gupta, Bhisham Gupta, and Sarjinder Singh. Estimation of sensitivity level of personal interview survey questions. *Journal of Statistical Planning and Inference*, 100(2):239–247, 2002.

[20] Sat Gupta, Geeta Kalucha, and Javid Shabbir. A regression estimator for finite population mean of a sensitive variable using an optional randomized response model. *Communications in Statistics - Simulation and Computation*, 46(3):2393–2405, 2017.

[21] Sat Gupta, Samridhi Mehta, Javid Shabbir, and Sadia Khalil. A unified measure of respondent privacy and model efficiency in quantitative rrt models. *Journal of Statistical Theory and Practice*, 12(3):506–511, 2018.

[22] Sat Gupta, Javid Shabbir, and Supriti Sehra. Mean and sensitivity estimation in optional randomized response models. *Journal of Statistical Planning and Inference*, 140(10):2870–2874, 2010.

[23] Sat Gupta, Javid Shabbir, Rita Sousa, and Pedro Corte-Real. Estimation of the mean of a sensitive variable in the presence of auxiliary information. *Communications in Statistics - Theory and Methods*, 41(13-14):2394–2404, 2012.

[24] Sat Gupta, Anna Tuck, Tracy Gill, and Mary Crowe. Optional unrelated-question randomized response models. *Involve: A Journal of Mathematics*, 6(4):483 – 492, 2013.

[25] Sat Gupta, Joia Zhang, Sadia Khalil, and Pujita Sapra. Mitigating lack of trust in quantitative randomized response technique models. *Communications in Statistics - Simulation and Computation*, 0(0):1–9, 2022.

[26] Amy Hinsley, Aidan Keane, Freya A. V. St. John, Harriet Ibbett, and Ana Nuno. Asking sensitive questions using the unmatched count technique: Applications and guidelines for conservation. *Methods in Ecology and Evolution*, 10(3):308–319, 2019.

[27] ICO-UK. Types of encryption. http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/, July 4 2021.

[28] Edward Jones and Harold Sigall. The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, 76:349–364, 11 1971.

[29] Geeta Kalucha, Sat Gupta, and B. K. Dass. Ratio estimation of finite population mean using optional randomized response models. *Journal of Statistical Theory and Practice*, 9(3):633–645, 2015.

[30] Geeta Kalucha, Sat Gupta, and Javid Shabbir. A two-step approach to ratio and regression estimation of finite population mean using optional randomized response models. *Hacettepe Journal of Mathematics and Statistics*, 45:1819 – 1830, 2016.

[31] Jiban Khadka. Sampling error in survey research. *International Journal of Science and Research (IJSR)*, 8(1):2214–2220, Jan 2019.

[32] Sadia Khalil, Muhammad Noor ul Amin, and Muhammad Hanif. Estimation of population mean for a sensitive variable in the presence of measurement error. *Journal of Statistics and Management Systems*, 21(1):81–91, 2018.

[33] Sadia Khalil, Qi Zhang, and Sat Gupta. Mean estimation of sensitive variables under measurement errors using optional rrt models. *Communications in Statistics - Simulation and Computation*, 50(5):1417–1426, 2021.

[34] Ryan Ko and Raymond Choo. *The Cloud Security Ecosystem*. Syngress, 2015.

[35] Ivar Krumpal. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4):2025–2047, Jun 2013.

[36] Samuel S. K. Kwan, Mike K. P. So, and Kar Yan Tam. Research note: Applying the randomized response technique to elicit truthful responses to sensitive questions in is research: The case of software piracy behavior. *Information Systems Research*, 21(4):941–959, 2010.

[37] A Lange. An overview of homomorphic encryption, May 9, 2011.

[38] Jan Lanke. On the degree of protection in randomized interviews. *International Statistical Review*, 44, 08 1976.

[39] Ronald B. Larson. Controlling social desirability bias. *International Journal of Market Research*, 61(5):534–547, 2019.

[40] K. Lauter, M. Naehrig, and V. Vaikuntanathan. *Can homomorphic encryption be practical?* Cryptology ePrint Archive. IACR, 2011.

[41] Maxwell Lovig, Sadia Khalil, Sumaita Rahman, Pujita Sapra, and Sat Gupta. A mixture binary rrt model with a unified measure of privacy and efficiency. *Communications in Statistics - Simulation and Computation*, 0(0):1–12, 2021.

[42] Samridhi Mehta and Priyanka Aggarwal. Bayesian estimation of sensitivity level and population proportion of a sensitive characteristic in a binary optional unrelated question rrt model. *Communications in Statistics - Theory and Methods*, 47(16):4021–4028, 2018.

[43] A. J. Menezes, Van Oorschot Paul C., and Scott A. Vanstone. *Chapter-3 Number-Theoretic Reference Problems*, page 90–98. CRC Press, 2001.

[44] Liam Morris. Analysis of partially and fully homomorphic encryption, May 10 2013.

[45] G. Narjis and J. Shabbir. Estimating the prevalence of sensitive attribute with optional unrelated question randomized response models under simple and stratified random sampling. *Scientia Iranica*, 28(5):2851–2867, 2021.

[46] Ghulam Narjis and Javid Shabbir. Estimation of population proportion and sensitivity level using optional unrelated question randomized response techniques. *Communications in Statistics - Simulation and Computation*, 49(12):3212–3226, 2020.

[47] Monique Ogburn, Claude Turner, and Pushkar Dahal. Homomorphic encryption. *Procedia Computer Science*, 20:502–509, 2013. Complex Adaptive Systems.

[48] Pascal Paillier. Public-key cryptosystems based on composite degree residuosity classes. In Jacques Stern, editor, *Advances in Cryptology — EUROCRYPT '99*, pages 223–238, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.

[49] Eric Plutzer. Privacy, Sensitive Questions, and Informed Consent: Their Impacts on Total Survey Error, and the Future of Survey Research. *Public Opinion Quarterly*, 83(S1):169–184, 06 2019.

[50] K. H. Pollock and Yuksel Bek. A comparison of three randomized response models for quantitative data. *Journal of the American Statistical Association*, 71(356):884–886, 1976.

[51] A. Quatember. Chapter 7 - a mixture of true and randomized responses in the estimation of the number of people having a certain attribute. In Arijit Chaudhuri, Tasos C. Christofides, and C.R. Rao, editors, *Data Gathering, Analysis and Protection of Privacy Through Randomized Response Techniques: Qualitative and Quantitative Human Traits*, volume 34 of *Handbook of Statistics*, pages 105–117. Elsevier, 2016.

[52] William Reynolds. Development of reliable and valid short forms of the marlow–crowne social desirability scale. *Journal of Clinical Psychology*, 38:119–125, 01 1982.

[53] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, feb 1978.

[54] Pujita Sapra, Sadia Khalil, and Sat Gupta. Accounting for lack of trust in optional binary rrt models using a unified measure of privacy and efficiency. *Journal of Statistical Theory and Practice*, 16(3):51, Jul 2022.

[55] Nj Scheers. A review of randomized response technique. *Measurement and Evaluation in Counseling and Development*, 25:27–41, 04 1992.

[56] Rakesh Shrestha and Shiho Kim. Chapter ten - integration of iot with blockchain and homomorphic encryption: Challenging issues and opportunities. In Shiho Kim, Ganesh Chandra Deka, and Peng Zhang, editors, *Role of Blockchain Technology in IoT Applications*, volume 115 of *Advances in Computers*, pages 293–331. Elsevier, 2019.

[57] Sarjinder Singh, Anwar Joarder, and Maxwell King. Regression analysis using scrambled responses. *Australian Journal of Statistics*, 38:201 – 211, 06 2008.

[58] Rita Sousa, Javid Shabbir, Pedro Corte Real, and Sat Gupta. Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information. *Journal of Statistical Theory and Practice*, 4(3):495–507, 2010.

[59] P.V. Sukhatme and B.V. Sukhatme. *Sampling Theory of Surveys with Applications*. Iowa State University Press, Ames, 1970.

[60] Steven K Thompson. *Sampling*. Wiley series in probability and statistics. Wiley, 3 edition, 2012.

[61] Roger Tourangeau and Ting Yan. Sensitive questions in surveys. *Psychological bulletin*, 133:859–83, 10 2007.

[62] Stanley L Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

[63] Stanley L. Warner. The linear randomized response model. *Journal of the American Statistical Association*, 66(336):884–888, 1971.

[64] Mark A. Will and Ryan K.L. Ko. Chapter 5 - a guide to homomorphic encryption. In Ryan Ko and Kim-Kwang Raymond Choo, editors, *The Cloud Security Ecosystem: Technical, Legal, Business and Management Issues*, pages 101–127. Syngress, Boston, 2015.

[65] Ting Yan. Consequences of asking sensitive questions in surveys. *Annual Review of Statistics and Its Application*, 8(1):109–127, 2021.

[66] Zaizai Yan, Jingyu Wang, and Junfeng Lai. An efficiency and protection degree-based comparison among the quantitative randomized response strategies. *Communications in Statistics - Theory and Methods*, 38(3):400–408, 2008.

[67] Amber Young, Sat N Gupta, and Ryan Parks. A binary unrelated-question rrt model accounting for untruthful responding. *Involve: A Journal of Mathematics*, 12(7):1163–1173, 2019.

[68] Qi Zhang, Sat Gupta, Geeta Kalucha, and Sadia Khalil. Ratio estimation of the mean under rrt models. *Journal of Statistics and Management Systems*, 22(1):97–113, 2019.

[69] Qi Zhang, Sadia Khalil, and Sat Gupta. Mean estimation in the simultaneous presence of measurement errors and non-response using optional RRT models under stratified sampling. *Journal of Statistical Computation and Simulation*, 06 2021. In Press.

[70] Hongchao Zhou and Gregory Wornell. Efficient homomorphic encryption on integer vectors and its applications. In *2014 Information Theory and Applications Workshop (ITA)*, pages 1–9, 02 2014.

Chapter A: List of Publications

1. Maxwell Lovig, Sadia Khalil, Sumaita Rahman, Pujita Sapra, and Sat Gupta. A mixture binary rrt model with a unified measure of privacy and efficiency. *Communications in Statistics - Simulation and Computation*, 0(0):1–12, 2021.

2. Sat Gupta, Joia Zhang, Sadia Khalil, and Pujita Sapra. Mitigating lack of trust in quantitative randomized response technique models. *Communications in Statistics - Simulation and Computation*, 0(0):1–9, 2022.

3. Pujita Sapra, Sadia Khalil, and Sat Gupta. Accounting for lack of trust in optional binary rrt models using a unified measure of privacy and efficiency. *Journal of Statistical Theory and Practice*, 16(3):51, Jul 2022.