THREE ESSAYS ON THE INCENTIVES AND DESIGN OF SURVEY TECHNIQUES

by

Timothy Flannery

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ECONOMICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2015

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Timothy Flannery entitled Three Essays on the Incentives and Design of Survey Techniques and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 14 April 2015

Andreas Blume

Date: 14 April 2015

Mark Walker

Date: 14 April 2015

Stanley S. Reynolds

Date: 14 April 2015

Asaf Plan

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.
I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 14 April 2015

Dissertation Director: Andreas Blume

STATEMENT BY AUTHOR

SIGNED: Timothy Flannery

ACKNOWLEDGEMENTS

I want to thank all those who supported and assisted me in completing this dissertation, particularly my committee members. Andreas Blume introduced me to response techniques in the classroom and continually gave me valuable insight throughout. Throughout my whole career as a graduate student, Mark Walker provided me guidance and direction, for which I cannot be grateful enough. Because of their support, I developed into the researcher I am today. Additionally, I am grateful to my other committee members Stanley Reynolds and Asaf Plan for their critical feedback.

I would also like to acknowledge my fellow graduate students for their role in the process. I enjoyed studying my first year preliminary exams with Michael Phizicky and David Reitzke. I will especially miss our late night study sessions often with deep discussions or simple games to provide a momentary distraction from our work. In my first years as a graduate student, Lindsey Nagy, Alex Roomets, Dina Shatnawi, Eric Cardella, and Trevor Kollmann gave me counsel from their experience as a young student. Throughout the years, many students gave me valuable insight on my work. David Rietzke, Stepen Roberts, Phil Hihn, Alex Hollingsworth, James Fisher, Ivan Rudik, and Patrick Cizek assisted me in my preparation for several presentations and often critiqued my work.

I am grateful to John Wooders, Rabah Amir, and again Mark Walker whose courses inspired me to become an applied theorist. In addition to developing into a researcher, I want to express gratitude to those who helped me in my development as an instructor. If I ever needed guidance in the classroom, I knew I could find John Drabicki with his deep knowledge of pedagogy diligently working in his office, but always willing to spend a few moments to suggest improvements to further develop my skills as an instructor.

Finally, I appreciate all the support my family provided, especially my beautiful wife Jackie and her family (the Neidkowskis), my parents–James and Tonja Flannery–and my brothers: Kevin, John, and Brian Flannery. My wife gave me support during tumultuous times for which I am forever grateful. Of course, my parents always provided encouragement and advice when solicited. In any time of distress, I knew my family would provide me with comfort.

DEDICATION

To my beautiful wife Jaclyn Flannery

I enjoyed 2014 more than any other year. Our wedding far exceeded my expectations as I never attended a better one. I look forward to sharing the rest of my life with you and will cherish every moment of it.

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

My dissertation focuses on the design and incentives of survey techniques. As many institutions use surveys to allocate funding or determine policy, ensuring surveys provide accurate information is essential. Though incentives certainly play a role in whether survey participants report information truthfully, economists have largely overlooked the issue while statisticians tend to focus on estimators without directly modeling incentive constraints. One of the chapters models and analyzes the incentives of a commonly used survey technique, randomized response, while the other two chapters of my dissertation design two response techniques which improve upon others found in the literature by obtaining more precise estimates and/or incentivizing participants better.

In Chapter One "A Game Theoretic Analysis of the Randomized Response Technique", I explicitly model the decision of participants to truthfully respond in the randomized response survey as a game. Randomized response techniques are used to determine the proportion of a population that belongs to a stigmatized group and introduce noise so the surveyor cannot perfectly infer whether a participant belongs to a stigmatized group, regardless of how a participant responds. The interviewer wants to reduce noise as much as possible while maintaining enough noise to ensure participants respond truthfully. Unlike prior literature, I find that the incentives of a participant depend on the number of participants; therefore, the amount of noise required under randomized response decreases when the number of participants increases as adding respondents relaxes truth-telling constraints. However, adding respondents only relaxes incentive constraints to a limit, so some noise remains even when there are a large number of participants.

I improve upon the original randomized response technique in two ways in Chapter 2: "Eliciting Private Information using Correlation: A Modification of Random-

ized Response". In standard randomized response techniques, participants receive questions independently by using a randomization device such as a die. With my technique, participants receive perfectly correlated questions which reduces the variance of the surveyor's estimator while still protecting the privacy of the subjects. Unlike with the randomized response technique, adding correlation allows the surveyor to use a dominant strategy mechanism though it provides limited information. In addition to correlation, my technique provides the surveyor with private information on the distribution of questions asked. Because of the private information, participants become more uncertain of which question is more associated with the stigmatizing characteristic giving them a stronger incentive to respond truthfully.

My final chapter, Chapter 3 "A Response Technique with Dominant Strategies in Forced Responses", improves upon a randomized response technique commonly used in practice. In the forced response technique, a fraction of survey participants are directly asked whether they belong to the stigmatizing group while the remaining participants either simply state "yes" or "no" according to a privately observed command. Unlike the original randomized response technique, the surveyor must worry whether participants obey the command in addition to answering truthfully. Psychologically, participants may feel more inclined to disobey than to lie. Therefore, I design a technique where obeying the command is a dominant strategy by providing the surveyor with private information. The paper then discusses a more general response technique with private information and suggests restrictions on the mechanisms to ensure the surveyor does not have an incentive to try to "trick" respondents into believing they have more privacy protection than they actually do. The chapter concludes with a discussion on privacy measures.

CHAPTER 1

A Game Theoretic Analysis of the Randomized Response Technique

1.1 Introduction

Many surveys use randomized response techniques so subjects can avoid being classified as belonging to a stigmatized group such as drug users or cheaters in the classroom[1]. Instead of directly asking respondents a question, randomized response techniques require each respondent to privately and independently use a device that randomizes his or her question so an interviewer is uncertain whether the participant belongs to a stigmatized group regardless of the response. Respondents may wish to tell the interviewer the truth but want the interviewer to view them positively, or they may fear retribution from the justice system, insurance companies, or other entities.[2] With randomized response techniques, participants can honestly answer a survey question yet still have some measure of privacy protection. A researcher can use the noisy data of the survey to estimate the proportion of the relevant population that belongs to a stigmatized class. This paper models the randomized response technique as a game with an interviewer and multiple respondents. The interviewer determines the amount of noise in the survey, which should be high enough to induce truth-telling but as low as possible to elicit as much information as possible. In other words, there is a tension between privacy preservation and improving estimation.

In contrast to previous literature on the subject, which does not directly address incentive constraints or assumes the participants and interviewer know the "stigmatized proportion"[3], in this paper the interviewer and participants update their beliefs

[1]See Goodstadt and Gruson (1975) and Martínez Sánchez and Rosa García (2012) respectively.

[2]See Tourangeau and Yan (2007) for a comprehensive discussion on the evidence of such deception in surveys with sensitive topics.

[3]For example, Blume, Lai, and Lim (2013) assume the proportion is known to be one half.

about the proportion and the incentives of the participants are explicitly modeled. Since the researcher performs the survey to estimate the proportion, presuming the proportion is unknown seems more natural. In addition to reducing the variance of the proportion estimate, adding more participants relaxes incentive constraints since each respondent affects the interviewer's estimation of the proportion less, unlike the case where the proportion is known and the constraint is independent of the number of respondents.

In Warner's (1965) original model, respondents independently use a private randomization device to determine whether to respond with a yes or no to one of the following statements: "I belong to the stigmatized group." or "I do not belong to a stigmatized group" For example, suppose an instructor wants to know if students cheated on an exam in a course. Students may personally like the professor and feel guilty for lying. However, no student wants the professor to believe he cheated, whether he actually cheated or not; thus, students may lie when asked directly. Suppose the instructor gives every student a die along with the statements "I cheated at least once on an exam in this course." and "I never once cheated on an exam in the course.". Each student privately rolls the die and responds with a yes or no to the first statement if the die lands on any number except a five or six while the student responds to the second statement when a five or six is rolled.

With the randomization, students may respond honestly because the noise created by the die creates reasonable doubt about whether the student cheated, but it still allows the instructor to estimate the number of cheaters. Though the instructor could use an anonymous survey, some students may fail to answer, causing selection bias. Unlike with anonymity, with the randomized response technique the instructor can identify each student's response, eliminating the issue of selection bias. Though there are many variants of this technique, this paper focuses on the one developed by Warner (1965).

If a respondent answers with a "yes" and if the likelihood of the statement "I belong to the stigmatized group" is greater than one half, the interviewer's estimate of the stigmatized proportion increases causing the interviewer to believe all members of the population studied are more likely to belong to the stigmatized group, even those outside of the survey. Thus, participants exert an externality on each other with their responses. As more respondents are added, each respondent affects the belief of the interviewer less so the benefit to lying decreases, but the cost of lying remains the same. Therefore, when more participants are added, each one has a greater net incentive to tell the interviewer the truth.

Adding another participant improves the interviewer's estimation of the stigmatized proportion by reducing variance but also relaxes the incentive constraints; thus, the interviewer can reduce the amount of noise when adding more participants, improving the estimation further. Though in many cases respondents may not be able to verify the number of survey participants, in others, such as in an experimental economics laboratory or in the instructor-and-student case described above, the respondents may observe each other and thus know how many engaged in the survey. Moreover, a consequence of this result is that if truth-telling is optimal in the single respondent case, then it is also optimal for any number of respondents.

When the probability of the statement "I belong to the stigmatized group" is greater than one half, a "yes" response to *either* question from a participant always makes the interviewer more likely to believe that participant belongs to the stigmatized group than a "no" statement regardless of the number of participants. Without lying aversion, each participant would answer "no" to avoid being associated with the stigmatized group. Even as the number of subjects grows large, some noise may be required to induce participants to truthfully respond.

The next section of the paper discusses the related literature. Section 1.3 introduces the model. In Section 1.4, there is a discussion about why adding more respondents to a survey affects the incentives of each respondent regarding truth-telling. Additionally, I derive and compare the amount of noise required to induce a truthful equilibrium when $N = 1$ and $N = 2$. Section 1.5 characterizes the truthful equilibria under the model, and Section 1.6 analyzes the effect of adding more subjects and the incentives as N approaches infinity. In Section 1.7, I find sufficient conditions for a truth-telling equilibria with a more general model. The paper concludes in Section 1.8.

1.2 Related Literature

Since the seminal paper of Crawford and Sobel (1982) on cheap talk, most economists have come to view costless communication as a vital form of communication even if incentives are not perfectly aligned. Blume, Board, and Kawamura (2007) demonstrate that communication can be improved by using noisy channels.[4] Goltsman, Horner, Pavlov, and Squintani (2009) proved the noisy procedure introduced by Blume et al. (2007) maximizes welfare in the standard Crawford-Sobel model with the uniform distribution and quadratic preferences. The introduction of noise relaxes incentive constraints for the communicator as the party receiving a message cannot infer whether the message came in error or is the intended message. Even though noise sometimes improves communication in theory, noisy communication is rarely used in practice–but one notable exception is the randomized response survey technique.

Over the past few decades, many researchers have conducted surveys using the randomized response technique. The technique or variants thereof have also been used to measure illegal resource use in Uganda (Solomon, Jacobson, Wald, and Gavin, 2007), falsification of income tax reports (Musch, Bröder, and Klauer, 2001), employee theft (Wimbush and Dalton, 1997), and more. In the field of development economics, Karlan and Zinman (2012) use a variant of the randomized response technique to study how borrowers spend their microfinance loans. Unfortunately, whether the randomized response technique succeeds in providing a better measurement than the direct response technique, where an interviewer simply asks the question directly without noise, is not clear.

Though many papers study the effectiveness of the randomized response technique, few explain why the technique sometimes fails to work effectively. Umesh and Peterson (1991) give a detailed evaluation of the randomized response technique and

[4]Myerson (1991) proposed the idea earlier with a simple example involving carrier pigeons.

find mixed results regarding its success. Blume, Lai, and Lim (2013) run an experiment which induces lying aversion and stigmatization aversion monetarily which also provides mixed results as subjects increase truth-telling with randomized response but systematically deviate. Moshagen, Hilbig, Erdfelder, and Moritz (2014) also run an experiment where respondents roll a die and only receive compensation for rolling a target outcome; thus, many respondents lied, creating a stigmatized group, and the experimenters measured this group using the randomized response technique.

In order to improve surveys and data collection, economists and statisticians need to determine why randomized response sometimes fails and other times it seemingly succeeds. Whether the failure stems from not enough privacy protection, confusion on the part of the subjects, or poor execution is unclear. Regrettably, few models have been introduced to study the incentives of respondents to help resolve the issue. Most statisticians, with the notable exception of Ljungqvist (1993), have focused their research on the statistical properties of estimators while economists, with the exception of Blume et al. (2013), have overlooked the subject from a theoretical perspective. In order to understand the problems and benefits of the randomized response technique, a strong theoretical foundation is required. The small theoretical research which does address the incentive issue usually contains the following assumptions: respondents know the parameter of interest and can therefore measure the degree of privacy protection; the responses of other subjects should have no effect on each respondent's own incentives; and the beliefs of respondents are independent of their real world experience (their types in my model). After dropping these assumptions, this paper demonstrates the amount of noise required for a truthful equilibrium depends on the number of survey participants and the beliefs of the participants. As truthful equilibria often require a certain level of noise even with a large sample, a likely candidate for the failure of the technique is insufficient privacy protection though an empirical and experimental investigation is needed to verify whether that is the case or not.

1.3 Model

An interviewer wants to know the proportion, ν, of a population that belongs to a stigmatized group $\mathcal{S}$.[5] The interviewer recruits N respondents of the population to take a survey. Each respondent i observes a type $\theta_i \in \{s, t\}$ where s is the stigmatized type and t is the accepted type. Let $\mathcal{T}$ represent the accepted group of the population. Respondent types are drawn independently conditional on knowing the parameter ν. The experimenter has the uniform prior on ν. Ex-ante subjects also have a uniform prior on ν; however, subjects update their beliefs about the distribution given their type. The interviewer is not a member of the population being studied.[6]

In order for the interviewer to elicit information from respondent i, the respondent answers a question $q_i \in \{s, t\}$ where s corresponds to the question "Are you in $\mathcal{S}$?" and t corresponds to the question "Are you in $\mathcal{T}$?". The respondent privately uses a randomization device to determine which question to answer where the probability of being asked "Are you in $\mathcal{S}$?" is p_s and the probability of being asked "Are you in $\mathcal{T}$?" is $1 - p_s$ where $p_s \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$.[7] Respondents answer either question with $r_i \in \{y, n\}$. Let $r \in \{y, n\}^N$ be the vector of responses, R the sum of y responses, and R_{-i} the number of y responses excluding i's response.

A participant responds truthfully when $q_i = \theta_i$ and $r_i = y$ or $q_i \neq \theta_i$ and $r_i = n$. Otherwise, a partcipant responds dishonestly.

If subjects truthfully report their types, the probability of a "yes" response from subject i conditional on ν, $p(r_i = y|\nu)$, is the probability of i belonging to $\mathcal{S}$ and $q_i = s$, $p_s\nu$, plus the probability of i belonging to $\mathcal{T}$ and $q_i = t$, $(1-p_s)(1-\nu)$; thus,

$$p(r_i = y|\nu) = p_s\nu + (1 - p_s)(1 - \nu)$$

[5] Technically, the interviewer actually measures the probability a random member of the population belongs to the stigmatized group; however, for large samples the two are the same.

[6] Even if the interviewer belongs to the population being studied, the beliefs are formed at the end of the experiment and with a large number of participants, the initial beliefs of the interviewer should be irrelevant as long as the prior beliefs have full support.

[7] p_s cannot be zero or one as the question would no longer be random. When p_s is one half, the interviewer cannot infer any information about the parameter ν from the responses of the players.

Due to independence, the sum of "yes" responses of subjects other than i conditional on ν, $R_{-i}|\nu$, is a random variable that is binomially distributed with parameters $(p_s\nu + (1-p_s)(1-\nu), N-1)$. Throughout the paper, notation is abused with $p(y_i)$ or $p(s_i)$ instead of $p(r_i = y)$ or $p(\theta_i = s)$.

Given all the responses, the interviewer estimates ν with $\hat{\nu}(R; p_s)$ and forms beliefs of each respondent's type given all responses assuming the partcipants responded honestly. The interviewer's belief that respondent i belongs to group $\mathcal{S}$ is written as $\mu(s_i|r) = \mu(s_i|r_i, R_{-i})$. Since the goal of the interviewer is to estimate ν as well as possible, the interviewer wants p_s as close to either zero or one as possible while still ensuring subjects truthfully report.[8]

Respondents prefer honesty over lying but also want the interviewer to perceive them as belonging to the accepted group instead of the stigmatized group. Thus, each respondent's utility function is decreasing in $\mu(s_i|r)$. As in Blume et al. (2013), we use the following utility specification. Let $\mathcal{H} = \{(s,s,y),(t,t,y),(s,t,n),(t,s,n)\}$ be the honesty set. If $(\theta_i, q_i, r_i) \in \mathcal{H}$, then respondent i responded honestly. The payoff function for each respondent to be,

$$U_i(\theta_i, q_i, r) = \lambda \mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i) - \xi\mu(s_i|r_i, R_{-i})$$

where $\lambda, \xi \geq 0$ are parameters measuring, respectively, the degrees of lying aversion and stigmatization aversion and $\mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i)$ is an indicator function with the value of one if $(\theta_i, q_i, r_i) \in \mathcal{H}$ and zero otherwise.

For simplicity, the $\lambda \mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i)$ component of the utility function is referred to as the lying cost. The cost may be due to a psychological cost of lying, or the respondent may have the same ultimate goal as the researcher and thus wants the researcher to learn the true value of the parameter ν.[9] The other component of the utility function, $\mu(s_i|r_i, R_{-i})$, is referred to as the interviewer's belief of i. Each

[8] The payoff of the interviewer is not explicitly modeled, but the objective of the interviewer could be to minimize variance or entropy subject to a cost concern.

[9] One must be careful with the second interpretation as λ could depend on N which is not the case in this model.

respondent prefers to be associated with the accepted group due to an embarrassment from stigmatization or worries about possible persecution if the stigmatized characteristic is related to illegal activity.

A strategy for respondent i is a function $\sigma_i : \{\theta_i, q_i\} \mapsto \{y, n\}$. The belief of respondent i about the parameter ν is written as $\mu_i(\nu)$. The beliefs of the stigmatized types differ from those of the accepted type and the interviewer as the subjects use Bayesian updating once types are realized. The stigmatized respondents believe others are more likely to be stigmatized as they change their beliefs from $\mu_i(\nu) = \mathbb{1}(0 \leq \nu \leq 1)$ to $\mu_i(\nu|s)$. Finding the new belief only requires a simple application of Baye's rule:

$$\mu_i(\nu|s) = \frac{p(s|\nu)\mu_i(\nu)}{\int_0^1 p(s|\nu)\mu_i(\nu)d\nu} = \frac{\nu}{\int_0^1 \nu d\nu} = 2\nu$$

Using the same technique, one can verify that the accepted respondents have the following beliefs $\mu_i(\nu|t) = 2(1-\nu)$ after updating. Respondents could have access to more information about the parameter ν. For example, drug users likely know other drug users providing more observations other than just oneself. However, for the sake of simplicity, this paper only considers the case where all subjects update their beliefs based on their own type. In some cases, such as tax evasion, this assumption is not unreasonable.

The subjects fail to know the responses of the other subjects when choosing whether to answer "yes" or "no." Respondents behave under the premise that all others reveal their types truthfully. Assuming subjects are risk-neutral, the expected utility of subject i can be rewritten as,

$$\lambda \mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i) - \xi E_{R_{-i}}[\mu(s_i|r_i, R_{-i})|\theta_i]$$

The probability of $R_{-i}|\theta_i$ is given below.

$$\int_0^1 \binom{N-1}{R_{-i}} (p_s\nu + (1-p_s)(1-\nu))^{R_{-i}}((1-p_s)\nu + (1-\nu)p_s)^{N-1-R_{-i}}\mu_i(\nu|\theta_i)d\nu$$

The set of p_s that induces truth-telling depends on the parameters λ, ξ, and N and let $A(\lambda, \xi, N)$ represents that set, sometimes referred to as the truth-telling set. Clearly, $A(\lambda, \xi, N) \subseteq [0, 1]$ and $A(\lambda, \xi, N)$ is convex. When $A(\lambda, \xi, N) = [0, 1]$, then truth-telling is an equilibrium for the direct response technique ($p_s \in \{0, 1\}$). Since the interviewer wants p_s as close to zero or one as possible, a larger set is preferred to a smaller one.

1.4 Discussion and Examples

Most of the early literature on randomized response considered privacy measures instead of focusing on truth-telling constraints.[10] The privacy measures were functions of the interviewer's belief about a participant given a response. Though privacy measures may be useful for institutional policy and certainly relate to the incentives in a survey, they alone do not demonstrate whether participants honestly respond. Until Ljungqvist (1993), no paper formally modeled the utility functions with lying aversion or the truth-telling constraints of subjects. This paper also uses a utility model but also incorporates the underlying Bayesian process which determines the constraints as neither the respondents nor interviewer know ν.

Many research papers that analyze the randomize response techniques come from the statistics field and only consider privacy measures of the subjects, which are often a function of the unknown parameter, ν.[11] These models allow an analysis of the different degrees of privacy for each value of ν. If respondents know ν, then many of these measures of privacy protection with some privacy threshold requirement could be modified as specific utility functions and incentive constraints. However, in most cases, the respondents should also have uncertainty about ν; otherwise, the interviewer could simply ask the respondents their estimates of ν instead of enquiring about their personal information. A previous paper by Zaizai and Zankan

[10] For example Leysieffer and Warner (1976) use the following measure of jeopardy, $\frac{P(R|A^c)}{P(R|A)}$ (or its inverse) where $P(R|A) = p(r_i|s_i)$ and $P(R|A^c) = p(r_i|t_i)$.

[11] For example, Fligner, Policello, and Singh (1977) use a privacy measure that is a function of the unknown parameter ν: $\frac{1-\max\{\mu(s_i|y_i,\nu),\mu(s_i|n_i,\nu)\}}{1-\nu}$.

(2004) uses a privacy measure that appears similar the utility function of this paper. In their analysis and most previous analysis, the game theoretic aspect of surveys is ignored even though the responses of subjects indirectly affect the beliefs about other respondents through the estimation of ν.[12] Additionally, Zaizai and Zankan (2004) and many prior research papers fail to take into consideration that respondents with different types likely have different beliefs as would be expected in any model with Bayesian rationality.[13]

To illustrate the model in as simple as a manner of possible, consider the case of one respondent. The following four constraints must hold for participants to respond honestly:

$$U_1(s, q_s, y) \geq U_1(s, q_s, n) \iff \lambda - \xi\mu(s_1|y_1) \geq -\xi\mu(s_1|n_1)$$

$$U_1(s, q_t, n) \geq U_1(s, q_t, y) \iff \lambda - \xi\mu(s_1|n_1) \geq -\xi\mu(s_1|y_1)$$

$$U_1(t, q_t, y) \geq U_1(t, q_t, n) \iff \lambda - \xi\mu(s_1|y_1) \geq -\xi\mu(s_1|n_1)$$

$$U_1(t, q_s, n) \geq U_1(t, q_s, y) \iff \lambda - \xi\mu(s_1|n_1) \geq -\xi\mu(s_1|y_1)$$

Thus, for truth-telling to be optimal for the single respondent case, the conditions for every type can be combined to the following,

$$\xi\big|\mu(s_1|y_1) - \mu(s_1|n_1)\big| \leq \lambda$$

The left hand side of the inequality represents the benefit to lying and while the right hand side represents the cost to lying. Lying benefits the subjects by changing the beliefs of the interviewer. Consider $\mu(s_1|y_1)$ which can be rewritten as follows:

$$\mu(s_1|y_1) = \int_0^1 \mu(s_1, \nu|y_1)d\nu = \int_0^1 \mu(s_1|\nu, y_1)\mu(\nu|y_1)d\nu$$

[12] Ljungqvist (1993) discusses that subject's decisions depend on their beliefs about the behavior of others but fails to explicitly include it in his analysis. (p. 101)

[13] Respondents could also have different payoff functions which has been discussed by Ljungqvist (1993) (though not analyzed) (p. 102). However, the model in this paper omits such a scenario for simplicity.

$$= \int_0^1 \frac{\mu(s_1, y_1|\nu)}{p(y_1|\nu)} \mu(\nu|y_1) d\nu = \int_0^1 \frac{p_s \nu}{p_s \nu + (1-p_s)(1-\nu)} \mu(\nu|y_1) d\nu$$

To calculate the posterior belief of ν, note that $\mu(\nu|y_1) \propto \mu(y_1|\nu)\mu(\nu) = p_s\nu + (1 - p_s)(1 - \nu)$. After solving for the constant of integration, it is easy to verify that $\mu(\nu|y_1) = 2(p_s\nu + (1 - p_s)(1 - \nu))$. After inputting this equality into the derived expression above, one can observe that $\mu(s_1|y_1) = p_s$ and using the same method it can be demonstrated that $\mu(s_1|n_1) = 1 - p_s$. Thus, the truth-telling set for the one respondent case is

$$A(\lambda, \xi, 1) = \left[\frac{\xi - \lambda}{2\xi}, 1 - \frac{\xi - \lambda}{2\xi}\right]$$

When the researcher only interviews a single respondent, the truth-telling set with this model turns out to be the same as Blume et al. (2013) even though in their model ν is known by both the interviewer and respondents. They assumed that the probability of each type, ν, is exactly equal to half. In this model, the ex-ante probability is also equal to one half so the similarity of the result should not be too surprising.[14]

With two respondents, each respondents must consider the response of the other respondent before deciding to tell the truth as both responses affect the interviewer's belief about each respondent. As in the case of one respondent, the incentive constraints for all types can be combined. The incentive constraint for the first respondent in the case where $N = 2$ is written as follows,

$$\left|p(y_2|\theta_1)(\mu(s_1|y_1, y_2) - \mu(s_1|n_1, y_2)) + p(n_2|\theta_1)(\mu(s_1|y_1, n_2) - \mu(s_1|n_1, n_2))\right| \leq \frac{\lambda}{\xi}$$

The left hand side of the expression in the absolute value bars represents the expected belief differential when the second subject responds with a "yes" while the right hand side covers the case when the other respondent says "no." In the case of one respondent, the type of that subject is irrelevant, but with many subjects, the type of

[14]Blume et al. (2013) also look at other equilibria including those of mixed strategy. Astonishingly, even in their simple model, the game has a plethora of equilibria which often include lying. Though analyzing all the equilibria of the game induced by the randomized response survey with many respondents is of interest, this paper only focuses on the truth-telling equilibria as they are the most studied and usually be the ones of interest for practical purposes.

the respondent may matter since it changes beliefs on the likelihood of the responses of other subjects. The probability of a “yes” response from the other subject given $\theta_1 = s$ is calculated as follows:

$$p(y_2|s_1) = \int_0^1 p(y_2, \nu|s_1)d\nu = \int_0^1 p(y_2|\nu)\mu_1(\nu|s)d\nu$$

$$= \int_0^1 (p_s\nu + (1-p_s)(1-\nu))2\nu d\nu = \frac{1}{3}(1+p_s)$$

The calculation of $p(n_2|s_1)$ can be done in a similar manner or one can use the relation $p(n_2|s_1) + p(y_2|s_1) = 1$. To calculate the belief of the interviewer when $r_1 = r_2 = y$, rewrite the belief using the definition of a conditional distribution:

$$\mu(s_1|y_1, y_2) = \frac{p(y_1, y_2, s_1)}{p(y_1, y_2)}$$

The denominator is calculated below using the substitution $\tau = p_s\nu + (1-p_s)(1-\nu)$:

$$p(y_1, y_2) = \int_0^1 p(y_1, y_2, \nu)d\nu = \int_0^1 p(y|\nu)^2 d\nu = \int_0^1 (p_s\nu + (1-p_s)(1-\nu))^2 d\nu$$

$$= \frac{1}{2p_s - 1}\int_{p_s}^{1-p_s} \tau^2 d\tau = \frac{p_s^2 - (1-p_s)^2}{2p_s - 1} = \frac{1}{3}(p_s^2 - p_s + 1)$$

The numerator is calculated below:

$$p(y_1, y_2, s_1) = \int_0^1 p(y_1, s_1|\nu)p(y_2|\nu)d\nu = \int_0^1 p_s\nu(p_s\nu + (1-p_s)(1-\nu))d\nu$$

$$= \frac{1}{6}p_s(1+p_s)$$

After calculating the other sets of beliefs in the same fashion,

$$A(\lambda, \xi, 2) = \left\{p_s : \left|\frac{(2p_s - 1)(2-p_s)(1+p_s)}{2(p_s^2 - p_s + 1)(2p_s - 2p_s^2 + 1)}\right| \le \frac{\lambda}{\xi}\right\}$$

There are a few important distinctions to note in the two respondent example

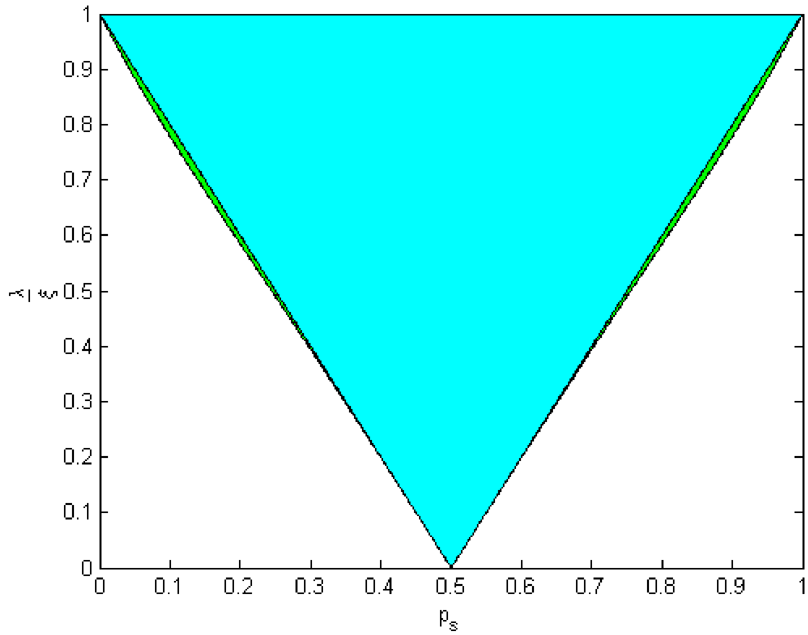

Figure 1.1: The blue area represents the set of parameter configurations of p_s, λ and ξ that permit truth-telling with a single respondent, $A(\lambda, \xi, 1)$. The union of the blue and green area represents the set of parameter configurations that permit truth-telling with two respondents, $A(\lambda, \xi, 2)$.

compared to the one respondent case. The constraints for the two respondent case are more relaxed than when there is only one respondent; therefore, the truth-telling set is larger. As the graph indicates, the sets hardly differ when p_s is close to zero, one, or half. It also turns out the truth-telling set is the same for $\theta_1 = t$; thus, the truth-telling set is type independent for the case where $N = 2$. Finally, the boundary defining the set becomes significantly messier when only one more respondent is added even though the utility function is not particularly complicated indicating that studying constraints with a very general utility model is likely intractable.

1.5 Characterization of Truth-Telling Equilibria

Though many other papers ignore the effect responses of other subjects on truth-telling incentives when studying the RRT which may seem reasonable since types are conditionally independent, as Example 1 demonstrates, responses of other subjects clearly affect incentives. When a subject answers the interviewer's question, the response has two effects, a direct effect and an indirect effect. The direct effect is the effect on the belief of the interviewer on the subject, $\mu(s_i|r)$, as if the parameter is known; therefore, if $p_s > \frac{1}{2}$, the interviewer always believes a "yes" response indicates a higher likelihood of a stigmatized type than a "no" response indicates. The responses of other participants are irrelevant for the direct effect. The indirect effect is the affect of a participant's response on the interviewer's belief of ν. If $p_s > \frac{1}{2}$, "yes" responses from all subjects increase the beliefs of the experimenter about every respondent belonging to the stigmatized group as the experimenter believes ν is larger while "no" responses from other subjects increase the beliefs of the experimenter about every respondent belonging to the accepted group as the experimenter believes ν is smaller in this case.

$$\mu(s_i|y_i, R_{-i}) = \int_0^1 \underbrace{\frac{p_s\nu}{p_s\nu + (1-p_s)(1-\nu)}}_{\text{Direct Effect}} \underbrace{\mu(\nu|R_{-i}+1)}_{\text{Indirect Effect}} d\nu$$

$$\mu(s_i|n_i, R_{-i}) = \int_0^1 \underbrace{\frac{(1-p_s)\nu}{(1-p_s)\nu + p_s(1-\nu)}}_{\text{Direct Effect}} \underbrace{\mu(\nu|R_{-i})}_{\text{Indirect Effect}} d\nu$$

If $p_s < \frac{1}{2}$, the roles of "yes" and "no" are simply reversed.

Theorem 1.1. *The set of probabilities of "Are you in $\mathcal{S}$?", p_s, which induce truth-telling as an equilibrium is characterized below:*

$$A(\lambda, \xi, N) = \left\{ p_s : \left| 2E_{R_{-i}}\mu(s_i|y_i, R_{-i}) - 1 \right| \leq \frac{\lambda}{\xi} \right\}$$

Proof: All proofs in this chapter are contained in Section 1.9.

Clearly, the responses of other subjects play a role in the incentives of each subject according to Theorem 1.1 as set p_s depends on N. Surprisingly, the set of truth-telling equilibria for both types of respondents is the same. Types fail to affect the incentives because of the use of a uniform prior. In fact, any symmetric beta prior would provide the same result as Theorem 1.1 since $P(R_{-i}|s) = P(N-1-R_{-i}|t)$ and $\mu(s_i|R_{-i}, y_i) = \mu(t_i|N-1-R_i, n_i)$. Inserting these identities into the constraint and simplifying gives the inequality described in Theorem 1.1.

Corollary 1.1. *Truth-telling is an equilibrium for all values of* p_s*,* $A(\lambda, \xi, N) = [0,1]$*, if* $\frac{\lambda}{\xi} \geq 1$*.*

When lying costs are high enough relative to stigmatization aversion, using a direct response technique where $p_s \in \{0, 1\}$ is clearly optimal since subjects always have an incentive to tell the truth. Any model analyzing the randomized response technique should have this property that using the direct response technique may be optimal. In most real world surveys, the direct response technique is used since responses often do not have a strong stigmatization associated with them. When a marketing company asks someone whether chicken or beef is preferred, neither response is likely to carry strong stigmatization. Additionally, for groups with large lying costs such as nuns or children, the direct response technique may be optimal. Ultimately, under which topics lying costs are high relative to stigmatization costs and whether different groups of individuals have different lying cost is an empirical and experimental question.

This paper only focuses on equilibria where subjects tell the truth; however, such equilibria are not necessarily unique and may not even provide the most information. Greenberg, Abul-Ela, Simmons, and Horvitz (1969) assume there is a known portion of subjects who lie to the interviewer when estimating the parameter of interest, but in many cases such an assumption seems unreasonable. Hout, Böckenholt, and Van Der Heijden (2010) use two samples, one with direct questioning and one with randomized response, in order to estimate the number of cheaters in a population. Blume et al. (2013) find that in a survey with one individual it may be optimal to allow for lying instead of focusing on truthful equilibria. Most of the research that uses the randomized response technique in practice assumes truth-telling; therefore, this paper only focuses on truthful equilibria even though there may be other equi-

Table 1.1: Boundaries of Truth-telling Sets

		p_s = .01	.1	.25	.4
N	1	.9800	.8000	.5000	.2000
	10	.9583	.7051	.4386	.1930
	20	.9500	.6788	.4109	.1868
	30	.9455	.6671	.3968	.1816
	40	.9425	.6602	.3883	.1774
	50	.9404	.6557	.3826	.1738

libria and some of the other equilibria may provide more information than truthful equilibria.

The boundary of the truth-telling set (when a subject is indifferent between lying and telling the truth) is numerically calculated using Matlab for certain values of N and p_s in Table 1.1. In other words, the table provides the minimum ratio of λ over ξ for which truth-telling is an equilibrium for a given N and p_s. Numerical results suggest that the set increases as N increases. Additionally, the absolute rate at which the set increases appears to be decreasing; thus, the set looks as if it approaches a limiting set that is unequal to the unit interval. Both of these conjectures suggested by the numerical results are proved in the next section.

1.6 Comparative Statics

Since the truth-telling constraint is not independent of N, a researcher may need to know the effect of increasing or decreasing the number of respondents on incentives. Additionally, a researcher could be interested in what occurs when a sample grows large (N tends toward infinity) in order to approximate incentives for groups with numerous respondents. If the boundary of $A(\lambda, \xi, N)$ is decreasing or non-monotonic in the number of respondents, then it may not be optimal to use more respondents in a survey as p_s would need to be closer to half in order to ensure truth-telling, which would cause the data to be noisier. Since adding respondents is typically costly, it is important to know whether adding more respondents provides enough additional information to warrant their addition.

Proposition 1.1. *The set of probabilities of "Are you in $\mathcal{S}$?", p_s, which induces a truth-telling equilibrium with sample size N is contained in the set which induces a truth-telling equilibrium of the sample size $N+1$, $A(\lambda, \xi, N) \subset A(\lambda, \xi, N+1)$.*

Proposition 1.1 states that increasing the sample size allows the researcher to reduce the amount of noise in the survey in addition to reducing variance of the proportion estimate. Since the researcher wants to estimate ν in as small an interval as possible, choosing p_s as close to zero or one is ideal as long as respondents truthfully respond. When the researcher uses more respondents, the expected change of the researcher's beliefs about a respondent's type from a "yes" or "no" answer decrease. With many participants, the researcher's beliefs about any individual subject are largely affected by the responses of the other subjects (through the indirect effect), and participants ignore the informational externality imposed on others from a "yes" or "no" response. Thus, the benefit to lying decreases as more subjects are added; however, the cost of lying (λ) remains the same. Therefore, the truth-telling constraint is less binding when the interviewer surveys more people allowing the interviewer to decrease p_s if more subjects are added.

Interviewers clearly have an incentive to lie to respondents about the number of participants in a survey as subjects are more inclined to tell the truth in surveys with more subjects. Thus, if a researcher can interview all the subjects simultaneously, then respondents would know the number of survey participants and behave accordingly. When a researcher cannot interview all subjects at the same time, then finding a p_s that guarantee truth-telling for all values of N is important.

Corollary 1.2. *If $p_s \in \left[\frac{\xi-\lambda}{2\xi}, 1-\frac{\xi-\lambda}{2\xi}\right]$, then respondents always have an incentive to tell the truth for any sample size, N.*

Since the truth-telling set expands in the number of participants, then simply choosing a p_s that yields truth-telling for a single respondent guarantees it holds for any larger number of participants. If a researcher is uncertain of the number of

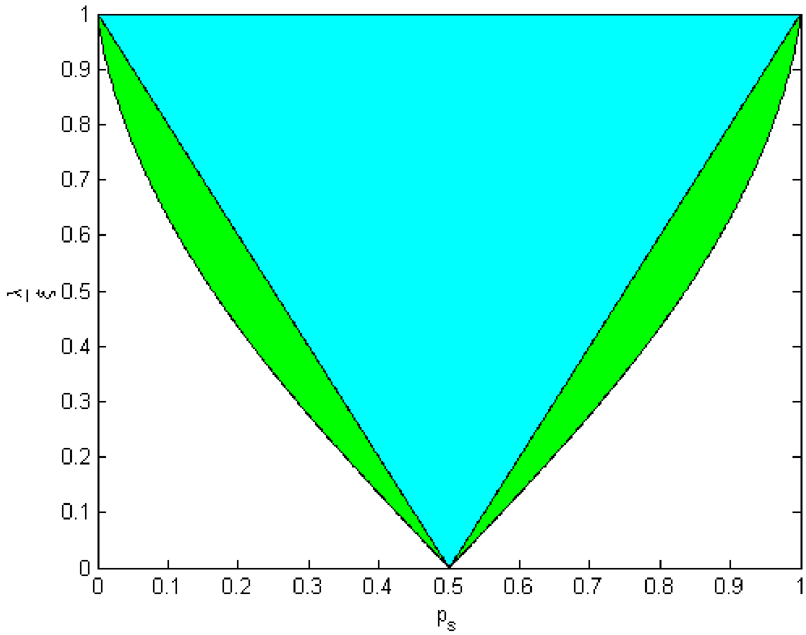

Figure 1.2: The blue area represents the set of parameter configurations of p_s, λ, and ξ that permit truth-telling with a single respondent, $A(\lambda, \xi, 1)$. The union of the blue and green area represents the set of parameter configurations that permit truth-telling with an infinite number of respondents, $\lim_{N\to\infty} A(\lambda, \xi, N)$.

respondents in a survey and wants to take a cautious approach, then using the p_s that induces truth-telling in the single respondent case may be the best approach.

Proposition 1.2. *If $\frac{\lambda}{\xi} < 1$, the direct response technique, $p_s \in \{0, 1\}$, always fails to induce truth-telling regardless of sample size, N (some noise is required to induce truth-telling). The set of probabilities of "Are you in $\mathcal{S}$?", p_s, that induce truth-telling as the sample size, N, approaches infinity, $\lim\limits_{N\to\infty} A(\lambda, \xi, N)$, equals*

$$\left\{ p_s : \left| 2\left(\frac{p_s}{2p_s - 1} - \frac{p_s(1-p_s)}{(2p_s-1)^2} \big(log(p_s) - log(1-p_s) \big) \right) - 1 \right| \le \frac{\lambda}{\xi} \right\}$$

As the number of subjects approaches infinity, the effect of the response of a participant on the estimation of the parameter ν is negligible. Thus, the participant only worries about the direct effect of the response on the interviewer's belief as the indirect effect is zero. Therefore, the truth-telling constraint only depends on the expected belief about the participant given a "yes" response. Since the interviewer knows ν in the limit, the belief about a response given "yes" is $\frac{p_s\nu}{p_s\nu+(1-p_s)(1-\nu)}$. Since the respondent fails to know the value of ν, the expected value of $\frac{p_s\nu}{p_s\nu+(1-p_s)(1-\nu)}$ depends on the prior of the respondent. Integrating $\frac{p_s\nu}{p_s\nu+(1-p_s)(1-\nu)}$ with respect to ν and inserting it into the constraint gives the inequality listed in Proposition 1.2.

Thus, unfortunately, p_s cannot arbitrarily approach zero or one as more respondents are added because the truth-telling set fails to approach the unit interval as N grows large since the direct affect always remains. In other words, unless the condition in Corollary 1.1 is met, the interviewer always needs to keep some level of noise even when N is large.

The effect of adding more respondents expands the truth-telling set the most when the amount of noise is intermediate. Intuitively, this occurs as the interviewer's belief about a subject remains close to zero or one when the amount of noise is small and close to half when the amount of noise is large. Thus, when a researcher is considering whether to add more respondents or increase the amount of noise to ensure the truth-telling, the benefit of adding a respondent is greater in the case of intermediate noise than the case when the amount of noise is small. When the amount of noise is small, adding more noise may be ideal.

Experimenters should use the limiting distribution as a measure for incentives with caution. Adding noise slows the rate of convergence as indicated by Table 1.1 in the previous section. Because of the slow rate of convergence, researchers not only need to be careful with worrying about the limiting incentives, but also must be careful not to use estimates of the parameter of interest, ν which depend on it. Thus, using the Bayesian approach studied by Winkler and Franklin (1979) is particularly appropriate with very noisy data.

1.7 A More General Model

Though the model incorporates the effect of the responses of other subjects and allows respondents to update beliefs unlike prior models, it contains two major drawbacks: the utility function is not very general and the initial beliefs of the subjects are uniform before they update their type. As stated in a footnote in Section 1.3, the uniform prior is not so restrictive for the experimenter as her beliefs are mostly formed from the data ex-post when the sample size is large, though sufficient conditions are also found for general beliefs of the experimenter. These two drawbacks are a result of simplification for the sake of tractability; however, in this section, a general payoff specification and different beliefs are considered.

With some extremely sensitive topics, such as illegal activity, subjects may be more concerned with having $\mu(s_i|r)$ close to one as the subject may only be concerned with maintaining some sort of "reasonable doubt." To account for this possibility and others, the expected payoff function is changed from a linear model to a more general model:

$$U_i(\theta_i, q_i, r) = \lambda \mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i) - g(\mu(s_i|r_i, R_{-i}))$$

In order to focus on the effect of the non-linearity, the ex-ante beliefs of the subjects are uniform and subjects update by type as before so $\mu_i(\nu|s) = 2\nu$ and $\mu_i(\nu|t) = 2(1-\nu)$. Additionally, assume g is strictly increasing, differentiable, and $g(0) = 0$. The set of p_s which induces truth-telling equilibria in this more general case is written as $A(\lambda, g, N)$.

Unfortunately, Theorem 1.1 and Proposition 1.1 fail to hold with this more general model.[15] For example, suppose $g(x) = x^3$. When $N = 1$ and $p_s = .01$, truth-telling is an equilibrium if $\lambda \geq .9703$ for both the s and t types, but when $N = 3$ and $p_s = .01$, truth-telling is an equilibrium if $\lambda \geq .9708$ for the s type and if $\lambda \geq .9465$ for the t type. Clearly, Theorem 1.1 fails to hold as truth-telling condition now depends on a respondent's type. Additionally, since the tightest

[15] Proposition 1.2 fails to hold in its current form, but it can easily be amended to allow for different payoff specifications.

constraint determines the equilibrium set, adding more respondents fails to expand the equilibrium set when $\lambda \in (.9703, .9708)$, which means that Proposition 1.1 also fails to hold. When $g(x) = x^3$, then the payoff is convex. Thus, respondent's suffer more when from a change in beliefs from 0.9 to 1.0 than 0 to 0.1. In other words, changing the experimenter's beliefs by lying benefits the respondent more in higher regions than in lower regions. This effect is referred to as the "convex effect." As in the linear case, each respondent has less of an effect on the belief of the interviewer when more participants are added. However, when respondent i is of the stigmatized type, he believes that the responses of others are likely to cause the interviewer to believe ν is larger causing the expected value of $\mu(s_i|r)$ to increase. Since the stigmatized respondent believes he is more likely to be in higher regions when $N = 3$ than when $N = 1$, the convex effect is stronger. Even though the indirect effect decreases as more respondents are added, the convex effect increases and its magnitude is stronger when $g(x) = x^3$ and one moves from $N = 1$ to $N = 3$.

Even though Theorem 1.1 and Proposition 1.1 do not hold in the general case, for any function g in the general case, one can find a corresponding ξ where the truth-telling set in the linear case, $A(\lambda, \xi, N)$, is a subset of the truth-telling set in the general case with g, $A(\lambda, g, N)$.

Theorem 1.2. *Suppose the stigmatization aversion in the linear case is set equal to the largest possible marginal change in the payoff function in the general case, $\xi = \max_{x \in [0,1]} g'(x)$. Then, any noise that induces a truth-telling equilibrium in the linear case also induces a truth-telling equilibrium in the general case, $A(\lambda, \xi, N) \subseteq A(\lambda, g, N)$.*

Theorem 1.2 is intuitive as the payoffs are simply linearized using the worst case scenario. In fact, the proof is simply an application of the mean value theorem. Unfortunately, the interviewer is unlikely to know g, so finding the optimal p_s could be difficult. In cases such as illegal activity, convex functions seem rather natural since respondents should suffer more when one eliminates reasonable doubt. Fortunately, finding sufficient conditions for a truth-telling equilibrium is simple in the case when the harm caused by the beliefs is convex.

Corollary 1.3. *Suppose a respondent's harm caused by the interviewer's beliefs, g, is convex (concave). Then, the set of probabilities of "Are you in S?", p_s, that induce a truth-telling equilibrium in the linear case using the marginal harm to the respondent when the belief of the interviewer is one (zero), $g'(1)$ ($g'(0)$), as the measure of stigmatization aversion, ξ, is a subset of the set that induces truth-telling, $A(\lambda, \xi, N)$, in the general case. Thus, $A(\lambda, \xi, N) \subseteq A(\lambda, g, N)$ when $\xi = g'(1)$ ($\xi = g'(0)$).*

Using the assumption of convexity or concavity, the interviewer only needs to know the ratio of lying cost to the marginal pain at one of the endpoints, $\frac{\lambda}{g'(1)}$ or $\frac{\lambda}{g'(0)}$, to ensure truth-telling is an equilibrium. Additionally, if g is concave and then convex, one only needs to know the larger of the two endpoints.

Corollary 1.4. *Suppose the harm caused by the beliefs, g, is concave then convex, i.e. $\exists x^* : g$ is concave if $x \in [0, x^*]$ and g is convex if $x \in [x^*, 1]$. The set of probabilities of "Are you in S?", p_s, which induce a truth-telling equilibrium of the linear case when the stigmatization aversion is equal to the largest marginal harm at the endpoints, $\xi = max\{g'(0), g'(1)\}$, is a subset of the set of probabilities in the general case, $A(\lambda, \xi, N) \subseteq A(\lambda, g, N)$.*

The concave-convex functions are natural candidates for the harm function, g. Suppose the interviewer asks the respondent whether he is a thief or not. The respondent may feel strong embarrassment from any association with theft, which means that the harm from the interviewers beliefs, $\mu(s_i|r)$ moving from 0.0 to 0.1 may be worse than the beliefs moving from 0.4 to 0.5; however, the respondent also may fear a criminal investigation if the belief, $\mu(s_i|r)$ is high enough so the harm from the interviewer's beliefs moving from 0.9 to 1.0 may be significantly worse than moving from 0.4 to 0.5. Even if the interviewer is uncertain of whether the harm function is concave or convex (but she know it is one of the two), she can use Corollary 1.4 to find sufficient conditions for a truth-telling equilibrium.

Of course, the beliefs of the experimenter or the subject could be different from those specified in the model before. If the experimenter wants to take a conservative approach, she may want to find sets of beliefs for which truth-telling under these beliefs is a sufficient condition for truth-telling for any possible set of beliefs.

Theorem 1.3. *Suppose the experimenter believes that $\nu = \frac{1}{2}$ with certainty and truth-telling is an equilibrium given p_s, λ, and ξ. Then, truth-telling with the parameters p_s, λ, and ξ is also an equilibrium for any number of subjects, any set of beliefs for the respondents, and any other set of beliefs for the experimenter (N, μ_1,..., μ_N,μ).*

If the experimenter believes the proportion is one half with certainty, the beliefs of the respondents fail to affect incentives. In the original model, beliefs of the respondents affected incentives because participants expected the beliefs of the interviewer to converge to their beliefs. However, since the prior of the experimenter places all the mass on one point, she ignores any evidence, which explains why truth-telling does not depend on N in Theorem 1.3. Given the beliefs of the interviewer, the indirect effect described in previous sections does not exist, which is an intuitive reason why these beliefs are sufficient for all others in regards to truth-telling.

Of course, the indirect effect disappears with any prior that places all its mass on one point. In order to understand why placing all the mass at half is a sufficient condition, one must consider placing all the mass at the other extremes, zero and one. In these extreme cases, the responses of the participants have no effect on the belief of the experimenter. When the proportion is one half, even though the respondents cannot affect the belief of the interviewer about ν, a participant's response affects the belief of the interviewer about that particular respondent.

The results of this section indicate that the simple model used by Blume et al. (2013) give fairly sufficient conditions for a truth-telling equilibrium with the appropriate chosen parameters despite some of its drawbacks.

1.8 Conclusion

When surveys are conducted, subjects place an informational externality on each other since each response allows an interviewer to estimate the parameter of interest better. Other research papers that study the randomized response usually explicitly or implicitly have the assumption that either there are so many subjects that one can use limiting constraints, there is only one subject, or subjects already know the parameter of interest. The difference in incentives mirrors the differences in incentives under Cournot versus monopoly or Cournot perfect competition. One could view the model presented by Blume et al. (2013) as analogous to the monopoly case since they deeply analyze the case with only one respondent, and the limiting case in this model as perfectly competitive since there are so many subjects that the effect of a single respondent on the estimation of ν is negligible. As in the Cournot model, subjects take into account the actions of other respondents since the answers of others impact the beliefs of the interviewer about every subject.

Whether subjects actually behave differently with different numbers of subjects is an empirical and experimental question. One could look at past surveys using the randomized response technique on a particular topic where the number of participants was disclosed to the subjects and the number of subjects varied between the surveys to study whether more truth-telling occurred when more respondents were used. Alternatively, one could run an experiment with a sensitive questions using the same p_s for each trial by varying N to test whether the subjects are more honest as N increases.

Many institutions use surveys to determine their policies; thus, surveys play a vital role in society and interpreting survey data correctly is important. Numerous survey techniques have been developed to encourage subjects to tell the truth when they may not under direct questioning but most papers focused on the statistical properties of estimators instead of the incentives. More theoretical, empirical, and experimental research on survey incentives needs to be conducted to ensure any policy developed or funding determined through the use of a survey is appropriate.

1.9 Proofs and Elaborations

Proof of Theorem 1.1: Given our model, the following condition must hold in equilibrium,

$$(1)\left|E_{R-i}\left[\mu(s_i|y_i, R_{-i}) - \mu(s_i|n_i, R_{-i})|\theta_i\right]\right| \leq \frac{\lambda}{\xi}$$

The term inside the absolute value bars will be referred to as the expected belief differential. First consider the case where $\theta_i = s$. Note that

$$P(R_{-i}|s_i; N-1) = \int_0^1 P(R_{-i}, \nu|s; N-1)d\nu = \int_0^1 P(R_{-i}|\nu; N-1)\mu_i(\nu|s)d\nu$$

Conditional on ν, R_{-i} is distributed binomially with parameters $N-1$ and $p_s\nu + (1-p_s)(1-\nu)$. Additionally, the belief of respondent i conditional on her type is $\mu_i(\nu|s) = 2\nu$. Therefore, the equation above can be written as

$$\int_0^1 \binom{N-1}{R_{-i}} [p_s\nu + (1-p_s)(1-\nu)]^{R_{-i}} [(1-p_s)\nu + p_s(1-\nu)]^{N-1-R_{-i}} 2\nu d\nu$$

When $\theta_i = t_i$, $\mu_i(\nu|t) = 2(1-\nu)$, using the same process, $P(R_{-i}|t; N-1)$ can be rewritten as

$$\int_0^1 \binom{N-1}{R_{-i}} [p_s\nu + (1-p_s)(1-\nu)]^{R_{-i}} [(1-p_s)\nu + p_s(1-\nu)]^{N-1-R_{-i}} 2(1-\nu) d\nu$$

Note that $P(R|s; N-1) = P(N-1-R|t; N-1)$ due to the symmetry of the beta function. Again, by symmetry, $\mu(s_i|R_{-i}, n_i) = \mu(t_i|N-1-R_{-i}, y_i)$. Given these identities and assuming $\theta_i = s$, the expected belief differential can be rewritten,

$$\sum_{R_{-i}=0}^{N-1} P(R_{-i}|s)\mu(s_i|R_{-i}, y_i) - P(R_{-i}|s)\mu(s_i|R_{-i}, n_i)$$

$$= \sum_{R_{-i}=0}^{N-1} P(R_{-i}|s)\mu(s_i|R_{-i}, y_i) - \sum_{R_{-i}=0}^{N-1} P(N-1-R_{-i}|t)\mu(t_i|N-1-R_{-i}, y_i)$$

Reordering the second summation gives the following:

$$\sum_{R_{-i}=0}^{N-1} P(R_{-i}|s)\mu(s_i|R_{-i}, y_i) - \sum_{R_{-i}=0}^{N-1} P(R_{-i}|t)\mu(t_i|R_{-i}, y_i)$$

$$= \sum_{R_{-i}=0}^{N-1} P(R_{-i}|s)\mu(s_i|R_{-i}, y_i) - \sum_{R_{-i}=0}^{N-1} P(R_{-i}|t)(1 - \mu(s_i|R_{-i}, y_i))$$

$$= \sum_{R_{-i}=0}^{N-1} (P(R_{-i}|s) + P(R_{-i}|t))\mu(s_i|R_{-i}, y_i) - 1$$

Note that $P(R_{-i}|s) + P(R_{-i}|t) = 2P(R_{-i})$ which simplifies the expected belief differential giving the constraint,

$$\left|2E_{R-i}\big[\mu(s_i|R_{-i}, y_i)\big] - 1\right| \leq \frac{\lambda}{\xi}$$

It is easy to verify that the same holds for $\theta_i = t$.

Proof of Proposition 1.1: In order to prove the optimal p_s is closer to zero or one, it suffices to show that the constraint weakens as the number of respondents increases. WLOG, suppose $p_s > \frac{1}{2} \implies E_{R_{-i}}[\mu(s_i|y_i, R_i)] > E_{R_{-i}}[\mu(s_i|n_i, R_i)]$. The preceding inequality holds since the respondent is more likely to be asked "Are you in $\mathcal{S}$?"; thus, the interviewer is more likely to believe the respondent belongs to the stigmatized group given a "yes" response. Additionally, $E_{R_{-i}}[\mu(s_i|y_i, R_i)] > \frac{1}{2}$ for the same reason. The truth-telling constraint can be rewritten without the absolute value bars with the assumption $p_s > \frac{1}{2}$. Therefore, the following must hold for the constraint for $N+1$ respondents to be less than the constraint for N respondents

$$2E_{R_{-i}}[\mu(s_i|y_i, R_i); N] - 1 - \big(2E_{R_{-i}}[\mu(s_i|y_i, R_i); N-1] - 1\big) < 0$$

$$\iff E_{R_{-i}}[\mu(s_i|y_i, R_i); N-1] - E_{R_{-i}}[\mu(s_i|y_i, R_i); N-2] < 0$$

For notational convenience, let y_N be the response of the nth subject if she answers "yes" and n_N be the response otherwise. The beliefs of the interviewer when there are only $N-1$ other subjects can be rewritten as,

$$\mu(s_i|y_i, R_{-i}) = \mu(s_i, y_N|y_i, R_{-i}) + \mu(s_i, n_N|y_i, R_{-i})$$

$$= p(y_N|R_{-i}, y_i)\mu(s_i|y_i, R_{-i}, y_N) + p(n_N|R_{-i}, y_i)\mu(s_i|y_i, R_{-i}, n_N)$$

Using this equality, the difference of the two expectations can be rewritten as

$$\sum_{R_{-i}=0}^{N-1} P(R_{-i}, y_N)\big((1-p(y_N|R_{-i}, y_i))\mu(s_i|y_i, R_{-i}, y_N) - p(n_N|R_{-i}, y_i)\mu(s_i|y_i, R_{-i}, n_N)\big)$$

$$+P(R_{-i}, n_N)\big((1-p(y_N|R_{-i}, y_i))\mu(s_i|y_i, R_{-i}, y_N) - p(n_N|R_{-i}, y_i)\mu(s_i|y_i, R_{-i}, n_N)\big)$$

The probability $p(y_N|R_{-i}, y_i)$ can be rewritten as $1-p(n_N|R_{-i}, y_i)$ and $p(n_N|R_{-i}, y_i)$ can be rewritten as $1-p(y_N|R_{-i}, y_i)$. Applying these identities to the equation above yields

$$\sum_{R_{-i}=0}^{N-1} P(R_{-i}, y_N)p(n_N|R_{-i}, y_i)\big(\mu(s_i|y_i, R_{-i}, y_N) - \mu(s_i|y_i, R_{-i}, n_N)\big)$$

$$+\sum_{R_{-i}=0}^{N-1} P(R_{-i}, n_N)p(y_N|R_{-i}, y_i)\big(\mu(s_i|y_i, R_{-i}, n_N) - \mu(s_i|y_i, R_{-i}, y_N)\big)$$

$$=\sum_{R_{-i}=0}^{N-1} P(R_{-i}, y_N)p(n_N|R_{-i}, y_i)\big(\mu(s_i|y_i, R_{-i}, y_N) - \mu(s_i|y_i, R_{-i}, n_N)\big)$$

$$-\sum_{R_{-i}=0}^{N-1} P(R_{-i}, n_N)p(y_N|R_{-i}, y_i)\big(\mu(s_i|y_i, R_{-i}, y_N) - \mu(s_i|y_i, R_{-i}, n_N)\big)$$

$$<\sum_{R_{-i}=0}^{N-1} P(R_{-i}, y_N)p(n_N|R_{-i}, y_i)\big(\mu(s_i|y_i, R_{-i}, y_N) - \mu(s_i|y_i, R_{-i}, n_N)\big)$$

$$-\sum_{R_{-i}=0}^{N-1} P(R_{-i}, n_N)p(y_N|R_{-i}, n_i)\big(\mu(s_i|y_i, R_{-i}, y_N) - \mu(s_i|y_i, R_{-i}, n_N)\big)$$

The last inequality come from the fact that $p(y_N|R_{-i}, y_i) > p(y_N|R_{-i}, n_i)$ which holds since $\mu(\nu|R_{-i}, y_i) > \mu(\nu|R_{-i}, n_i)$ (remember $p_s > \frac{1}{2}$). Given that respondents are treated symmetrically, the expression above can be rewritten as:

$$\sum_{R_{-i}=0}^{N-1} P(R_{-i}, y_N)p(n_i|R_{-i}, y_N)\big(\mu(s_i|y_i, R_{-i}, y_N) - \mu(s_i|y_i, R_{-i}, n_N)\big)$$

$$- \sum_{R_{-i}=0}^{N-1} P(R_{-i}, n_N)p(y_i|R_{-i}, n_N)\big(\mu(s_i|y_i, R_{-i}, y_N) - \mu(s_i|y_i, R_{-i}, n_N)\big)$$

$$= 0$$

Proof of Proposition 1.2: The following three equations are used to in the proof.

$$\text{a)} p_s E[\nu|R_{-i}] = p_s \int_0^1 \nu\mu(\nu|R_{-i})d\nu = \int_0^1 \mu(s_i, y_i, \nu|R_{-i})d\nu = \mu(s_i, y_i|R_{-i})$$

$$\text{b)} \tau = p(y_i|\nu) \implies E[\tau|R_{-i}] = \int_{1-p_s}^{p_s} \tau\mu(\tau|R_{-i})d\tau = p(y_i|R_{-i})$$

$$\text{c)} E[\nu|R_i] = \frac{E[\tau|R_{-i}] - (1-p_s)}{2p_s - 1}$$

The only portion of the truth-telling constraint which depends on N is $\mu(s_i|y_i, R_{-i})$.

$$\lim_{N\to\infty} \sum_{R_{-i}=0}^{N-1} P(R_{-i})\mu(s_i|y_i, R_{-i})$$

$$\lim_{N\to\infty} \sum_{R_{-i}=0}^{N-1} P(R_{-i})\frac{\mu(s_i, y_i|R_{-i})}{p(y_i|R_{-i})}$$

Using, a) and b), the expression above can be rewritten as:

$$\lim_{N\to\infty} \sum_{R_{-i}=0}^{N-1} P(R_{-i})\frac{p_s E[\nu|R_{-i}]}{E[\tau|R_{-i}]}$$

Inserting the identity in c) yeilds the following,

$$\lim_{N\to\infty}\sum_{R_{-i}=0}^{N-1} P(R_{-i})\frac{p_s(E[\tau|R_{-i}]-(1-p_s))}{(2p_s-1)E[\tau|R_{-i}]}$$

$$\frac{p_s}{2p_s-1}-\frac{p_s(1-p_s)}{2p_s-1}\lim_{N\to\infty}\sum_{R_{-i}=0}^{N-1}\frac{P(R_{-i})}{E[\tau|R_{-i}]}$$

$$\frac{p_s}{2p_s-1}-\frac{p_s(1-p_s)}{2p_s-1}\lim_{N\to\infty}\sum_{R_{-i}=0}^{N-1}\int_{1-p_s}^{p_s}\binom{N-1}{R_{-i}}\frac{\tau^{R_{-i}}(1-\tau)^{N-1-R_{-i}}}{E[\tau|R_{-i}]}$$

$$\frac{p_s}{2p_s-1}-\frac{p_s(1-p_s)}{(2p_s-1)^2}\int_{1-p_s}^{p_s}\lim_{N\to\infty}\sum_{R_{-i}=0}^{N-1}\binom{N-1}{R_{-i}}\frac{\tau^{R_{-i}}(1-\tau)^{N-1-R_{-i}}}{E[\tau|R_{-i}]}d\tau$$

The Bayesian estimator, $E[\tau|R_{-i}]$ is consistent and $f(z)=\frac{1}{z}$ is a one to one function; thus, the limit above is $\frac{1}{\tau}$.

$$\frac{p_s}{2p_s-1}-\frac{p_s(1-p_s)}{(2p_s-1)^2}\int_{1-p_s}^{p_s}\frac{1}{\tau}d\tau$$

$$\frac{p_s}{2p_s-1}-\frac{p_s(1-p_s)}{(2p_s-1)^2}\big(log(p_s)-log(1-p_s)\big)$$

Inserting the expression above for $\lim_{N\to\infty}E_{R_{-i}}[\mu(s_i|y_i,R_{-i})]$ into the constraint in Theorem 1.1 gives the desired result.

Proof of Thereom 1.2: The truth-telling constraint for player i is given below:

$$\big|g(\mu(s_i|y_i,R_{-i}))-g(\mu(s_i|n_i,R_{-i}))\big|\le\lambda$$

Using the mean value theorem, the problem can be rewritten as follows:

$$\big|g'(x)(\mu(s_i|y_i,R_{-i})-\mu(s_i|n_i,R_{-i}))\big|\le\lambda \text{ for some } x\in[\mu(s_i|n_i,R_{-i}),\mu(s_i|y_i,R_{-i})]$$

Thus, setting $\xi=\max\limits_{x\in[0,1]} g'(x)$ is a sufficient condition for truth-telling.

Proof of Theorem 1.3: Note the following equalities.

1. $\mu(s_i|y_i, R_{-i}) = \frac{p_s E[\nu|R_{-i}]}{p_s E[\nu|R_{-i}]+(1-p_s)(1-E[\nu|R_{-i}])}$

2. $\mu(s_i|n_i, R_{-i}) = \frac{(1-p_s)E[\nu|R_{-i}]}{p_s(1-E[\nu|R_{-i}])+(1-p_s)E[\nu|R_{-i}]}$

The difference between $\mu(s_i|y_i, R_{-i})$ and $\mu(s_i|n_i, R_{-i})$ is maximized when $E[\nu|R_{-i}] = \frac{1}{2}$. If the interviewer believes $\nu = \frac{1}{2}$ with certainty, the condition $E[\nu|R_{-i}] = \frac{1}{2}$ is satisfied regardless of N or the beliefs of the participants.

CHAPTER 2

Eliciting Private Information Using Correlation: A Modification of Randomized Response

2.1 Introduction

Respondents often lie to interviewers when asked about stigmatizing characteristics or activities such as anti-Semitism, tax evasion, cheating by economics students, and illicit drug use.[1] Researchers want to learn information about the proportion of individuals with such characteristics but cannot ask survey participants directly due to their incentive to respond dishonestly. Respondents often wish to tell the researcher the truth but want to avoid association with the stigmatized group due to embarrassment or, in the case of illegal activity, fear of future litigation.

In order to elicit truthful responses from subjects, researchers have developed several techniques to extract relevant information while still preserving the privacy of the subject. One of the most famous and earliest techniques is the randomized response technique (RRT)[2]. Warner (1965) developed the procedure and analyzed its statistical properties in a seminal paper. In the RRT, subjects privately and independently use a randomization device, such as a die, to determine whether they respond to the statement "I am a member of a group X." or "I am a member of group X^c." where X is a particular stigmatized group and X^c is its complement. For example, if the die lands on one through four, then the participant responds to the first statement and if it lands on five or six, the participant responds to the second statement. The researcher observes the response but does not know the statement received in order to protect the privacy of the subject while still allowing

[1]See Krumpal (2012), Houston and Tran (2001), Kerkvliet (1994) and Fisher, Kupferman, and Lesser (1992) respectively.

[2]There are several variants of the technique such as the forced response technique which is more common in practice, but this paper focuses on randomized response as it is the most statistically related.

the researcher to acquire some information about the parameter of interest. The method is used since many subjects have an incentive to lie under the direct response technique (DRT), causing estimation problems, or subjects refusing to engage in the survey, causing a selection bias problem.[3]

This paper introduces a new method, the correlated response technique (CRT), for surveys when subjects fail to provide truthful responses when asked directly. Under the CRT, the questions asked to respondents are perfectly correlated. The following example illustrates the way to implement the method. Suppose the interviewer wants to know the number of students who cheated in a class where the number of students in the course is twenty. The interviewer brings a deck of cards to the school where fifteen of the cards state "I cheated at least once in the course." and five cards that state "I never cheated in the course." The students privately observe their statement and return a sheet of paper only stating "yes" or "no." The interviewer can estimate the proportion of students cheating while the privacy of the students is still protected since the interviewer fails to know which statement each student received. Unlike in the RRT, the researcher knows the exact number of "I am a member of group X." and "I am a member of group X^c." statements. In the example above, the interviewer knows there are exactly fifteen cards stating the subject cheated and five stating the subject did not cheat.

Knowing the exact number of statements makes the data less noisy than the RRT allowing the researcher to obtain more information about the proportion of stigmatized subjects. As the sample size approaches infinity in the RRT, the researcher would also know the exact number statements; however, adding noise slows the rate of convergence and some studies may not be able to use a large number of subjects due to lack of funding or a small population of interest as in the example. The variance of the estimate obtained under the CRT is less than of the RRT for comparable amounts of randomization.

In addition to using correlation, the procedure in the model also allows for the

[3]Though the RRT clearly protects privacy better than the DRT in theory, whether the RRT actually incentivizes subjects to answer truthfully is a subject of debate. See John, Loewenstein, Acquisti, and Vosgerau for a discussion.

researcher to have private information, which is a novel concept. To understand the process by which incomplete information is added, consider the example again except suppose the interviewer brings two decks of cards instead of one where the first deck contains fifteen questions stating "I cheated at least once in the course." and five questions stating "I never cheated in the course." while the second deck has fifteen questions stating "I never cheated in the course." and five questions stating "I cheated at least once in the course.". The interviewer can distinguish between the two decks while the subjects cannot. The interviewer then flips a coin to determine which deck to use.

By giving the researcher private information about the distribution of questions asked, subjects have a stronger incentive to tell the truth because they are unsure which answer, "yes" or "no", indicates they more likely belong to the stigmatized group.[4] In the example, if the students knew there were fifteen "I cheated at least once in the course." questions, they would likely be more hesitant to give a "yes" answer as the interviewer views a "yes" answer as an indication that the student cheated. With the incomplete information, uncertainty over the distribution of questions creates uncertainty over which answers are more likely to be stigmatizing which encourages subjects to tell the truth as they do not know which answer associates them with the stigmatized group more.

The CRT allows truth-telling to be a dominant strategy. In the example, this amounts to giving ten questions stating "I cheated at least once in the course." and ten questions stating "I never cheated in the course." In the dominant strategy case, there is no uncertainty over the distribution of questions. Since each question is equally likely, neither a "yes" or a "no" response indicates a respondent is more likely belong to a stigmatized group; thus, respondents have a dominant strategy to honestly respond. The method allows the interviewer to determine whether the proportion is extreme or moderate, though it does not allow the researcher to estimate the proportion unless the proportion equals one half. For example, if the

[4]Subjects do, however, know which distribution is more likely given their card which is an aspect that is incorporated in the model.

instructor receives twenty "yes" or twenty "no" responses, then exactly half the students cheated. Of course, if the interviewer receives ten "yes" response and ten "no" responses, any number of the students could have cheated; however, the interviewer updates his belief to reflect the fact that with ten "yes" and ten "no" response the proportion is more likely to be extreme (close to zero or one) than before. Additionally, the technique can be run multiple times with multiple respondents to measure the exact distance of the proportion from one half while still maintaining the dominant strategy property. For example, the researcher could learn the distance of the proportion from one half is 0.4 indicating the proportion is either 0.1 or 0.9.

A researcher may only care whether a parameter is extreme or moderate when trying to measure if a committee is diverse over some dichotomous trait. The dominant strategy technique provides other statistical information such as whether the proportion is above or below a certain threshold.

The next section discusses related literature. Section 2.3 introduces the model. Section 2.4 describes a natural way to implement the model in a practical and effective manner. Section 2.5 provides the posterior distribution and other statistics for a researcher once the data is collected. In Section 2.6, the specific version of the correlated response technique is discussed which allows for truth-telling to be a dominant strategy for any parameter values in my model and even works well in some cases outside of the scope of the model, though in Nash implementation instead of dominant strategy implementation. Section 2.7 analyzes the truth-telling constraint and emphasizes it importance. In Section 2.8, a the variance of a common estimator of the RRT is compared with the analogous estimator of the CRT; additionally, a specific example is used to compare the CRT to the RRT. The factors which decide whether a researcher should use the DRT, the RRT, or the CRT are examined in Section 2.9. Section 2.10 concludes by discussing a few drawbacks to my model as well as future research opportunities related to response techniques.

2.2 Related Literature

Several research papers have introduced new methods similar to the RRT but in all of them, the questions for subjects are determined independently and none of them provide the interviewer with private information. Both Bouza, Herrera, and Pasha (2010) and Chaudhuri and Mukerjee (1988) give a detailed description of many of the proposed procedures for estimating a binary random variable as in the model of this paper. Many of these methods have several randomization devices and though statistical estimators have been developed with these models, the incentive structures have often not been analyzed in detail.

Economists have long known that allowing for correlation can improve information transmission relative to direct communication. Though focused on games in general instead of solely on information transmission, Aumann (1974) extended the concept of Nash Equilibria to allow for correlation which can improve payoffs even in a simple game such as chicken. Goltsman, Hörner, Pavlov, and Squintani (2009) demonstrate that using a strategic mediator can improve payoffs in a standard sender receiver when the bias of the sender is intermediate relative to direct communication. Though theoretically interesting, there are few applications of using correlation, with the exception of direct communication or public signals, in practice. One can view the questions given in the model as the recommendations[5] for players, the answers of the subjects as their actions, and the interviewer as the mediator; therefore, the correlated response technique is one an area to apply Aumann's original concept of correlated equilibria (1974). Correlation can improve communication in survey techniques and can even provide information in cases where the RRT analyzed by Warner cannot as a dominant strategy mechanism exists with CRT.

The response technique most related to the CRT is the group response technique developed by Chaubey and Li (1995). In the group response technique, respondents privately answer a question in a group of size $K > 1$. If one or more respondents

[5] One can view the statement "I cheated in the classroom at least once in the course" as a recommendation for the cheaters to respond "yes" and a recommendation for the others to respond "no".

in the group are of the stigmatized type, the group is considered stigmatized. The researcher can only distinguish between groups with zero stigmatized individuals and groups with more than zero stigmatized individuals. This mechanism requires much more information loss than the one proposed in this paper for a given number of respondents. Therefore, the CRT gives better parameter estimates at a lower cost and can protect privacy just as well or better with correctly chosen parameters.

In addition to introducing a new technique for eliciting information, the paper also discusses the incentive structure under this new technique unlike most other papers which analyze response techniques, as they often only focus on the statistical properties of estimators. The only other three papers that study the incentives from a "utility" standpoint are Flannery (2014), Blume et al. (2013), and Ljungqvist (1993). Ljungqvist (1993) studies the RRT with a general utility model but fails to account for the effect of other respondent's answers. Blume et al. (2013) use a specific utility function, also used by this author, but they only look at the incentives under the RRT with one respondent when the distribution of the parameter of interest is known. Flannery (2014) uses the same model as this paper for the RRT and analyzes equilibria and comparative statics of the procedure with a general number of subjects. Though statisticians have developed many excellent methods of surveys with noise, the incentives for truth-telling under most of these procedures have yet to be studied from a utilitarian perspective. This paper uses a simple linear utility function to study the CRT to make the incentives transparent under the procedure.

2.3 Model

An interviewer wants to know the proportion (or information about the proportion), ν, of a population that belongs to a stigmatized group $\mathcal{S}$. Each respondent i observes a type $\theta_i \in \{s, t\}$ where t is the accepted type and s is the stigmatized type. Let $\mathcal{T}$ denote the set of accepted types in the population. Respondent types are independent conditional on knowing ν. For simplicity and tractability, both the subjects and the experimenter have a uniform prior over ν on the unit interval;

however, subjects update their priors conditional on their types. Thus, those in group $\mathcal{S}$ believe ν is larger than those in $\mathcal{T}$.

In order to elicit information from the respondents, the interviewer gives a question $q_i \in \{q_s, q_t\}$ to respondent i where q_s corresponds to the question "Are you in $\mathcal{S}$?" and q_t corresponds to the question "Are you in $\mathcal{T}$?". Respondents know that the interviewer asks T subjects q_t and the remaining subjects q_s with probability one half and asks T subjects q_s and the remaining subjects q_t with probability one half. Each respondent privately observes a question from the joint randomization device. The interviewer privately observes the distribution of the questions (only the interviewer knows whether T q_t questions were asked or $N - T$ were).

Let r be the vector of responses, R the number of "yes" responses, R_{-i} the number of "yes" responses excluding i's response, and Q the number of q_s questions, ($Q = T$ or $Q = N - T$). Given all the responses, the interviewer estimates ν and forms beliefs of each respondent's type given all responses. The interviewer's belief that respondent i belongs to group s given the data r is written as $\mu(s_i|r)$. Respondents prefer to tell the interviewer the truth but also prefer to be associated with the regular group. Therefore, respondents want the interviewer to believe they belong to group $\mathcal{T}$ and each respondent's utility is decreasing in the expected value of $\mu(s_i|r)$. The payoff to the respondents is similar to that of Blume, Lai, and Lim (2013). Let $\mathcal{H} = \{(s, q_s, y), (t, q_t, y), (s, q_t, n), (t, q_s, n)\}$ be the honesty set and the payoff function for each respondent to be,

$$U_i(\theta, q, r) = \lambda \mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i) - \xi \mu(s_i|R_{-i}, r_i, Q)$$

where $\lambda, \xi \geq 0$ are parameters measuring, respectively, the degrees of lying aversion and stigmatization aversion and $\mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i)$ is an indicator function with the value of one if $(\theta_i, q_i, r_i) \in \mathcal{H}$ and zero otherwise.

For simplicity, the $\lambda \mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i)$ component of the utility function is referred to as the lying cost. The cost may be due to a psychological cost to lying, or the respondent may have the same ultimate goal as the researcher and thus wants the researcher to learn the true value of the parameter ν. The other component of the utility function, $\mu(s_i|r_i, R_{-i})$, is referred to as the belief. Each respondent prefers to be associated with the normal group due to an embarrassment from stigmatization or worries about possible persecution if the stigmatized characteristic is related to illegal activity.

Respondents are assumed to maximize expected utility. Thus, each respondent chooses the response that maximizes the following expected payoff:

$$\lambda \mathbb{I}_{\mathcal{H}}(\theta_i, q_i, r_i) - \xi E_{R_{-i},Q}[\mu(s_i|R_{-i}, r_i, Q)|q_i, \theta_i]$$

The expected sum of "yes" responses depends on the question of the respondent due to the correlation and the type of the respondent since a respondent's type affects beliefs about ν. A strategy for respondent i is a function $\sigma_i : \{\theta_i, q_i\} \mapsto \{y, n\}$. Both stigmatized and normal types update their beliefs given their type where μ_i represents the belief of respondent i. The normal respondents believe others are more likely to be normal as they change their beliefs from $\mu_i(\nu) = 1$ to $\mu_i(\nu|t) = 2(1-\nu)$ while the stigmatized respondents have the following beliefs $\mu_i(\nu|s) = 2\nu$ after updating. Finding the new belief only requires a simple application of Bayes' rule:

$$\mu_i(\nu|t) = \frac{\mu_i(t|\nu)\mu_i(\nu)}{\int_0^1 \mu_i(t|\nu)\mu_i(\nu)d\nu} = \frac{(1-\nu)}{\int_0^1 (1-\nu)d\nu} = 2(1-\nu)$$

Using the same technique, one can verify that the stigmatized respondents have the following beliefs $\mu_i(\nu|s) = 2\nu$ after updating. Additionally, the subjects update their beliefs about the responses of others given their question since there are a set amount of "Are you in $\mathcal{S}$?" questions and "Are you in $\mathcal{T}$?" questions. In the RRT, this is unnecessary as subjects receive their questions independently.

2.4 Implementation of CRT and Discussion

The correlated response technique described above can be implemented in an intuitive, simple, inexpensive way. The interviewer meets all subjects at once in a room and brings two decks of cards. In one deck of cards, the interviewer has T cards with question q_t and in the other deck there are $N - T$ cards with question q_t. The interviewer sets both decks of cards on a table and is aware of the contents of each deck. However, the subjects only know the properties of both decks but lack knowledge on which deck has T q_t questions and which does not. As the subjects may fail to fully trust the interviewer[6], she displays T 'q_s cards from each deck and T q_t cards from each deck in order to convince the subjects the distribution of the deck is as stated.

The interviewer then flips a coin to determine which deck to use. Once the deck of use is determined, the interviewer shuffles the cards and deals one to each subject. Subjects then privately observe a questions and then return a different card possibly with their name and certainly with a "yes" or "no" checked, presumably answering their question truthfully. The card with a "yes" or "no" checked and is returned to the interviewer who then estimates the parameter of interest with the given data. The cards with the question can be returned to the interviewer if they are shuffled first or destroyed by the subjects if they have reservations about the interviewer keeping them.

Giving the interviewer private information may cause one to wonder what prevents the interviewer from trying to create a false sense of randomization such as a weighted coin to use the deck of cards which allows the most information elicitation. Fortunately, neither deck of cards provides more information as they are isomorphic. Since both decks of cards offer the same amount of information, the interviewer has no incentive to deviate from the implementation method discussed.

One might object to this method and the CRT in general by noting the respondents could simply answer the question directly and have their responses shuffled so the responses remain anonymous. Though in some cases, this method is a better technique, it might not always work. If some respondents refuse to take the survey seriously and send an offensive or inappropriate response, the researcher will not be able to identify which subjects wrote such statements. If some subjects refuse to

[6]If the subjects fully trusted the interviewer, the direct response technique is optimal.

respond and return a blank card, again the researcher is unable to identify them. A researcher may be interested in other characteristics such as race, gender, or income, and, therefore, wants the subjects to identify themselves in case one of these characteristics needs to be confirmed. For example, a researcher may want to view a subjects tax returns to verify their income. From a cynical perspective, requiring the respondents to identify themselves helps prevent researchers from falsifying data. Finally, if all respondents admit to being in a stigmatized group under the direct response technique with anonymity, the researcher knows with certainty they all belong to the stigmatized group. Under no circumstances is the researcher ever certain a specific individual belongs to a stigmatized group under the correlated response technique (though the interviewer may know at least a certain number of subject belong to a stigmatized group but lacks knowledge on who they are).

2.5 Beliefs of the Interviewer and Parameter Estimation

The interviewer updates his belief using all the data at once, and the interviewer assumes respondents answer truthfully. Without loss of generality, assume the interviewer asks T q_t questions and $N-T$ q_s questions ($Q = N - T$). Therefore, for notational convenience, the conditioning on Q is suppressed in this section. The distribution of responses given the parameter ν is the following:

$$p(R|\nu) = \sum_{j=max\{R-N+T,0\}}^{min\{R,T\}} \binom{T}{j}\binom{N-T}{R-j}\nu^{T+R-2j}(1-\nu)^{N-T-(R-2j)}$$

To derive the distribution, notice that each player can be divided into two groups, the q_s group and q_t group, and each yes must come from a person in one of these groups. Thus, R is the sum of two binomial random variables, one with parameters $(N-T,\nu)$ and the other with parameters $(T, 1-\nu)$. If everyone responded with a no ($R = 0$), then the distribution above amounts to a beta distribution with parameters $T+1$ and $N-T+1$. However, if everyone responds with no except one person, the interviewer is unsure whether the respondent who answered yes belongs to group q_s or group q_t. If all respond with no except two players, then either both yes respondents belonged to group q_s, both belonged to group q_t, or one belonged to each group. Thus, the binomial distribution is used as the process above describes different combinations of respondents belonging to the two different groups. The *max* and *min* function on the summation are used as no more than T players from group q_t could have answered yes and no more than $N - T$ players from group q_s could have answered yes. The first power corresponds to the number of people who are in group $\mathcal{S}$, **not** those receiving q_s, while the second corresponds to the number of people in group $\mathcal{T}$, again **not** those receiving q_t. Naturally the sum of the powers is N as each respondent must belong to one group and cannot belong to both. For the sake of clarity, and example is given below.

Example 2.1. *Suppose N=5, T=2, and the interviewer asks three respondents "Are you in $\mathcal{S}$?" and two "Are you in $\mathcal{T}$?" Then, the distribution of R given ν can be written as follows:*

$$p(R|\nu) = \begin{cases} \nu^2(1-\nu)^3 & \textit{if } R = 0 \\ 3\nu^3(1-\nu)^2 + 2\nu(1-\nu)^4 & \textit{if } R = 1 \\ 3\nu^4(1-\nu) + 6\nu^2(1-\nu)^3 + (1-\nu)^5 & \textit{if } R = 2 \\ \nu^5 + 6\nu^3(1-\nu)^2 + 3\nu(1-\nu)^4 & \textit{if } R = 3 \\ 2\nu^4(1-\nu) + 3\nu^2(1-\nu)^3 & \textit{if } R = 4 \\ \nu^3(1-\nu)^2 & \textit{if } R = 5 \end{cases}$$

Using Bayes' rule, the posterior of ν given the responses is proportional to $p(R|\nu)$ and the prior.

$$\mu(\nu|R) \propto p(R|\nu)\mu(\nu)$$

Given that the prior is uniform, the distribution is simply proportional to the distribution of $p(R|\nu)$; thus, one only needs to find the constant of integration in order to calculate $\mu(\nu|R)$. The probability of each R given ν is the sum of a weighted beta distributions of ν. Thus, the constant of integration is simply the sum of a set of beta functions evaluated at specific numbers. The formula is given below.

$$\mu(\nu|R) = \frac{\sum_{j=max\{R-N+T,0\}}^{min\{R,T\}} \binom{T}{j}\binom{N-T}{R-j}\nu^{T+R-2j}(1-\nu)^{N-T-(R-2j)}}{\sum_{j=max\{R-N+T,0\}}^{min\{R,T\}} \binom{T}{j}\binom{N-T}{R-j}B(T+R-2j+1, N-T-R+2j+1)}$$

Here the B represents the beta function where the first argument in the parenthesis is the alpha parameter and the second is the beta parameter. Unlike the Warner randomized response technique when numbers are large, the posterior in our model is always tractable and easy to calculate even for a general Beta function as the prior instead of the uniform.[7] The reason the posterior is more tractable is that it is the weighted sum of beta functions unlike in the Warner method.

Ultimately, the goal of the interviewer is to estimate the parameter ν and possibly analyze the posterior distribution. The next two propositions in this section provide methods for estimating ν.

Proposition 2.1. *The method of moments estimator, labeled ν_{MM}, in the CRT model is equal to $\frac{R-T}{N-2T}$.*

One can derive the v_{MM} in an intuitive manner. Because the responses come from one of two groups, those asked q_s or those asked q_t, one only needs to look at the expected number of yes responses from each group. From those asked q_t, the expected number of yes responses is $(1-\nu)T$ and the expected number of yes responses from the other group is $\nu(N-T)$; therefore, the total expected number of yes responses is the sum of the number of yes responses of those asked q_s and those asked q_t combined. Rearranging the equation in terms of ν gives the estimator above. In the case $T = \frac{N}{2}$, then the method of moments estimator fails to exist. Importantly, the method of moments estimator is an unbiased estimator of ν.

[7]See Winkler and Franklin, 1979 for details on deriving the posterior under the Warner method.

Theorem 2.1. *ν_{MM} is an unbiased estimator of ν.*

Proof: This proof and all proofs not in the body of this chapter are in Section 2.11.

Another technique for estimating the parameter is to use the mean of the posterior distribution as it seems to be a more reasonable estimator since values of ν_{MM} are extreme when R is below T or above $N-T$.[8]

Proposition 2.2.

$$E[\nu|R] = \frac{\sum_{j=max\{R-N+T,0\}}^{min\{R,T\}} \binom{T}{j}\binom{N-T}{R-j} B(T+R-2j+2, N-T-R+2j+1)}{\sum_{j=max\{R-N+T,0\}}^{min\{R,T\}} \binom{T}{j}\binom{N-T}{R-j} B(T+R-2j+1, N-T-R+2j+1)}$$

Though the above calculation appears tedious, its formula is much simpler than the mean of the posterior with the Warner randomized response technique which requires the use of a double summation.[9] Though $E[\nu|R]$ always gives interior values for estimating ν, ν_{MM} is easier to calculate and more intuitive. When $Q=T$, simply reverse the roles of T and $N-T$ in the expressions above for estimation.

2.6 Dominant Strategy CRT Mechanism

Organizations often value employing a diverse group of individuals, and similarly, society often wants diversity in institutions. Though characteristics such as race and gender are observable, others such as political affiliation and spirituality may be unknown. A researcher may want to know if a university committee discussing affirmative action or free speech has political or religious diversity. Unfortunately, in some schools having liberal or conservative viewpoints could be considered stigma-

[8] The estimator can easily be amended so it always takes values between zero and one; however, even if it is amended, it has corner solutions when R is below T or above $N-T$.

[9] See Hussain, Shabbir, and Riaz, 2010 for details.

tizing; thus, many individuals may lie or refuse to engage in the survey[10]. However, if the researcher uses the method discussed in this paper and sets $T = \frac{N}{2}$ (assuming N is even)[11], then respondents always have an incentive to tell the truth.

Proposition 2.3. *If $T = \frac{N}{2}$ and N is even, then truth-telling is a dominant strategy regardless of the value of ξ and λ.*

The proof of Proposition 2.3 is rather intuitive. Regardless of the response of a subject, the belief of the interviewer about that subject is unaffected as neither question is more likely. In fact, for any vector of responses, the interviewer's belief about each respondent is one half. Therefore, respondents cannot possibly benefit from lying as it fails to affect the interviewer's beliefs about their types, and since lying cost are positive ($\lambda > 0$), lying is a dominated strategy.

Once the researcher obtains the responses, the researcher can use the posterior distribution from the previous section to perform hypothesis tests. The example and graph below illustrates the posterior distributions for the case when $N = 4$. Clearly, the posterior for R is the same as that for $N - R$ under the dominant strategy approach.

Example 2.2. *Let N=4 and T=2. The following are the conditional probabilities of R:*

$$P(R = 0|\nu) = P(R = 4|\nu) = \nu^2(1-\nu)^2$$

$$P(R = 3|\nu) = P(R = 1|\nu) = 2\nu^3(1-\nu) + 2\nu(1-\nu)^3$$

$$P(R = 2|\nu) = \nu^4 + 4\nu^2(1-\nu)^2 + (1-\nu)^4$$

The posterior distributions given R are graphed in Figure 2.1.

[10] Though not in the framework of this model, the method could also be used if describing a characteristic is taboo.

[11] For simplicity, the case of even N is discussed. When N is odd, a new mechanism must be used that slightly differs from that discussed in the model where the researcher asks $\frac{N-1}{2}$ individuals q_s and $\frac{N-1}{2}$ individuals q_t. The final individual privately flips a coin to determine which question to answer.

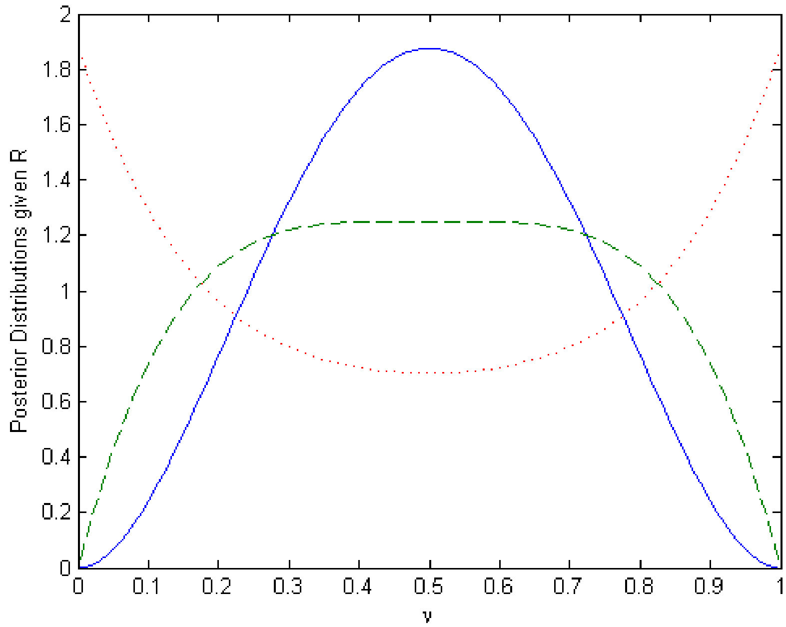

Figure 2.1: Plots of $\mu(\nu|R=0)=\mu(\nu|R=4)$ (blue solid curve), $\mu(\nu|R=1)=\mu(\nu|R=3)$ (green dashed curve),$\mu(\nu|R=2)$ (red dotted curve). The variable ν is on the x-axis while the posteriors are on the y-axis.

In general, if a researcher wants know whether the parameter of interest, ν, is "moderate" (close to half) or "extreme" (far from half) and wants to take a cautious approach since lying cost could be low or stigmatization costs could be high, then, the dominant strategy implementation approach is ideal. Whenever the interviewer has a symmetric prior, responding truthfully is a dominant strategy; however, with other priors responding honestly may not be a dominant strategy. For example, suppose the interviewer believes ν is close to zero and $N=2$, then a respondent belonging to $\mathcal{S}$ may hesitate to respond truthfully as the other respondent likely belongs to $\mathcal{T}$. A truthful response would change the interviewer's belief about the respondent from close to a zero percent chance of belonging to the stigmatized group to fifty percent of belonging to the stigmatized group.

When there is a large number of subjects, truth-telling is a Nash equilibrium even with priors with a lot of density near zero and one. Suppose the prior over

ν puts all mass "close" to zero and one respondent belongs to the stigmatized group and received question q_s; thus, this respondent responds with "yes". The $\frac{N}{2}$ respondents who received the question q_t will respond with "yes" while the other $\frac{N}{2} - 1$ respondents who received q_s will respond "no". The interviewer will believe that one of the respondents who responded "yes" belongs to the stigmatized group but is uncertain which one. His belief about each of the respondents who stated "yes" will change from near zero to $\frac{1}{\frac{N}{2}+1}$. Therefore, in the limit the stigmatized respondent will have an incentive to respond truthfully.

Unfortunately, the dominant strategy approach fails to give a single estimate or interval estimate; thus, the dominant strategy approach has strong limitations. Additionally, the probability of receiving responses close to $\frac{N}{2}$ is always high often giving the test little statistical power. In Example 2 above, if there are two stigmatized subjects and two normal subjects, the probability of $R = 2$ is four times as likely as $R = 0$ or $R = 4$; thus, the likelihood of $R = 0$ or $R = 4$ is rather low. With larger N, the problem exacerbates. In order to overcome this statistical problem, the experiment can be run multiple times.

Theorem 2.2. *If the dominant strategy mechanism is run m times, subjects have a dominant strategy to tell the truth every time.*

Again, since either a "yes" or "no" response fails to affect beliefs about an individual respondent, subjects have no incentive to lie; however, they always have an incentive to tell the truth. Let W_i^m be a vector of responses for respondent i and W^m represent the vector of all responses from all subjects if the experiment has been run m times and S be the true number of stigmatized subjects in the experiment.

Proposition 2.4. *As $m \to \infty$, the experimenter believes*

$$\mu(S|W^m) \xrightarrow{d} \begin{cases} \frac{1}{2} & \textit{if } S = A \\ \frac{1}{2} & \textit{if } S = B \end{cases}$$

where $A + B = N$.

Proposition 2.4 simply states that if the experimenter runs the dominant strategy mechanism multiple times, she knows $|\frac{N}{2} - S|$. Of course she fails to know the sign of the argument, and, thus, only infers whether S is "extreme" or "moderate." For clarity, refer to Example 2. If $R = 1$ or $R = 3$, the interviewer immediately knows that $|2 - S| = 1$, so $S = 1$ or $S = 3$. When the mechanism is only run once and $R = 2$, then the experimenter knows $S \in \{0, 2, 4\}$; however, if the experiment is run enough times and $S = 2$, then $R = 0$ or $R = 4$ eventually occurs, which is not the case when $S \in \{0, 4\}$. Thus, after running the experiment enough times, the experimenter learns whether $S \in \{0, 4\}$ indicating ν is more likely to be close to zero or one, $S \in \{1, 3\}$ indicating ν is more likely close to one quarter or three quarters, or $S = 2$ indicating ν is more likely close to half. If enough subjects are used, then the interviewer knows the distance of the true parameter, ν from half.

Corollary 2.1. *As $m \to \infty$ and $N \to \infty$, then*

$$\mu(\nu|W^m) \xrightarrow{d} \begin{cases} \frac{1}{2} & \textit{if } \nu = a \\ \frac{1}{2} & \textit{if } \nu = b \end{cases}$$

where $a + b = 1$

Corollary 2.1 is an obvious extension of Proposition 2.4. By performing the dominant strategy mechanism on a large number of subjects with an infinite number of repetitions, the interviewer should know the exact distance of the parameter from one half. For example, the experimenter may learn the distance from one half is 0.1 indicating that ν is either 0.4 or 0.6 with each being equally likely. An obvious extension of this model is to consider beliefs different from the uniform. If the prior puts probability zero on the interval from zero to half or half to one, this mechanism would give an exact estimation of ν.[12] Any continuous prior with full support converges to a discrete posterior distribution with support a and b where $a + b = 1$ since the researcher learns the exact distance of ν from one half.

[12]Though implementation would not be in dominant strategy and, even with a large number of respondents truth-telling may not be a Nash equilibrium with so many iterations.

The implementation of the dominant strategy technique requires no private information on the part of the interviewer (coin flipping in the description in Section 2.4) unlike the correlated response technique for other values of T since there is no uncertainty over the distribution of questions. If subjects have lexicographic preferences over lying and stigmatization aversion or lying cost are zero, this is the only CRT mechanism that elicits any information from subjects.

2.7 Truth-telling Equilibrium in CRT Mechanism

Clearly, if respondents always answered truthfully, the optimal CRT mechanism would simply set $T = 0$ and all subjects would answer the same question. Thus, the size of $T \in \{1, ..., K\}$ where $K = \frac{N}{2}$ if N is even and $K = \frac{N-1}{2}$ if N is odd can be regarded as the degree of privacy protection. A researcher wants to set T as low as possible in order to elicit the maximum amount of information while at the same time incentivizing the subjects to truthfully respond.

Theorem 2.3. *Truth-telling is an equilibrium for respondents of both types with either question if*

$$\left| 2E_{R_{-i},Q}[\mu(s_i|R_{-i}, y_i, Q)|q_s] - 1 \right| \leq \frac{\lambda}{\xi}$$

The condition above is similar to that of the randomized response technique studied by Flannery (2014). Unlike the randomized response model, subjects must take into account that the number of "Are you in $\mathcal{S}$?" questions could either be T or $N - T$; additionally, subjects take into account that each of them receive one of the N questions as questions are not drawn independently. Shockingly, the constraint is the same for all four types of respondents which is a product of the linearity of the model and the beta prior.

If the constraint in Theorem 2.3 fails to hold, the methods in this paper cannot be used to analyze the data as all estimations techniques discussed in this paper rely on truth-telling as a condition. Methods have been suggested to overcome cases where subjects lie. Statisticians such as Mangat (1994) and Greenberg et al. (1969) try to take lying into account in their analysis of their own randomized response procedures; however, both authors assume the probability a subject lies, or alternatively the proportion of subjects lying, is known to the researcher. In cases where empirical or experimental evidence exists to support the likelihood someone lies, then a survey method which takes this into account is valid. Unfortunately, this author believes such an occurrence is rare although a recent paper by Clark and Desharnais (1998) addresses a method to accomplish such an estimation. Their estimation, however, requires the assumption that the number of subjects who lie does not change much with the randomization of the questions which is not always likely the case. Alternatively, one could assume subjects use mixed strategies and analyze the data under this assumption, but Blume et al. (2013) take this approach and show that in a randomized response model with only one subject a plethora of equilibria exists making it unlikely the researcher can analyze the data in this case due to a lack of knowledge on which equilibria is being played by the subjects. Though often unstressed by statisticians, the importance of ensuring the constraint holds is vital in almost any survey procedure. Even though this paper focuses on the truth-telling equilibria, these equilbria may not necessary elicit the most information in terms of entropy, the lowest variance unbiased estimator, or some other measure.

Unlike most previously studied survey methods, the method in this paper provides the interviewer with private information to relax the incentives of the subjects. To understand why this changes the incentives of the respondents, one must consider a method similar to the one in this paper except there is no randomization over the distribution of questions. In the implementation method in Section 2.4, this would be analogous to using one deck of cards with a known number of "Are you in $\mathcal{S}$?" and "Are you in $\mathcal{T}$?" questions by the subjects, and the interviewer would no longer flip a coin since there is no randomization over two decks of cards. Consider the case where $N = 3$ the interviewer asks one subject "Are you in $\mathcal{T}$?" questions and two subjects "Are you in $\mathcal{S}$?" questions with no randomization over the decks of cards. An example of such a case is provided below.

Example 2.3. *Suppose the interviewer asks two subjects "Are you in $\mathcal{S}$?" and one subject "Are you in $\mathcal{T}$?" Below are the constraints for each type:*

$\{(s, q_s); (s, q_t); (t, q_s); (t, q_t)\}$.

1. *The constraint for* (s, q_s) *and* (t, q_s) *requires the following holds:* $\frac{19}{45} \leq \frac{\lambda}{\xi}$

2. *The constraint for* (s, q_t) *and* (t, q_t) *requires the following holds:* $\frac{11}{45} \leq \frac{\lambda}{\xi}$

The details of the example and all remaining examples are in Section 2.11.

Subjects who receive the question which is asked more frequently have a stronger incentive to lie. In Example 2.3, the "Are you in $\mathcal{S}$?" question is asked more than "Are you in $\mathcal{T}$?" so the constraint is tighter when q_i is equal to q_s. In Example 2.3 when a subject receives the "Are you in $\mathcal{T}$?" question, then the other two subjects receive an "Are you in $\mathcal{S}$?" questions and it is likely the other respondents will respond with the same answer since types are positively correlated without knowing ν; however, when a subject receives an "Are you in $\mathcal{S}$?" question, the other two subjects received different questions and likely give different answers due to the positive correlation again. The change in the beliefs of the interviewer from three "yes" or "no" responses to two "yes" or "no" responses is not as strong as the change in beliefs from two "yes" responses to two "no" responses; thus, the subjects with the "Are you in s?" questions are more likely to have a stronger effect on the belief so the benefit to lying is greater and, therefore, the constraint is tighter. Since the estimation techniques in Section 2.5 require truth-telling, the following constraint must hold: $\frac{\lambda}{\xi} \geq \frac{19}{45}$.

When the interviewer has private information over the distribution of questions, the incentives are the same for all groups of subjects since neither question is more likely than the other. The example below illustrates how the constraints change when $N = 3$, $T = 1$, and the implementation method used from Section 2.4 is used instead of that from Example 2.3.

Example 2.4. *When the interviewer uses the CRT technique,* $N = 3$*, and* $T = 1$*, the following constraint must hold to ensure truth-telling:* $\frac{1}{5} \leq \frac{\lambda}{\xi}$

The constraint in Example 2.4 is a weighted difference of the two constraints in Example 2.3. Respondents cannot tell for sure which response is more jeopardizing as they fail to know whether $Q = 1$ or $Q = 2$; nevertheless, they know the likelihood of receiving the more jeopardizing questions is twice as much as receiving the less jeopardizing question. Consequently, their constraint is two thirds multiplied by the more jeopardizing constraint minus one third times the less jeopardizing constraint of Example 2.3. Clearly, the private information greatly reduces the constraint as it moves from $\frac{19}{45}$ to $\frac{1}{5}$.

If N is even, a truthful equilibria always exists by using the dominant strategy mechanism described in the previous section and the method can easily be amended (see footnote in previous section) to always elicit truth-telling in the case of an odd number of respondents.

2.8 Comparing CRT to RRT

In the randomized response model, subjects independently use a private randomization device to determine whether to answer q_s or q_t. Let p_s be the probability a subject is asked “Are you in $\mathcal{S}$?”. Warner (1965) solved the maximum likelihood estimator, denoted $\hat{\nu}$ in this paper. The formula for $\hat{\nu}$ is given below:

$$\hat{\nu} = \frac{p_s - 1}{2p_s - 1} + \frac{R}{(2p_s - 1)N}$$

When $p_s = \frac{N-T}{N}$, the estimator is the same as the ν_{MM} in Proposition 2.2. Both estimators are an unbiased estimate of ν and comparing them is a natural way to compare the RRT and CRT.

Theorem 2.4. *If $p_s = \frac{N-T}{N}$ and $T \neq \frac{N}{2}$, then $Var(\nu_{MM}) < Var(\hat{\nu})$.*

As the CRT has a lower variance for its estimator, it may be better to use it than the RRT as long as the incentives remain the same. Clearly, when $p_s = \frac{1}{2}$, the RRT is completely uninformative while the CRT provides information on whether ν is moderate or extreme when $T = \frac{N}{2}$. Thus, the CRT gives better estimates of ν

than the RRT when the probability of the question q_s is the same.

The reason the CRT provides a better estimator is quite intuitive. With the RRT, the underlying distribution of questions given to the respondents is a binomial distribution with parameters p_s and N. Therefore, the interviewer is uncertain how many of each question, q_s or q_t, was asked. Under the CRT, the interviewer knows exactly how many q_s and q_t questions were asked. By reducing the noise associated with the questions, the interviewer can estimate ν more accurately.

In order to compare the incentives and posterior of the CRT and RRT, the case of $N = 3$ is used with $T = 1$ and the probability a subject is asked "Are you in $\mathcal{S}$?" in the RRT is two thirds. Two thirds is chosen as it one of the most natural comparisons with $T = 1$ along with one third. The reason $N = 1$ or $N = 2$ is not analyzed in this section is that the natural probability for comparison with $N = 1$ is zero or one which is the direct response technique while the natural probability for comparison for $N = 2$ is one half in which case the response in the RRT is completely uninformative (though not in the CRT).

Since subjects independently use a private randomization device to determine whether to answer the question "Are you in $\mathcal{S}$?" or "Are you in $\mathcal{T}$?" under the RRT, conditioning on the question is unnecessary since it provides subjects no information about the responses of others unlike in the CRT.

Example 2.5. *Under the RRT, when the probability of being asked "Are you in $\mathcal{S}$?" is two thirds and $N = 3$, then truth-telling is an equilibrium if $\frac{5{,}153}{15{,}795} \leq \frac{\lambda}{\xi}$.*

The constraint of Example 5 is more binding that that of Example 4. Without the private information on the part of the interviewer, however, this is not the case as the constraint in Example 3 is more binding than Example 5.

The graphs of the posteriors of each distribution are given above when $R = 2$ and $R = 3$. The posterior for $R = 0$ is simply a rotation of $R = 3$ around $x = \frac{1}{2}$ and ditto for $R = 1$ and $R = 2$. In the graphs above, it appears the change in beliefs under the CRT is larger than the change in beliefs under the RRT indicating the CRT produces more information than the RRT.

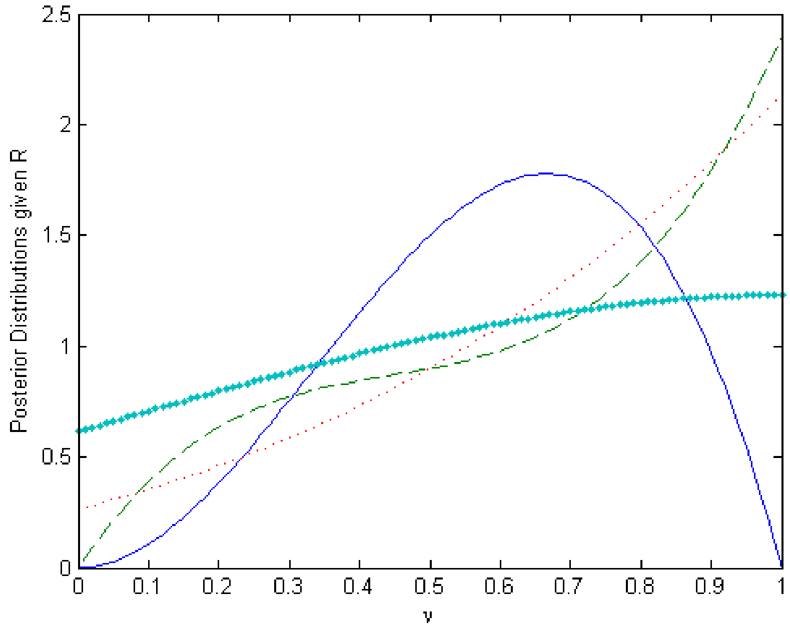

Figure 2.2: Plots of the posteriors from Example 4 and Example 5 of RRT when R=3 and $p_s = \frac{2}{3}$ (red dotted curve) and R=2 and $p_s = \frac{2}{3}$ (light blue studded curve) and CRT when R=3 and Q=2 (dark blue solid curve) and when R=2 and Q=2 (green dashed curve). The x-axis of the graph is ν while the y-axis is $\mu(\nu|\cdot)$.

2.9 Choosing a Mechanism

When studying a binary characteristic in society through the use of a survey, the choice of the mechanism is paramount as the researcher wants to accurately measure the parameter of interest. In order to induce truth-telling on the part of the subjects, randomization may be needed; however, more randomization creates more noise which then makes the researcher require more subjects for estimation. Since a researcher likely has limited resources, minimizing the noise in a mechanism and choosing the correct mechanism is important.

When lying costs are high relative to stigmatization aversion, using the DRT is clearly optimal as it provides more information when subjects truthfully reveal their types. Unsurprisingly, the DRT is the most used in practice. Additionally, if the subjects trust the researcher to keep their responses private and do not feel embarrassed telling the truth, it is also optimal to use the DRT. One large advantage of the DRT is the ease of which it can be implemented over the phone, internet, or mail unlike the other response techniques.

When lying costs are low relative to stigmatization aversion and the interviewer cannot be trusted or subjects feel embarrassed telling the interviewer the truth, there are a few factors that determine when the CRT is better to use than the RRT. In cases where the researcher wants only to know whether the parameter of interest is moderate or extreme, the CRT is optimal as it provides a dominant strategy mechanism and allows the interviewer to extract much information in this regard since it can be run multiple times. Additionally, as the amount of subjects and runs approaches infinity, the interviewer can calculate the exact distance of the parameter of interest, ν, from one half. When there are few subjects, the CRT may provide more information than the RRT. Finally, the posterior distribution is less computationally complex than that of the RRT as it consists of only weighted sums of beta functions though with modern computing the issue looms less large than before.

Many times, however, the RRT may be better to use than the CRT. In the model of this paper, the payoff function is linear. With a non-linear payoff function, the RRT may be better. Payoffs are likely non-linear for characteristics such as rape, murder, or other criminal behavior as beliefs close to one can be thought of as "beyond a reasonable doubt" in the legal system. For example, if $N = 3$ and $T = 1$, there are some combinations of responses, ($R = 0$ or $R = 3$), which guarantee the researcher knows at least one individual belongs to the stigmatized group. In the RRT, a researcher always remains uncertain about the number of people belonging to the stigmatized group regardless of responses. If an interviewer only wants to insert a small amount of noise, the RRT may be better as the CRT has a lower bound of $T = 1$ while the noise of randomized response can be arbitrarily close to zero; additionally, the noise in the randomized response technique can be changed infinitesimally while the correlated response technique requires changes in discrete increments as $T \in \{1, 2, 3...\}$ giving more freedom of choice with noise with the randomized response technique. Finally, the researcher may want to interview the subjects at different times as all may not be available at one time or the interviewer may want to add more respondents to the survey later which cannot be done with the CRT.

Though the model has a lying aversion parameter, λ, which is independent of the technique used, subjects may feel more comfortable giving a truthful a response in one of the two settings. Whether individuals feel more comfortable sharing their private information in a group setting, such as alcoholics anonymous, or a private setting, such as a confession, is largely an experimental and empirical question that could depend on the characteristic being studied. Subjects might find it easier to understand one of the techniques better than the other and could feel more comfortable with the mechanism that is simpler to comprehend, which again, is largely an experimental and empirical question.

2.10 Conclusion

This paper develops a new survey technique which uses correlation and private information in order to elicit more information from subjects while still incentivizing them to reply truthfully. Methods for estimating the parameter of interest under the CRT are also provided. Under the proposed scheme, a special case exists where truth-telling is a dominant strategy for subjects allowing an interviewer to elicit information, regardless of the lying cost or stigmatization aversion. The paper also emphasizes the importance of incentives under any survey technique as the data may simply be noise when they are ignored.

A natural extension is to look at a model with a more general payoff function, to allow any form of correlation instead of perfect correlation, or both. One benefit to the method discussed in this paper versus a method that allows a very general form of correlation is the natural way which the technique can be implemented, which is likely not the case for a model which allows general correlation. Linear payoffs are used for simplicity, but looking at models with a general payoff function, polynomial payoff function, or even a quadratic payoff function would be an interesting extension, though tractability may be a issue.

An experimental investigation of the technique needs to be undertaken for it to be used in practice as it is important to know whether subjects actually behave as the model predicts. Better yet, a field study should be done to validate whether the method works well or at least better than the randomized response technique. Validation studies are difficult as they require the parameter of interest to be known, but such a study is essential before the CRT is taken seriously.

Several fields including economics use surveys for important studies which sometimes provide information that has important policy implications for governments, businesses, or charities; hence, the manner in which surveys are conducted and the incentives of any survey technique need to be studied. Unfortunately, few papers on survey techniques give a detailed analysis on the incentive structure even though truth-telling constraints usually must hold for the estimation to be valid. Hopefully, this paper inspires more literature on the subject to improve data gathering and its policy implications.

2.11 Proofs and Elaborations

Proof of Theorem 2.1:

$$E[\nu_{MM}|\nu] = E\left[\frac{R-T}{N-2T}\middle|\nu\right] = \frac{E[R|\nu]-T}{N-2T}$$

$$= \frac{(1-\nu)T+\nu(N-T)-T}{N-2T} = \frac{\nu(N-2T)}{N-2T} = \nu$$

Proof of Proposition 2.3: Using independence and the definition of conditional probability, the belief the of experimenter about respondent i given a “yes” response from respondent i can be rewritten as follows:

$$\mu(s_i|R_{-i},y_i)=\int_0^1\frac{\nu P(R_{-i}|q_s,\nu)}{\nu P(R_{-i}|q_s,\nu)+(1-\nu)P(R_{-i}|q_t,\nu)}\mu(\nu|R_{-i}+1)d\nu$$

$P(R_{-i}|q_s,\nu)$ represents the probability of R_{-i} responses from the $N-1$ respondents other than i given respondent i received the question q_s and the parameter ν, and a similar definition applies for $P(R_{-i}|q_t,\nu)$. The denominator above is equal to $2P(R_{-i},y_i|\nu)$ and using the identity $\mu(\nu|R_{-i}+1)=\frac{P(R_{-i},y_i|\nu)\mu(\nu)}{P(R_{-i},y_i)}$. Inserting this identity into the equation above yields:

$$\mu(s_i|R_{-i},y_i)=\int_0^1\frac{\nu P(R_{-i}|q_s,\nu)\mu(\nu)d\nu}{2P(R_{-i},y_i)}$$

$$=\frac{\int_0^1\nu P(R_{-i}|q_s,\nu)\mu(\nu)d\nu}{\int_0^1\nu P(R_{-i}|q_s,\nu)\mu(\nu)d\nu+\int_0^1(1-\nu)P(R_{-i}|q_t,\nu)\mu(\nu)d\nu}$$

Due to the symmetry of the prior, one can use the substitution $u=1-\nu$ for $\int_0^1(1-\nu)P(R_{-i}|q_t,\nu)\mu(\nu)d\nu$ to show it is equal to $\int_0^1\nu P(R_{-i}|q_s,\nu)\mu(\nu)d\nu$, so $\mu(s_i|R_{-i},y_i)$ is equal to one half. One can use the same methodology to show $\mu(s_i|R_{-i},n_i)$ also equals one half. Since a response has no effect on the interviewer's belief, responding honestly is a dominant strategy with any positive lying cost.

Proof of Proposition 2.4: Each $S\in\big\{\{N,0\},\{N-1,1\},\{N-2,2\},..,\{\frac{N}{2}\}\big\}$ has a unique support. The probability $R=0$ or $R=N$ is only positive if $S=\frac{N}{2}$. The probability $R=N-1$ or $R=1$ is only positive when $S\in\{\frac{N+2}{2},\frac{N-2}{2}\}$. When $S\in\{\frac{N+4}{2},\frac{N-4}{2}\}$, the probability that $R=2$ and $R=N-2$ is positive while the probability that $R=0$ or $R=N$ is zero which is unique to $S\in\{\frac{N+2}{2},\frac{N-2}{2}\}$. A similar argument applies to all such $S\in\big\{\{N,0\},\{N-1,1\},\{N-2,2\},..,\{\frac{N}{2}\}\big\}$. With enough runs of the experiment, all events with a positive probability occur allowing the interviewer to learn the support of R. In each set listed the first element plus the second element equals N (with the exception of $S=\frac{N}{2}$ but one can view this set as $\{\frac{N}{2},\frac{N}{2}\}$ to overcome this problem). Since the prior over ν is uniform, the first and second element of each set are equally likely; thus, the interviewer places a probability of half on each in the posterior.

Proof of Theorem 2.3: Truth telling is an equilibrium if the following constraint holds given my model:

$$\left| E_{R_{-i},Q}[\mu(s_i|R_{-i}, y_i, Q) - \mu(s_i|R_{-i}, n_i, Q)|\theta_i, q_i] \right| \leq \frac{\lambda}{\xi}$$

The mathematical statement above simply means the expected benefit to lying (the expected belief differential) should be larger than the lying cost. Assume $q_i = q_s$ and $\theta_i = s$ The expected belief differential can be rewritten as,

$$\sum_{R_{-i}=0}^{N-1} P(T|q_s)P(R_{-i}|s, q_s, T)\mu(s_i|R_{-i}, y_i, T)$$
$$- \sum_{R_{-i}=0}^{N-1} P(T|q_s)P(R_{-i}|s, q_s, T)\mu(s_i|R_{-i}, n_i, T)$$
$$+ \sum_{R_{-i}=0}^{N-1} P(N-T|q_s)P(R_{-i}|s, q_s, N-T)\mu(s_i|R_{-i}, y_i, N-T)$$
$$- \sum_{R_{-i}=0}^{N-1} P(N-T|q_s)P(R_{-i}|s, q_s, N-T)\mu(s_i|R_{-i}, n_i, N-T)$$

To further simplify the problem, the following identities are used.

1. $P(R_{-i}|s, q_s, Q) = P(N-1-R_{-i}|t, q_s, Q)$: To prove this identity, note that

   $$P(R_{-i}|s, q_s, Q) = \int_0^1 \mu_i(\nu|s)P(R_{-i}|q_s, Q)$$

   Simply use the substitution $x = 1 - \nu$ and reintegrate to obtain the identity.

2. $P(T|q_s) = P(q_s|T) = P(N-T|q_t) = P(q_t|N-T)$ and $P(N-T|q_s) = P(q_s|N-T) = P(q_t|T) = P(T|q_t)$. This identity can easily be shown by applying Bayes' rule.

3. $\mu(s_i|R_{-i}, y_i, Q) = \mu(t_i|N-1-R_{-i}, n_i, Q) = 1 - \mu(s_i|N-1-R_{-i}, n_i, Q)$: This

identity comes from the underlying symmetry of the model.

Applying identity one and two and reordering the summation to the expected belief differential yields:

$$\sum_{R_{-i}=0}^{N-1} P(q_s|T)P(R_{-i}|s,q_s,T)\mu(s_i|R_{-i},y_i,T)$$

$$-\sum_{R_{-i}=0}^{N-1} P(q_s|T)P(R_{-i}|t,q_s,T)\mu(t_i|R_{-i},y_i,T)$$

$$+\sum_{R_{-i}=0}^{N-1} P(q_s|N-T)P(R_{-i}|s,q_s,N-T)\mu(s_i|R_{-i},y_i,N-T)$$

$$-\sum_{R_{-i}=0}^{N-1} P(q_s|N-T)P(R_{-i}|t,q_s,N-T)\mu(t_i|R_{-i},y_i,N-T)$$

Using identity two again and three as well to the expected belief differential simplifies the expression to the following:

$$P(q_s|T)\left[\sum_{R_{-i}=0}^{N-1}(P(R_{-i}|s,q_s,T)+P(R_{-i}|t,q_s,T))\mu(s_i|R_{-i},y_i,T)-1\right]+$$

$$P(q_s|N-T)\left[\sum_{R_{-i}=0}^{N-1}(P(R_{-i}|s,q_s,N-T)+P(R_{-i}|t,q_s,T))\mu(s_i|R_{-i},y_i,N-T)-1\right]$$

The expression $P(R_{-i}|s,q_s,T) + P(R_{-i}|t,q_s,T) = \int_0^1 \mu_i(\nu|s)P(R_{-i}|q_s,T,\nu) + \mu_i(\nu|t)P(R_{-i}|q_s,T,\nu)d\nu = \int_0^1 P(R_{-i}|q_s,T)(2\nu+2(1-\nu))d\nu = 2P(R_{-i}|q_s,T)$. Similarly, the expression $P(R_{-i}|s,q_s,N-T)+P(R_{-i}|t,q_s,N-T) = 2P(R_{-i}|q_s,N-T)$. Applying these equalities to the expected belief differential gives:

$$P(q_s|T)\left[\sum_{R_{-i}=0}^{N-1} 2P(R_{-i}|q_s,T)\mu(s_i|R_{-i},y_i,T)-1\right]+$$

$$P(q_s|N-T)\Big[\sum_{R_{-i}=0}^{N-1} 2P(R_{-i}|q_s, N-T)\mu(s_i|R_{-i}, y_i, N-T) - 1\Big]$$

The above expression, using identity two and algebra, can be rewritten as:

$$\sum_{R_{-i}}^{N-1} P(R_{-i}, T|q_s)\mu(s_i|R_{-i}, y_i, T) + P(R_{-i}, N-T|q_s)\mu(s_i|R_{-i}, y_i, T)$$

Inserting the expression above for the expected belief differential into the constraint gives the proposed result. Clearly, the incentives are symmetric for each question type under the incentive scheme so the same holds for $q_i = q_t$, and by the symmetry of the beta function, it is the same for $\theta_i = t$.

Elaboration of Example 2.3: Let $(\theta_1, q_1) = (s, q_s)$. The following constraint must hold for truth-telling to be an equilibrium.

$$\Big|\sum_{r_2, r_3} p(r_2, r_3|s, q_s)[\mu(s_1|y_1, r_2, r_3) - \mu(s_1|n_1, r_2, r_3)]\Big| \leq \frac{\lambda}{\xi}$$

The probabilities when $r_2 = r_3 = y$ are derived below:

$$p(y_2, y_3|s, q_s) = \int_0^1 p(y_2, y_3|\nu, q_s)\mu_1(\nu|s)d\nu = \int_0^1 \nu(1-\nu)2\nu d\nu = \frac{1}{6}$$

Since the first subject received the "Are you in s?" question, he knows the only way the other subjects received opposite questions and only both answer "yes" if they belong to different groups which is why $p(y_2, y_3|\nu, q_s) = \nu(1-\nu)$. To calculate the beliefs of the interviewer, note that

$$\mu(s_1|r_1, r_2, r_3) = \frac{p(r_1, r_2, r_3|s_1)p(s_1)}{p(r_1, r_2, r_3)} = \frac{.5p(r_1, r_2, r_3|s_1)}{p(r_1, r_2, r_3)}$$

All "yes" responses only occur when two of the subjects belong to the stigmatized group and one belongs to the normal group and both subjects in the stigmatized group receive the "Are you in $\mathcal{S}$?" question and the normal subject receives the question "Are you in t?."

$$p(y_1, y_2, y_3) = \int_0^1 p(R = 3|\nu) = \int_0^1 \nu^2(1-\nu)d\nu = \frac{1}{12}$$

To calculate $p(y_1, y_2, y_3|s_1)$, one must note that respondent one only answers yes if $q_1 = q_s$; therefore, the vector of questions must be (q_s, q_s, q_t) or (q_s, q_t, q_s), and each occurs with probability one third. Then, $p(y_1, y_2, y_3|s_1)$ can be written as:

$$p(q_s, q_s, q_t)p(y_1, y_2, y_3|s_1, q_s, q_s, q_t) + p(q_s, q_t, q_s)p(y_1, y_2, y_3|s_1, q_s, q_t, q_s)$$

$$= 2p(q_s, q_s, q_t)p(y_1, y_2, y_3|s_1, q_s, q_s, q_t) = \frac{2}{3}\int_0^1 p(y_1, y_2, y_3|s_1, q_s, q_s, q_t, \nu)$$

$$= \frac{2}{3}\int_0^1 \nu(1-\nu)\mu_1(\nu|s) = \frac{2}{3}\int_0^1 \nu(1-\nu)2\nu = \frac{1}{9}$$

Hence, $\mu(s_1|y_1, y_2, y_3) = \frac{2}{3}$ which should be intuitive since the interviewer knows two subjects are stigmatized. All other probabilities are calculated in the same fashion and inserting them into the constraint gives: $\frac{19}{45} \leq \frac{\lambda}{\xi}$. One can do the same for all other combinations of θ_1 and q_i to solve for the constraints.

Elaboration of Example 2.4: The probabilities for when $Q = 2$ are the same as those from the previous example. For the probabilities when $Q = 1$, note that by the symmetry of the beta function $p(r_1, r_2, r_3|Q = 1) = p(r_1, r_2, r_3|Q = 2)$ and by the isomorphic relationship on $\mu(s|r, N-T)$ and $\mu(s|r, T)$, $\mu(s_1|y_1, X) = \mu(s_1|n_1, N-1-X)$ for all $X \in \{0, 1, 2, ..., N-1\}$. The constraint with the CRT mechanism is the sum of the weighted constraints over Q. To calculate the conditional probabilities of Q:

$$p(Q = 2|q_s) = \frac{p(q_s|Q = 2)p(Q = 2)}{p(q_s|Q = 2)p(Q = 2) + p(q_s|Q = 1)p(Q = 1)} = \frac{\frac{2}{3} * \frac{1}{2}}{\frac{2}{3} * \frac{1}{2} + \frac{1}{3} * \frac{1}{2}}$$

$$= \frac{2}{3}$$

After solving for all the probabilities in the same fashion, one can insert them into the constraint to solve.

Proof of Theorem 2.4: In order to calculate the variance of each estimator, the decomposition of variance formula is used.

$$E[Var(\nu_{MM}|\nu)] + Var[E(\nu_{MM}|\nu)] < E[Var(\hat{\nu}|\nu)] + Var[E(\hat{\nu}|\nu)]$$

Each estimator is unbiased so $Var[E(\nu_{MM}|\nu)] = Var[E(\hat{\nu}|\nu)] = Var(\nu)$ leaving the following,

$$E[Var(\nu_{MM}|\nu)] < E[Var(\hat{\nu}|\nu)]$$
$$\frac{E[Var(R|\nu)]}{(N-2T)^2} < E[Var(\hat{\nu}|\nu)]$$

The random variable $R|\nu$ is the sum of two independent binomial variables with parameters $((1-\nu), T)$ and $(\nu, N-T)$; therefore, $Var(R|\nu) = \nu(1-\nu)T + \nu(1-\nu)(N-T) = \nu(1-\nu)N$. The $Var(\hat{\nu}|\nu)$ is derived in Warner (1965). Inserting these two identities yields,

$$\frac{NE[\nu(1-\nu)]}{(N-2T)^2} < \frac{E[(p_s\nu + (1-p_s)(1-\nu))(p_s(1-\nu) + (1-p_s)\nu)]}{N(2p_s-1)^2}$$

After simplifying the expectations, the expression above can be rewritten as

$$\frac{N}{6(N-2T)^2} < \frac{1+2p_s-2p_s^2}{6N(2p_s-1)^2}$$

Insert the fact that $p_s = \frac{N-T}{N}$,

$$\frac{N}{6(N-2T)^2} < \frac{1+2\frac{N-T}{N} - 2(\frac{N-T}{N})^2}{6N(2\frac{N-T}{N}-1)^2}$$

Multiply the right hand side by $\frac{N}{N}$ and simplification yields,

$$1 < 1 + 2\frac{N-T}{N} - 2(\frac{N-T}{N})^2$$

The identity above clearly holds which completes the proof.

Elaboration of Example 2.5: Solving for the constraints is different in this example than the previous two since the questions are generated independently instead of being perfectly correlated. The constraint under the RRT for $\theta_1 = s$ is:

$$|\sum_{r_2,r_3} p(r_2, r_3|s)[\mu(s_1|y_1, r_2, r_3) - \mu(s_1|n_1, r_2, r_3)]| \leq \frac{\lambda}{\xi}$$

The probabilities are calculated below in the case when $r_2 = r_3 = y$. Note that

$$p(y_i|\nu) = p(y_i, s_i|\nu) + p(y_i, t_i|\nu) = \frac{2}{3}\nu + \frac{1}{3}(1-\nu) = \frac{1}{3} + \frac{1}{3}\nu$$

$$p(y_2, y_3|s_1) = \int_0^1 p(y_2, y_3|\nu)(\mu(\nu|s_1)d\nu = \int_0^1 (\frac{1}{3} + \frac{1}{3}\nu)^2 2\nu d\nu = \frac{17}{54}$$

Again, the beliefs of the interviwer can be rewritten:

$$\mu(s_1|r_1, r_2, r_3) = \frac{p(r_1, r_2, r_3, s_1)}{p(r_1, r_2, r_3)}$$

Solving for the numerator when $r_1 = r_2 = r_3 = y$ gives

$$p(y_1, y_2, y_3, s_1) = \int_0^1 p(y_1, s_1|\nu)p(y_2|\nu)p(y_3|\nu)d\nu = \int_0^1 \frac{2}{3}\nu(\frac{1}{3} + \frac{1}{3}\nu)^2 d\nu = \frac{17}{162}$$

Solving for the denominator when $r_1 = r_2 = r_3 = y$ gives

$$p(y_1, y_2, y_3) = \int_0^1 p(y|\nu)^3 d\nu = \int_0^1 (\frac{1}{3} + \frac{1}{3}\nu)^3 d\nu = \frac{5}{36}$$

Therefore, $\mu(s_1|y_1, y_2, y_3) = \frac{34}{45}$. All other probabilities are calculated in the same fashion and inserting them into the constraint gives: $\frac{5{,}153}{15{,}795} \leq \frac{\lambda}{\xi}$. One can do the same for $\theta_1 = t$ to find the same constraint.

CHAPTER 3

A Response Technique With Dominant Strategies in Forced Responses

3.1 Introduction

> "You have no power over the dice. When they fall on twelve, then you have to say no. Something is being forced on you. And then it's just like eating, you have to eat spinach but you don't like it. Well, then I won't do it. That thing [the computer] is a dead thing. It obliges me to press a key I don't want to. My answer belongs to me."
>
> –Lensvelt-Mulders and Boeije 2007

The above quote comes from a sociological study on the effectiveness of the forced response technique—a statistical technique used to measure the proportion of a population with a stigmatized characteristic. The quote demonstrates that the subject in the study feels uncomfortable following a command given to them from the technique, and, therefore, chooses to disobey the command. To overcome this problem of disobedience, this paper develops a technique where following a command is always optimal regardless of how other respondents in the survey behave. Obedience in the technique is critical to providing an accurate estimate of the proportion studied.

In the forced response technique, respondents *privately* use a randomization device to determine whether to answer a sensitive question or obey a command, either "Just say yes." or "Just say no.". For example, a respondent could use a die for the randomization device and answer the question "Did you vote in the last presidential election?" when the die lands on one through four, "Just say no" if the die lands on five, and "Just say yes" if a die lands on six.[1] Researchers use

[1] Holbrook and Krosnick (2010) test the validity of a different response technique by using an approach similar to this.

the technique instead of a standard survey where they directly ask participants questions, a technique referred to as the direct response technique, because the technique creates uncertainty about which group a participant belongs to regardless of response. When the researcher observes a “yes” or “no” response, the researcher fails to know whether the respondent answered the question or simply obeyed one of the commands. When respondents answer honestly, the method reduces both estimation bias from lying and selection bias as those in the stigmatized group should be more willing to participate in the survey. If survey participants fail to obey commands in the forced response technique, participants in the stigmatized group may forgo participation or lie when given a question since the uncertainty provided by those obeying the commands no longer exists. These failures could create estimation bias and selection bias, eliminating the purpose of introducing the technique. If participants always have an incentive to obey a command as under the technique in this paper, those in the stigmatized group will be more inclined to answer honestly.

In order to create this dominant strategy incentive, the interviewer has private information on the underlying question. For example, the when the respondent receives a command, he or she is uncertain whether the question is “Did you vote in the last presidential election?” or “Did you abstain from voting in the last presidential election?”. Because of this uncertainty, the respondent cannot determine which command is more associated with the stigmatized group, “yes” or “no”.

I analyze the technique I develop in this paper with a utility model as recommended by Ljungqvist (1993). Using a utility model allows me to demonstrate that obeying a command is a dominant strategy—a choice that is optimal regardless of how the other respondents behave. Other models of analyzing forced response techniques, known as privacy measures, usually fail to distinguish the incentives with this technique versus one where obeying a command is not a dominant strategy, and, consequently not necessarily optimal.

The next section covers much of the literature on the forced response technique and its effectiveness. Section 3.3 describes the model of the paper. In Section 3.4, I discuss a natural way to implement the technique. In Section 3.5, I demonstrate following a command is a dominant strategy and give the incentive constraint required to ensure that answering a question honestly is also optimal. Section 3.6 provides some drawbacks of the technique developed in this paper. Section 3.7 discusses a more general way to introduce incomplete information to response techniques; however, it also suggests restrictions on the ways to implement the technique. In Section 3.8, I provide a discussion on why many of the privacy measures commonly used fail to apply to the method developed in this paper. Lastly, Section 3.9 concludes.

3.2 Literature Review

Warner (1965) developed the first model that introduced randomization in the survey process to protect the privacy of the individuals surveyed. Several variants of the technique have been developed in the decades since including the forced response technique.[1] Researchers have used these techniques to measure doping (Striegel, Ulrich, and Simon, 2010), induced abortions in urban North Carolina (Abernathy, Greenberg, and Horvitz, 1970), cigarette smoking among adolescents (Lauer, Akers, Massey, and Clarke, 1982), angler noncompliance with fishing regulations (Schill and Kline, 1995), and dear poaching (Wright, 1980). Though response techniques, such as the forced response technique, theoretically reduce selection and estimation bias, academics debate whether the technique actually works well in practice.

In order to test the validity of the technique, researchers use the "more is better criterion" or direct validation studies. The more is better criterion simply states that a higher incidence of reporting using a response technique versus direct questioning indicates that response techniques are better.[2] Though direct validation studies are considered a better method to test the technique, they are more difficult to perform as the technique inquires about personal and sensitive information. Lensvelt-Mulders, Hox, Van der Heijden, and Maas (2005) perform a meta-analysis on thirty-two previous studies, six of which are validation studies, and find some

[1] See Chaudhuri (2010) chapters 3,5, and 8.
[2] See Umesh and Peterson (1991) for a critique of this evaluation procedure.

evidence that indicates response techniques improve estimation; however, there is strong variation across studies.

Edgell, Himmelfarb, and Duchan (1982) perform a study to test the validity of the forced response technique by asking fifty-four college students several personally invasive questions using the forced response technique. Unfortunately, they used a computer to perform the randomization, which sometimes failed to actually randomize by their design. However, they find that a large amount of respondents failed to comply with the forced response when it associated them with the stigmatized group (Edgell et al., 1982). More recent validation studies employing the forced response technique also suggest the method fails to provide accurate estimates of the parameter of interest.

Wolter and Preisendörfer (2013) interview a group of known convicts and individually asked whether each one was convicted with both the forced response technique and direct response technique[3], and both techniques severely underestimated the proportion (each gave an estimate close to 0.6 instead of 1). Unfortunately, the study contained no "innocents" to provide cover for the convicts with forced "yes" responses, which the authors admitted was a weakness of their study (Wolter and Preisendörfer, 2013).

Van der Heijden, van Gils, Bouts, and Hox (1998) use the forced response technique when interviewing individuals known to engage in social security fraud. They find that the forced response technique performs better than traditional surveys but still severely underestimates the proportion as they receive an estimate of 0.43 when the true value is 1.0. Again, the study contained no "innocents" just like the study by Wolter and Preisendörfer (2013). Because these studies contained only individuals from the stigmatized group, one must be cautious to interpret them as demonstrating the forced response technique does not work.

Lensvelt-Mulders and Boeije (2007) interviewed eighteen welfare recipients in The Netherlands using the forced response technique and inquired about the participant's opinions regarding the technique afterwards. Two of ten respondents

[3]The interviewees did not know the researchers knew of their conviction.

who used the technique admitted to breaking the rules: one rolled the dice multiple times—two dice were used for the randomization device–and the other simply refused to follow instructions as the quote in the beginning indicated (Lensvelt-Mulders and Boeije, 2007). These recent studies indicate a hesitation to follow commands when a researcher employs the forced response technique. The technique designed in this paper eliminates this problem and could replace the standard forced response technique for this reason.

3.3 Model

An interviewer want to know the proportion, ν, of a population, which belongs to a stigmatized group, $\mathcal{S}$. To estimate this proportion, the interviewer employs a forced response technique where the probability of the commands “Just say yes.”, denoted c_y, and “Just say no.”, denoted c_n, are each equal to $p \in [0, \frac{1}{2}]$. The probability of a question is $1 - 2p$, where the question is either “Are you in $\mathcal{S}$?”, denoted q_s, or “Are you in $\mathcal{T}$?”, denoted q_t, where $\mathcal{T} = \mathcal{S}^c$ and is referred to as the accepted group. The interviewer randomly chooses which question to use with equal probability and knows which question is used before employing the technique.

The interviewer recruits N respondents to take the survey with the forced response technique. Respondent i has a type, θ_i, which indicates the respondent belongs to $\mathcal{S}$, in which case, $\theta_i = s$, or $\mathcal{T}$, in which case $\theta_i = t$. The probabilities that $\theta_i = s$ and $\theta_i = t$ are ν and $1 - \nu$ respectively. Types are drawn independently conditional on the parameter ν. Respondent i also receives a private and independent random draw, $x_i \in \{c_y, c_n, q_s, q_t\}$. The respondent knows that the probability of each command, c_y and c_n, is equal to p and the probability of a question, *either* q_s or q_t, but not both, is $1 - 2p$. The respondent fails to know whether the question is q_s or q_t, but the respondent knows each question is equally likely. Therefore, from the perspective of the respondent, the probability of each question is $\frac{1}{2} - p$ before the respondent knows x_i. Whether the support of x_i contains q_s or q_t is independent among the participants. The random variable $q_i \in \{q_s, q_t\}$ is the question in the support of i's draw. The capital letter Q represents a vector of all questions, $(q_1, q_2, ..., q_N)$.

Each respondent replies to the survey with $r_i \in \{yes, no\}$ and r represents the vector of all responses, $(r_1, r_2, ..., r_N)$ and r_{-i} be the response vector of all respondents but i. A respondent is obedient if he or she follows a command and disobedient when failing to follow a command. A respondent is honest if he or she responds to a question about his or her type truthfully. Finally, a respondent is dishonest if he or she lies in response to a question.

Given the all the responses, the interviewer forms a belief about respondent i belonging to the stigmatized group, $\mu(s_i|r, Q)$. The interviewer and respondents each have a prior on ν over the unit interval denoted $\mu(\nu)$ for the interviewer and $\mu_i(\nu)$ for respondent i. Each respondent updates his or her beliefs based upon type; thus, if respondents had a uniform prior, the posterior for the stigmatized type would be:

$$\mu_i(\nu|s) = \frac{p(s|\nu)\mu_i(\nu)}{\int_0^1 p(s|\nu)\mu_i(\nu)d\nu} = \frac{\nu}{\int_0^1 \nu d\nu} = 2\nu$$

Similarly, $\mu_i(\nu|t) = 2(1 - \nu)$.

Throughout the paper, the notation $p(r_i = yes)$, $p(x_i = c_y)$, $p(\theta_i = s)$, etc. is written as $p(yes)$, $p(c_y)$, $p(s_i)$ etc.. When the experimenter uses the question q_s, the probability of a "yes" response given the parameter ν under the assumption of obedience and honesty of the participants is derived below:

$$p(yes|\nu, q_s) = p(c_y) + p(q_s)p(s_i|\nu) = p + (1 - 2p)\nu$$

An honest and obedient respondent only answers "yes" when receiving the question q_s when he or she belongs to group $\mathcal{S}$ or obeying the command c_y. Of course, one can similarly show $p(no|\nu, q_t)$ is the same. Since each respondent receives the draw, x_i, and type, θ_i, independently, conditional on ν, the sum of the "yes" responses with question $q_i = q_s$ plus the sum of "no" respondents with question $q_i = q_t$, represented by R, is binomially distributed with parameters $p + (1 - 2p)\nu$ and N. Similarly, the sum of all other such "yes" and "no" responses excluding r_i, labeled R_{-i}, is binomially distributed with parameters $p + (1 - 2p)\nu$ and $N - 1$. The belief of the interviewer about respondent i depends on R and r_i instead of the entire vector, r, as shown below:

$$\mu(s_i|r, Q) = \int_0^1 \mu(s_i, \nu|r, Q)d\nu = \int_0^1 \mu(s_i|r_i, \nu, q_i)\mu(\nu|R)d\nu$$

Above, $\mu(\nu|R)$ is the researcher's posterior distribution given the data. The interviewer's belief about participant i and the parameter ν can be separated due to the underlying independence assumption.

The utility of each respondent depends on the respondent's honesty or obedience and the interviewer's belief about the respondent $\mu(s_i|r, Q)$. Let $U_i(\mu_i(s_i|r, Q), r_i|x_i, \theta_i)$ represent each player's utility function. The utility of each respondent is decreasing in the interviewer's belief about i, $\mu(s_i|r, Q)$. For simplicity, y and n are written to represent "yes" and "no" respectively. For any given fixed belief k, respondents prefer obedience:

$$U_i(k, y|c_y, \theta_i) \geq U_i(k, n|c_y, \theta_i) \forall k \in [0, 1]$$

$$U_i(k, n|c_n, \theta_i) \geq U_i(k, y|c_n, \theta_i) \forall k \in [0, 1]$$

For any given fixed belief k, respondents prefer honesty:

$$U_i(k, y|q_s, s) \geq U_i(k, n|q_s, s) \forall k \in [0, 1]$$

$$U_i(k, n|q_s, t) \geq U_i(k, y|q_s, t) \forall k \in [0, 1]$$

$$U_i(k, y|q_t, t) \geq U_i(k, n|q_t, t) \forall k \in [0, 1]$$

$$U_i(k, n|q_t, s) \geq U_i(k, y|q_t, s) \forall k \in [0, 1]$$

Each respondent maximizes their expected utility over the random variable Q and the expected responses of all other participants, r_{-i}, and the expected utility is denoted $E_Q\big[E_{r_{-i}}\big[U_i(\mu(s_i|r_i, r_{-i}, Q), r_i|x_i, \theta_i)\big]\big]$. To simplify notation,

$u_i(\mu(s_i|r_i, r_{-i}), r_i|x_i, \theta_i)$ represents the expected utility of player i (over the random variables Q) from choosing r_i given the draw x_i and type θ_i. A strategy for respondent i is a function that maps the respondent's type and draw, (θ_i, x_i), to a response, r_i.

Instead of questions and commands, the interviewer could use cards which specify a command or question that could be the same or different for each type, s or t. With this method, the interviewer would only need to either use commands or questions, not both. For example, instead of asking "Are you in $\mathcal{S}$?", a card could state "Just say yes if you belong to $\mathcal{S}$ and Just say no if you belong to $\mathcal{T}$." However, using both commands and questions is the most common in both the theoretical literature and field studies. In particular, many studies employ the forced response technique which often consists of only using q_s and c_y.

The payoffs are different for commands than questions even though it can be modeled with either all commands or all questions. Psychologically, a person may feel worse for lying than disobeying or visa versa. For this reason, respondents receive different payoffs for lying and disobeying; however, the model also allows for equal payoffs to lying and disobeying. The goal of the interviewer is to estimate ν. The interviewer faces a tradeoff when choosing p as a large p increases the privacy protection of the respondents, which incentivizes them to respond truthfully; however, a large p also decreases the amount of information transmission with each respondent whether that information is measured in terms of the variance of an estimator, the expected entropy of the posterior distribution, or some other measure. In the extreme case, choosing $p = \frac{1}{2}$ fails to transmit any useful information since participants only receive commands; conversely, choosing $p = 0$ eliminates all privacy protection because this is equivalent to the direct response technique.

3.4 Implementation

A researcher can implement the variant of the forced response technique discussed in this paper by simply using a deck of cards. For demonstration purposes, assume that $p = \frac{1}{3}$; therefore, the researcher would bring four cards to the interview but only use three cards on each subject, two of which are the different commands and the third of which is a question. Before the interview begins, the researcher flips a coin to determine whether to use the question q_s or q_t. The researcher has an incentive to only "stack" the deck with questions instead of commands; thus, the researcher must demonstrate that two of the cards are the different commands. Once the respondent is fully aware that two of the cards are the different commands, the researcher shuffles all three cards after which the respondent privately draws one. The respondent then observes the card and writes down a simple "yes" or "no" answer on a piece of paper. The respondent hands the researcher the paper and then returns the card and shuffles it or simply takes all the cards at the end of the trial. Either way the researcher is always uncertain over which card the respondent drew.

It is important to note that the researcher does not have a preference for the question q_s or q_t as they elicit the same amount of information. This occurs because the researcher simply reverses the roles of "yes" and "no" when the question is q_t as opposed to q_t in the estimation process. Therefore, the participants have no reason to believe that the researcher would prefer one question over the other; thus, the prior of the participants over the distribution of the third card, the question, is fifty-fifty over q_s and q_t. The implementation method described here can easily be modified with any p that is a rational number.

3.5 Incentives for Obedience and Honesty

Most variants of the forced response technique have one command that is more associated with the stigmatized group. For example, suppose the respondent knows the question beneath one of the cards, and, furthermore, assume the question q_s. In this case, respondents may be hesitant to obey the "yes" command since a "yes" response indicates the respondent is more likely to belong to the stigmatized group than a "no" response. With the technique described in this paper, obeying the command is a dominant strategy. A dominant strategy is a strategy that is always better than any other regardless of what the other players (in this case the other respondents and the researcher) choose to do. Thus, when $x_i \in \{c_n, c_y\}$, it is optimal for respondent i to obey the command regardless of whether $\theta_i = s$ or $\theta_i = t$.

Proposition 3.1. *Suppose respondent i receives a command, $x_i \in \{c_y, c_n\}$. Then, obeying the command is optimal regardless of the responses of others, a respondent's type, or choice of p by the interviewer:*
$u_i(\mu(s_i|y, r_{-i}), y|c_y, \theta_i) \geq u_i(\mu(s_i|n, r_{-i}), n|c_y, \theta_i)$ *and*
$u_i(\mu(s_i|n, r_{-i}), n|c_n, \theta_i) \geq u_i(\mu(s_i|y, r_{-i}), y|c_n, \theta_i)$, $\forall r_{-i}$, $\forall p$, *and* $\forall \theta_i$.

See Section 3.10 for proof.

In order for the researcher's estimate to be correct, the researcher must ensure that participants both obey the commands and respond honestly to the question. If participants fail to obey the command, this may eliminate all privacy protection and cause the participants who receive the question to lie when the otherwise would not have.

Proposition 3.2. *For any possible vector of questions, Q, there exists a Nash Equilibirum where each participant responds honestly if the following two inequalities hold for each i:*

$$E_{R_{-i}}\big[U_i\big(\int_0^1 \frac{(1-p)\nu}{p+(1-2p)\nu}\mu(\nu|R_{-i}+1)d\nu, y|q_s, s\big)\big] \geq$$

$$E_{R_{-i}}\big[U_i\big(\int_0^1 \frac{p\nu}{p+(1-2p)(1-\nu)}\mu(\nu|R_{-i})d\nu, n|q_s, s\big)\big]$$

and

$$E_{R_{-i}}\big[U_i\big(\int_0^1 \frac{(1-p)\nu}{p+(1-2p)\nu}\mu(\nu|R_{-i}+1)d\nu, n|q_t, s\big)\big] \geq$$

$$E_{R_{-i}}\big[U_i\big(\int_0^1 \frac{p\nu}{p+(1-2p)(1-\nu)}\mu(\nu|R_{-i})d\nu, y|q_t, s\big)\big]$$

See Section 3.10 for proof.

Only the constraints for the s types matter as answering the questions as an honest response is always optimal for the t type. This occurs since the an honest response by a member of $\mathcal{T}$ always increases the interviewer's belief that the participant's type is t. Thus, an experimental design only needs to take into account the s types utility and the beliefs of the s types to ensure truth-telling. The first constraint ensures that the s types respondent answer "yes" when given the question "Are you in $\mathcal{S}$?" while the second ensures the s types answer "no" when given the question "Are you in $\mathcal{T}$?".

According to Proposition 3.2, the payoff to each participant only depends on the prior of the researcher and R_{-i}, the sum of "yes" responses to q_s plus the sum of "no" responses to q_t. This random variable only affects the payoffs indirectly through the researcher's posterior belief of ν.[4] Additionally, the constraints above depend on the beliefs of respondent i implicitly as beliefs skewed to the right more (i believes ν is close to one) increase the likelihood R_{-i} is larger. In the two extreme cases, $N = 1$ and $N = \infty$, only one prior on ν affects the incentives to respond truthfully. When $N = 1$ only the prior of the researcher affects the incentive to respond honestly. The case of the uniform prior is given in Corollary 3.1.

Corollary 3.1. *When N=1 and the researcher has a uniform prior ($\mu(\nu) = 1$), truth-telling is an equilibrium if:*

$$U_i\big(1-p, y|q_s, s\big) \geq U_i\big(p, n|q_s, s\big)$$

$$U_i\big(1-p, n|q_t, s\big) \geq U_i\big(p, y|q_t, s\big)$$

When $N = \infty$, only the prior of the respondent affects the incentive to respond honestly as long as the researcher's prior has full support. This occurs since the researcher learns ν with an infinite sample if respondents are honest and obedient so the prior is irrelevant. However, the participant remains uncertain of ν at the time of the response, and since he is an expected utility maximizer, takes the expectation over his or her prior.

[4]See Flannery (2014) for a more detailed discussion.

Corollary 3.2. *When $N = \infty$, truth-telling is a Nash Equilibrium if the following inequalities hold:*

$$\int_0^1 U_i\left(\frac{(1-p)\nu}{p+(1-2p)\nu}, y|q_s, s\right) - U_i\left(\frac{p\nu}{p+(1-2p)(1-\nu)}, n|q_s, s\right)\mu_i(\nu|s)d\nu \geq 0$$

$$\int_0^1 U_i\left(\frac{(1-p)\nu}{p+(1-2p)\nu}, n|q_t, s\right) - U_i\left(\frac{p\nu}{p+(1-2p)(1-\nu)}, y|q_t, s\right)\mu_i(\nu|s)d\nu \geq 0$$

Before applying Corollary 3.2, one must take caution as the rate of convergence slows as more noise is added. Thus, even with a large number of subjects, using Corollary 3.2 to approximate incentives may be inadequate.

3.6 Drawbacks to Technique

As with any randomized response technique, the method described in this paper burdens a researcher with higher costs than the direct response technique. Requiring subjects to use a randomization device is inherently more costly as it takes more time to implement and more time for the respondent to comprehend. However, if respondents lie or avoid participation in a survey with the direct response technique, then it may be necessary to use a randomized response technique.

Additionally, the implementation process described here is more complicated that a standard forced response technique where a respondent only uses a die for example. Explaining the distributions of each deck to the participants may be confusing for them, and they may worry that the interviewer is somehow trying to fool them into revealing his or her private information without any privacy protection. Thus, interviewers must be careful to be transparent and patient when employing the technique to avoid these problems.

Finally, the technique described above allows less freedom in design choice than the standard force response technique since the probability of a yes command, c_y, must be the same as the probability of a no command, c_n. Researchers have often employed the technique with only the question q_s and the command c_y. By using privacy measures, several authors have demonstrated that only using one command

and one question, q_s and c_t, is statistically the most efficient in the sense of producing estimators with the least variance.[5] However, if respondents fail to obey the command under this particular forced response model, then the estimate found using the technique provides little to no meaningful information.

If respondents follow the commands with this technique as the theory predicts but disobey the commands with the standard forced response techniques, then the benefit to using the technique could easily outweigh these costs.

3.7 General Mechanism with Interviewer Private Information

The method described in the previous three sections is just a specific case of a randomized response technique that gives the interviewer private information over the distribution of questions asked. There are many more ways in which an interviewer could have private information and this section generalizes the method. Of course, a researcher may want to place restrictions on the different types of private information to ensure there is no incentive to try to "trick" the participant. For example, if a researcher uses two decks of cards, but one deck has two cards that contain q_s and the other deck has one of each type of question, q_s and q_t, then the interviewer prefers the first deck over the second as it provides more information. Even if a researcher intends to be honest, a participant may be hesitant to obey or respond honestly if there is an opportunity for the researcher to obtain better information by "stacking" a deck of cards for example.

To estimate the proportion, ν, with this more general technique, the interviewer uses M decks of cards containing both questions and commands. For deck $m \in \{1, 2, ..., M\}$, the probability of the commands "Just say yes.", c_y, and "Just say no.", c_n, are labeled p_y^m and p_n^m respectively. Let p_s^m and p_t^m represent the probability of the questions, "Are you in $\mathcal{S}$?", q_s, or "Are you in $\mathcal{T}$?", q_t respectively. The probability the interviewer uses deck m is denoted $p(m)$.

Respondent i also receives a private and independent random draw from deck

[5] See Loynes (1976) for example.

m, $x_i \in \{c_y, c_n, q_s, q_t\}$. Respondents observe their draw and know the distributions of decks, but they fail to know which deck they drew from. The interviewer knows deck the respondent uses; however, the interviewer never observes which card the respondent draws. The interviewer's draw of deck for player i, denoted by m_i, is drawn independently for each respondent. Let $\mathcal{M}$ represent the vector of deck draws for each respondent, $\mathcal{M} = (m_1, m_2, ..., m_M)$, and $\mathcal{M}_{-i}$ represent the vector of deck draws for all respondents other than i, $\mathcal{M}_{-i} = (m_1, m_2, .., m_i, m_{i+1}, .., m_M)$.

The payoffs and belief structure are the same as that of Section 3.3. The definition below provides one of many possible information restrictions to place on each deck of cards, m.

Definition 3.1. *Two decks of cards, m and $m^{'}$, have information equivalence if, when N=1 and the respondent is obedient and honest, there is a bijective function f: $\{yes, no\} \mapsto \{yes, no\}$ such that the interviewer's posterior belief is the same under a response r with m as $f(r)$ with deck $m^{'}$, $\mu(\nu|r, m) = \mu(\nu|f(r), m^{'})$, for any given prior of the interviewer, $\mu(\nu)$. Additionally, the probability of r with deck m is the same as the probability of $f(r)$ under deck $m^{'}$, $p(r|m) = p(f(r)|m^{'})$ for any given prior of the interviewer, $\mu(\nu)$.*

In the technique described in this paper, $M = 2$, $p_y^1 = p_n^1 = p_y^2 = p_n^2 = p$, $p_s^1 = p_t^2 = 1 - 2p$, and $p(1) = p(2) = \frac{1}{2}$. Using the function f where $f(yes) = no$ and $f(no) = yes$, one can show that it satisfies the strong information equivalence condition above. There are other sets of parameter configurations which satisfy information equivalence. Let $M = 2$, $p_s^1 = p_t^1 = p_s^2 = p_t^2 = z$, $p_y^1 = p_n^2 = 1 - 2z$, and the function f has the following property: $f(yes) = no$ and $f(no) = yes$ for any $z \in [0, 1]$. This method would create a stronger incentive to respond truthfully, but it would also have a weaker incentive to obey. To implement this technique, one could simply use an analogous method to that of Section 3.4.

One could also modify Warner's technique to allow for incomplete information using two decks, $M = 2$ and by setting $p_s^1 = p_t^2 = q$ and $p_t^1 = p_s^2 = 1 - q$ for $q \in [0, 1]$ and use the bijection $f(yes) = no$ and $f(no) = yes$ to ensure information equivalence. The probability of using each deck is equally likely. For example, if $q = \frac{1}{3}$, a researcher could use two decks of cards to interview respondents where one has two q_t and one q_s question while the other has two q_s and one q_t question. The interviewer would use the same process described in Section 3.4 to implement the procedure. The incentive to respond truthfully would be stronger under this technique than Warner's since the respondent is unsure which question is more associated with the stigmatized type. Consider the case of $N = 1$ with the uniform prior, $\mu(\nu) = 1$. The following constraint must hold to ensure truth-telling when $x_1 = q_s$ and $\theta_1 = s$:

$$\frac{2}{3}U\left(\frac{2}{3}, yes|q_s, s\right) + \frac{1}{3}U\left(\frac{1}{3}, yes|q_s, s\right) \geq \frac{2}{3}U\left(\frac{1}{3}, no|q_s, s\right) + \frac{1}{3}U\left(\frac{2}{3}, no|q_s, s\right)$$

Under the standard Warner technique with two q_s questions and one q_t question, the incentives to respond truthfully hold if:

$$U\left(\frac{2}{3}, yes|q_s, s\right) \geq U\left(\frac{1}{3}, no|q_s, s\right)$$

Under the standard Warner technique with two q_t questions and one q_s question, the incentives to respond truthfully hold if:

$$U\left(\frac{1}{3}, yes|q_s, s\right) \geq U\left(\frac{2}{3}, no|q_s, s\right)$$

Adding the incomplete information convexifies the truth-telling constraints but weights the one with two q_s questions more because the participant updates the probability of each deck conditional on the draw, x_i, which in this case is q_s. With the standard Warner technique, the respondents always have an incentive to respond to one of the questions honestly but not the other. For example, with two q_s questions and one q_t question, the s-type always has an incentive to respond honestly to q_t and the t-type to q_s, while the researcher must ensure the s-type responds honestly to q_s and the t-type to q_t. Convexifying these constraints makes the incentives for each question the same for both types. Since truth-telling depends on the most binding constraints, this method increases the likelihood that respondents will report truthfully.

The definition of information equivalence in this paper is particularly strong. For example, an alternative measure could only require the same amount of entropy for each posterior distribution or an equal variance of the Bayesian mean instead of requiring the posterior distributions to be the same. Alternatively, taking a frequentist approach, one could hold the variance of an estimator, such as the MLE or moments estimator, constant instead. All of these measures seem reasonable as well, but the definition of information equivalence above is a sufficient condition for all the others described to hold and a natural place to begin. One could also use a similar definition, but use the uniform distribution for the prior instead of requiring the equalities to hold for every prior.

Definition 3.2. *Two decks of cards, m and $m^{'}$, have uniform equivalence if, when N=1 and the respondent is obedient and honest, there is a bijective function f: $\{yes, no\} \mapsto \{yes, no\}$ such that the interviewer's posterior belief is the same under a response r with m as $f(r)$ with deck $m^{'}$, $\mu(\nu|r, m) = \mu(\nu|f(r), m^{'})$, under the uniform prior, $\mu(\nu) = 1$. Additionally, the probability of r with deck m is the same as the probability of $f(r)$ under deck $m^{'}$, $p(r|m) = p(f(r)|m^{'})$ when the interviewer has a uniform prior over the unit interval, $\mu(\nu) = 1$.*

Clearly, if information equivalence holds, then uniform equivalence also holds. Uniform equivalence may be more useful if all respondents are interviewed simultaneously. There is one obvious rule that fits the second definition which fails to fit the first definition. Let $M = 4$, $p_s^1 = p_y^1 = p_s^2 = p_n^2 = p_t^3 = p_y^3 = p_t^4 = p_n^4 = \frac{1}{2}$, and $p(1) = p(2) = p(3) = p(4) = \frac{1}{4}$. Decks one and four and decks two and three would use the identity function to prove uniform equivalence while decks one and two, decks one and three, decks two and four, and decks three and four would use the mapping $f(yes) = no$ and $f(no) = yes$.

To see why this specification of probabilities and decks has uniform equivalence but not information equivalence, consider the first two decks: $p_s^1 = p_y^1 = p_s^2 = p_n^2$ with the prior $\mu(\nu) = 2\nu$. This prior has an expected value of two-thirds and puts more density on values of ν closer to one. The probability of a yes response for the first deck and no from the second deck are given below:

$$p(yes|m=1) = p(c_y|m=1) + p(q_s|m=1)p(s_i|m=1) = \int_0^1 \big(\frac{1}{2} + \frac{1}{2}\nu\big)2\nu d\nu = \frac{5}{6}$$

$$p(no|m=2) = \int_0^1 \big(\frac{1}{2} + \frac{1}{2}(1-\nu)\big)2\nu d\nu = \frac{2}{3}$$

The probability of a "yes" response with question q_s is higher than the probability of a "no" response with question q_t since the interviewer believes the stigmatized type is more likely. Because of this, information equivalence fails. Just as with the technique described in earlier sections, this technique with four decks also has a dominant strategy in forced responses for similar reasons. Unfortunately, this technique can lead to a researcher believing that respondent belongs to a stigmatized group when the respondent responds truthfully. If the respondent draws the question from the second deck and answers "yes" or the fourth deck and answers "no", then the interviewer knows the respondent belongs to the stigmatized group. If the respondent is sufficiently harmed from being perceived as belonging to the stigmatized group with probability one, then this technique probably works poorly. For example, if litigation is possible from belonging to the stigmatized group, then the respondent may always want to leave some level of doubt in the interviewer's mind. In these cases, $U_i(1, r_i|x_i, \theta_i)$ may be so low that respondents never have an incentive to truthfully.

Obviously, adding more decks makes running the experiment more complex; and, thus, more costly and confusing. Therefore, the technique with four decks described in the last paragraph may be a poor choice when conducting a survey. The technique described in detail in this paper is fairly simple and ensures obedience is optimal. However, each of these techniques should be analyzed using validation studies to determine whether they improve upon the current response techniques.

3.8 Problem with Privacy Measures

The techniques of this paper highlight the importance of looking at incentives instead of using "privacy measures." Ljungqvist (1993) introduced this concept over two decades ago; however, most statisticians continue to use privacy measures instead. When respondents are confronted with a command in the technique discussed in this paper, they always have an incentive to obey it due to the interviewer's private information about the question in the deck. With a few exceptions, it is unclear how privacy measures can incorporate the interviewer's private information. By using utility to model incentives, adding private information can easily be added by using a von Neumann-Morgenstern utility function. To understand why the privacy measures often fail, consider the method in this paper with $p = \frac{1}{3}$ as discussed in Section 3.3. Below are a list of different privacy measures (using the notation of this paper):

1. $max\{\mu(s_i|yes,\nu), \mu(s_i|no,\nu)\}$ (Lanke, 1976)

2. $Prob(y|s_i,\nu)\mu(s_i|y,\nu) + Prob(n|s_i,\nu)\mu(s_i|n,\nu)$ (termed hazard, Greenberg, Kuebler, Abernathy, and Horvitz (1977))

3. According to Nayak (1994), "a design d_1 is better than another design d_2" if $\mu(s_i|y,\nu;d_1) \leq \mu(s_i|y,\nu;d_2)$, $\mu(s_i|n,\nu;d_1) \leq \mu(s_i|n,\nu;d_2)$, and $var(\hat{\nu};d_1) \leq var(\hat{\nu};d_2)$ where $\hat{\nu}$ is an estimate of ν such as the MLE.

All of these privacy measures would treat the method in Section 3.3 the same as that of a technique where the question is known by the respondent, whether it be q_s or q_t.[6] In other words, even though this mechanism gives a stronger incentive to obey the command compared to one where the respondent knows the question in the deck, these privacy measures treat them the same. In particular, if $q_i = q_s$, the privacy measure of Lanke (1976) would be $\mu(s_i|y,\nu) = \frac{p_s\nu}{p_s\nu+(1-p_s)(1-\nu)}$ regardless

[6] There are a few privacy measure which can be amended to account for incomplete information such as the privacy measure proposed by Zaizai and Zankan (2004), $|\mu(s_i|y,\nu) - \mu(s_i|n,\nu)|$. However, these are rare and not extensively used in the literature.

of whether the respondent knows $q_i = q_s$ or has uncertainly as in the procedure discussed in Section 3.3.

Of course, each of these privacy measures serve a purpose. For example, *hazard* by Greenberg et al. (1977) demonstrated that protecting the privacy of those in group $\mathcal{S}$ comes at a cost to those in group $\mathcal{T}$ as they may be perceived as belonging to the stigmatized group when responding honestly, even though they would not be in a standard survey (Direct Response Technique). Additionally, Nayak (1994) himself states that privacy protection and incentives (referred to as honest participation by him) are two distinct issues.

Ultimately, most privacy measures focus on the posterior beliefs about the researcher for different values of ν. As Theorem 3.2 shows, the decision of respondent i to report honestly also depends on his or her prior beliefs about the parameter ν, the responses of others, and the prior of the interviewer. Obeying a command clearly affects the beliefs of the interviewer, but the respondents are uncertain about how they affect the interviewer's beliefs, which cannot be modeled using the privacy measures listed above.

Additionally, as discussed by Flannery (2014), privacy measures fail to take into account the estimation process of ν by the interviewer. Through the interviewer's belief updating process, respondents impose an externality on each other as a "yes" response when $q_i = q_s$ increases the interviewer's belief about ν, which in turn increases the interviewer's belief about other respondents. Because of these issues associated with "privacy measures", any response technique should also be analyzed using a utility model like the one of this paper.

3.9 Conclusion

This paper develops a technique to overcome the problem of disobedience with the forced response technique. Though the technique has some drawbacks mostly due to the complexity of its implementation, it may be necessary to implement since prior research demonstrates that respondents often fail to obey the commands that associate them with the stigmatized group. Although providing the interviewer with private information can provide a robust incentive for participants to obey commands, one can develop other forms of response techniques with private information such as those in Section 3.7. The paper provides two different reasonable criterion to restrict the set of techniques available to ensure that the interviewer has an incentive to perform the procedure honestly. Other sets of restrictions should be further discussed and developed once the initial technique discussed in this paper is common in practice.

The incentives under this technique cannot be adequately studied using the privacy measures that are still pervasive in the statistics literature. Hopefully, this paper solves the problems related to the standard forced response technique in practice as well as in theory in order to improve the estimation of stigmatized populations.

3.10 Proofs

Proof of Proposition 3.1: Let Q^i denote a vector of questions where each element of the vector is the same as that of the vector Q except the ith element. For example if $Q = (q_s, q_t, q_s)$, then $Q^2 = (q_s, q_s, q_s)$. The probability of the vector Q^i is equal to the probability of the vector Q since each question, q_s or q_t, is equally likely to be in the support of x_i for all participants and due to the independence of questions among subjects. The belief of the interviewer about subject i given $r_i = y$ with vector Q is the same as the belief given $r_i = n$ with vector Q^i, $\mu(s_i|y, r_{-i}, Q) = \mu(s_i|n, r_{-i}, Q^i)$. Since the probability of Q and Q^i are the same and $\mu(s_i|y, r_{-i}, Q) = \mu(s_i|n, r_{-i}, Q^i)$, choosing to obey the command is optimal since disobeying a command fails to change the expected beliefs of the researcher.

Proof of Proposition 3.2: As stated in Section 3.3, the belief of the interviewer can be written in the following way:

$$\mu(s_i|r, Q) = \int_0^1 \mu(s_i, \nu|r, Q)d\nu = \int_0^1 \mu(s_i|r_i, \nu, q_i)\mu(\nu|R)d\nu$$

The belief $\mu(s_i|r_i, \nu, q_i)$ can be rewritten as below using the definition of a conditional distribution:

$$\mu(s_i|r_i, \nu, q_i) = \frac{p(s_i, r_i|q_i, \nu)}{p(r_i|q_i, \nu)} = \frac{p(s_i, r_i|q_i, \nu)}{p(s_i, r_i|q_i, \nu) + p(t_i, r_i|q_i, \nu)}$$

When $q_i = q_s$ and $r_i = yes$, the equation above is,

$$\frac{(1-p)\nu}{(1-p)\nu + p(1-\nu)} = \frac{(1-p)\nu}{\nu(1-2p) + p}$$

This value is in the left-hand side of the first constraint in Proposition 3.2. One can verify the other values in the constraint by inserting the other values of q and r_i into the same equation. The experimenter changes his or her posterior belief depending on the response of the respondent. Hence, if r_i is "yes", then the experimenter has the posterior belief $\mu(\nu|R_{-i}+1)$ and if r_i is "no", the experimenter has the posterior belief $\mu(\nu|R_{-i})$. Inserting both the derived posterior belief about the respondent given ν and about the posterior belief about the parameter ν gives the constraints in Proposition 3.2.

REFERENCES

Abernathy, J. R., B. G. Greenberg, and D. G. Horvitz (1970). Estimates of induced abortion in urban North Carolina. *Demography*, **7**(1), pp. 19–29.

Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics*, **1**(1), pp. 67–96.

Blume, A., O. J. Board, and K. Kawamura (2007). Noisy talk. *Theoretical Economics*, **2**(4), pp. 395–440.

Blume, A., E. K. Lai, and W. Lim (2013). Eliciting Private Information with Noise: The Case of Randomized Response.

Bouza, C. N., C. Herrera, and G. Pasha (2010). A review of randomized response procedures: the qualitative variable case. *Rev InvOper*, **31**(3), pp. 240–247.

Chaubey, Y. P. and W. Li (1995). Comparison between Maximum Likelihood and Bayes Methods for Estimation of Binominal Probability with Sample Compositing. *JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM-*, **11**, pp. 379–392.

Chaudhuri, A. (2010). *Randomized response and indirect questioning techniques in surveys*. CRC Press.

Chaudhuri, A. and R. Mukerjee (1988). *Randomized response: Theory and techniques*. Marcel Dekker New York.

Clark, S. J. and R. A. Desharnais (1998). Honest answers to embarrassing questions: Detecting cheating in the randomized response model. *Psychological Methods*, **3**(2), p. 160.

Crawford, V. P. and J. Sobel (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pp. 1431–1451.

Edgell, S. E., S. Himmelfarb, and K. L. Duchan (1982). Validity of forced responses in a randomized response model. *Sociological Methods & Research*, **11**(1), pp. 89–100.

Fisher, M., L. B. Kupferman, and M. Lesser (1992). Substance use in a school-based clinic population: Use of the randomized response technique to estimate prevalence. *Journal of Adolescent Health*, **13**(4), pp. 281–285.

Flannery, T. (2014). A Game Theoretic Analysis of the Randomized Response Technique.

Fligner, M. A., G. E. Policello, and J. Singh (1977). A comparison of two randomized response survey methods with consideration for the level of respondent protection. *Communications in Statistics-Theory and Methods*, **6**(15), pp. 1511–1524.

Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani (2009). Mediation, arbitration and negotiation. *Journal of Economic Theory*, **144**(4), pp. 1397–1420.

Goodstadt, M. S. and V. Gruson (1975). The randomized response technique: A test on drug use. *Journal of the American Statistical Association*, **70**(352), pp. 814–818.

Greenberg, B. G., A.-L. A. Abul-Ela, W. R. Simmons, and D. G. Horvitz (1969). The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association*, **64**(326), pp. 520–539.

Greenberg, B. G., R. R. Kuebler, J. R. Abernathy, and D. G. Horvitz (1977). Respondent hazards in the unrelated question randomized response model. *Journal of Statistical Planning and Inference*, **1**(1), pp. 53–60.

Holbrook, A. L. and J. A. Krosnick (2010). Measuring Voter Turnout By Using The Randomized Response Technique Evidence Calling Into Question The Methods Validity. *Public Opinion Quarterly*, **74**(2), pp. 328–343.

Houston, J. and A. Tran (2001). A survey of tax evasion using the randomized response technique. *Advances in taxation*, **13**, pp. 69–94.

Hout, A. v. d., U. Böckenholt, and P. G. Van Der Heijden (2010). Estimating the prevalence of sensitive behaviour and cheating with a dual design for direct questioning and randomized response. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **59**(4), pp. 723–736.

Hussain, Z., J. Shabbir, and M. Riaz (2010). Bayesian Estimation Using Warner's Randomized Response Model through Simple and Mixture Prior Distributions. *Communications in StatisticsSimulation and Computation®*, **40**(1), pp. 147–164.

John, L. K., G. Loewenstein, A. Acquisti, and J. Vosgerau (????). Paradoxical Effects of Randomized Response Techniques.

Karlan, D. S. and J. Zinman (2012). List randomization for sensitive behavior: An application for measuring use of loan proceeds. *Journal of Development Economics*, **98**(1), pp. 71–75.

Kerkvliet, J. (1994). Cheating by economics students: A comparison of survey results. *The Journal of Economic Education*, **25**(2), pp. 121–133.

Krumpal, I. (2012). Estimating the prevalence of xenophobia and anti-Semitism in Germany: A comparison of randomized response and direct questioning. *Social science research*, **41**(6), pp. 1387–1403.

Lanke, J. (1976). On the degree of protection in randomized interviews. *International Statistical Review/Revue Internationale de Statistique*, pp. 197–203.

Lauer, R. M., R. L. Akers, J. Massey, and W. R. Clarke (1982). Evaluation of cigarette smoking among adolescents: the Muscatine study. *Preventive Medicine*, **11**(4), pp. 417–428.

Lensvelt-Mulders, G. J. and H. R. Boeije (2007). Evaluating compliance with a computer assisted randomized response technique: a qualitative study into the origins of lying and cheating. *Computers in human behavior*, **23**(1), pp. 591–608.

Lensvelt-Mulders, G. J., J. J. Hox, P. G. Van der Heijden, and C. J. Maas (2005). Meta-analysis of randomized response research thirty-five years of validation. *Sociological Methods & Research*, **33**(3), pp. 319–348.

Leysieffer, F. W. and S. L. Warner (1976). Respondent jeopardy and optimal designs in randomized response models. *Journal of the American Statistical Association*, **71**(355), pp. 649–656.

Ljungqvist, L. (1993). A unified approach to measures of privacy in randomized response models: A utilitarian perspective. *Journal of the American Statistical Association*, **88**(421), pp. 97–103.

Loynes, R. (1976). Asymptotically optimal randomized response procedures. *Journal of the American Statistical Association*, **71**(356), pp. 924–928.

Mangat, N. (1994). An improved randomized response strategy. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 93–95.

Martínez Sánchez, F. and A. Rosa García (2012). Measuring the Frequency of Cheating Among Students Using a Randomized List Technique. *@ tic. revista d'innovació educativa*, (10), pp. 55–59.

Moshagen, M., B. E. Hilbig, E. Erdfelder, and A. Moritz (2014). An experimental validation method for questioning techniques that assess sensitive issues. *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)*, **61**(1), pp. 48–54.

Musch, J., A. Bröder, and K. C. Klauer (2001). Improving survey research on the World-Wide Web using the randomized response technique. *Dimensions of Internet science*, pp. 179–192.

Myerson, R. B. (1991). Game theory: analysis of conflict. *Harvard University.*

Nayak, T. K. (1994). On randomized response surveys for estimating a proportion. *Communications in Statistics-Theory and Methods*, **23**(11), pp. 3303–3321.

Schill, D. and P. A. Kline (1995). Use of random response to estimate angler non-compliance with fishing regulations. *North American Journal of Fisheries Management*, **15**(4), pp. 721–731.

Solomon, J., S. K. Jacobson, K. D. Wald, and M. Gavin (2007). Estimating illegal resource use at a Ugandan park with the randomized response technique. *Human Dimensions of Wildlife*, **12**(2), pp. 75–88.

Striegel, H., R. Ulrich, and P. Simon (2010). Randomized response estimates for doping and illicit drug use in elite athletes. *Drug and alcohol dependence*, **106**(2), pp. 230–232.

Tourangeau, R. and T. Yan (2007). Sensitive questions in surveys. *Psychological bulletin*, **133**(5), p. 859.

Umesh, U. and R. A. Peterson (1991). A Critical Evaluation of the Randomized Response Method Applications, Validation, and Research Agenda. *Sociological Methods & Research*, **20**(1), pp. 104–138.

Van der Heijden, P. G., G. van Gils, J. Bouts, and J. Hox (1998). A comparison of randomized response, CASAQ, and direct questioning; eliciting sensitive information in the context of social security fraud. *Kwantitatieve Methoden*, **59**, pp. 15–34.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, **60**(309), pp. 63–69.

Wimbush, J. C. and D. R. Dalton (1997). Base rate for employee theft: Convergence of multiple methods. *Journal of Applied Psychology*, **82**(5), p. 756.

Winkler, R. L. and L. A. Franklin (1979). Warner's randomized response model: A Bayesian approach. *Journal of the American Statistical Association*, **74**(365), pp. 207–214.

Wolter, F. and P. Preisendörfer (2013). Asking Sensitive Questions An Evaluation of the Randomized Response Technique Versus Direct Questioning Using Individual Validation Data. *Sociological Methods & Research*, **42**(3), pp. 321–353.

Wright, V. L. (1980). Use of randomized response technique to estimate deer poaching. *Wildlife Society Bulletin*, pp. 342–344.

Zaizai, Y. and N. Zankan (2004). A fair comparison of the randomized response strategies. *Acta Mathematica Scientia A*, **24**(3), pp. 362–368.