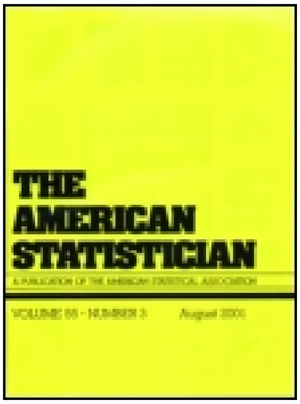

The American Statistician
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/utas20

Sample Surveys With Sensitive Questions: A Nonrandomized Response Approach

Ming T. Tan, , Guo-Liang Tian, and Man-Lai Tang
Ming T. Tan is Professor, Division of Biostatistics, University of Maryland Greenebaum Cancer Center, and Department of Epidemiology and Preventive Medicine, 10 South Pine Street, Baltimore, MD 21201 . Guo-Liang Tian is Associate Professor, Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong, P. R. China. . Man-Lai Tang is Associate Professor, Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, P. R. China. . The authors thank the editor, an associate editor, and a referee for their comments and suggestions. The research of M. T. Tan was supported partially by NIH grant RO3CA119758. The research of M. L. Tang was fully supported by a grant (project no. HKBU261508) from the Research Grant Council of the Hong Kong Special Administrative Region.

Version of record first published: 01 Jan 2012

To cite this article: Ming T. Tan, , Guo-Liang Tian, and Man-Lai Tang (2009): Sample Surveys With Sensitive Questions: A Nonrandomized Response Approach, The American Statistician, 63:1, 9-16

To link to this article: http://dx.doi.org/10.1198/tast.2009.0002

Sample Surveys With Sensitive Questions: A Nonrandomized Response Approach

Ming T. TAN, Guo-Liang TIAN, and Man-Lai TANG

Since the Warner's randomized response (RR) model to solicit sensitive information was proposed in 1965, it has been used and extended in a broad range of surveys involving sensitive questions. However, it is limited, for example, by a lack of reproducibility and trust from the interviewees as well as higher cost due to the use of randomizing devices. Recent developments of the alternative non-randomized response (NRR) approach have shown the promise to alleviate or eliminate such limitations. However, the efficiency and feasibility of the NRR models have not been adequately studied. This article introduces briefly the NRR approach, proposes several new NRR models, compares the efficiency of the NRR and RR models and studies the feasibility of the NRR models. In addition, we propose the concept of the degree of privacy protection between the NRR model and the Warner model to reflect the extent the privacy is protected. These studies show that not only the NRR approach is free of the limitations of the randomized approach but also the NRR model actually increases the relative efficiency and the degree of privacy protection. Thus, the nonrandomized response approach offers an attractive alternative to the randomized response approach.

KEY WORDS: Nonrandomized response models; Randomized response technique; Randomizing device; Sensitive questions; Warner model.

1. INTRODUCTION

Acquirement of sensitive information is often needed in a broad range of statistical applications. For instance, some behavioral and social studies may need to solicit information on reproductive history, sexual behavior, illegal drug usage, family violence, and income. When being asked these sensitive survey questions directly, some respondents may refuse to answer or may provide untruthful answers to protect their privacy. The problem is even more complicated with surveys in diverse populations, because of the interaction of sensitivity and respondent diversity. It is therefore difficult to draw valid inferences from these inaccurate data, which include refusal bias, response bias, and perhaps both. It has long been a challenge to obtain such information while having the privacy of the respondent protected and the resulting data analyzed properly.

The first attempt to overcome the difficulty was Warner's (1965) randomized response (RR) approach, which aims to encourage truthful answers from the respondents. The RR technique uses a randomization device (RD) to diffuse sensitivity (see Section 2.1). Since the introduction of the Warner model, voluminous works related to the RR technique have been done and they can be classified roughly into four major areas: (1) efficiency improvement by refining the Warner RD, (2) extensions to multichotomous categories, (3) introduction of nonsensitive questions into the RR models, and (4) inclusion of multiple sensitive questions (see Section 2.2 for a more detailed review). On the other hand, the RR models have been criticized because of their (1) inefficiency, (2) lack of reproductivity, (3) resulting lack of confidence/trust from the respondents, (4) involvement of sensitive questions or their complements, and (5) dependence on the RD (see Section 2.3).

To overcome some of the aforementioned inadequacies, recently, a totally different approach—the nonrandomized response (NRR) model—was proposed by Yu, Tian, and Tang (2008) for a single sensitive question, and by Tian, Tang, and Geng (2007) for two sensitive questions. Unlike the RR models, the NRR models use an independent nonsensitive question (e.g., season of birth) in the questionnaire to obtain indirectly a respondent's answer to a sensitive question. Obviously, the NRR designs reduce costs by doing away with the RD. However, the efficiency of the NRR models over the RR and the feasibility (i.e., the magnitude of the resulting inflation of sample size) of the NRR models were not addressed in these two initial articles, which has hindered the application of this promising approach. In addition, the two articles did not discuss the degree of privacy protection.

The main objective of this article is to study further the efficiency and feasibility of the NRR models while further evaluating the Warner model and its related designs. In Section 2, we reformulate the original Warner model to study its efficiency and then review other RR models to summarize their pros and cons. In Section 3, we introduce an alternative nonrandomized triangular model and derive the variance of the estimator, the relative efficiency (sample size) of the design to the design of direct questioning (DDQ) and the degree of privacy protection (DPP), and then we briefly review other NRR models. In Section 4, we propose a nonrandomized version for the Warner model that places the comparison of efficiency of the randomized model and nonrandomized model on the same footing. We then compare the relative efficiency of the Warner model with the

Ming T. Tan is Professor, Division of Biostatistics, University of Maryland Greenebaum Cancer Center, and Department of Epidemiology and Preventive Medicine, 10 South Pine Street, Baltimore, MD 21201 (E-mail: *mttan@som.umaryland.edu*). Guo-Liang Tian is Associate Professor, Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong, P. R. China. (E-mail: *gltian@hku.hk*). Man-Lai Tang is Associate Professor, Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, P. R. China. (E-mail: *mltang@math.hkbu.edu.hk*). The authors thank the editor, an associate editor, and a referee for their comments and suggestions. The research of M. T. Tan was supported partially by NIH grant RO3CA119758. The research of M. L. Tang was fully supported by a grant (project no. HKBU261508) from the Research Grant Council of the Hong Kong Special Administrative Region.

DOI: 10.1198/tast.2009.0002

triangular model theoretically and their degree of privacy protection numerically. A summary is given in Section 5.

2. RANDOMIZED RESPONSE MODELS

2.1 The Warner Model

Warner (1965) considered the situation in which the respondents in a population can be divided into two mutually exclusive groups: one group with a stigmatizing characteristic (e.g., ever cheated in an examination) and the other group without such a characteristic (never cheated in an examination). The following sensitive questions are then presented to respondents:

(a) I have cheated in an examination (i.e., I belong to class $\mathcal{A}$).

(b) I have never cheated in an examination (i.e., I do not belong to class $\mathcal{A}$).

2.1.1 Survey design. The respondent is directed to answer statement (a) or (b) privately by means of an RD (e.g., a dice or spinner) without indicating to the interviewer which question is being answered. Thus, the interviewer receives the response (e.g., yes or no) from a respondent without the knowledge of which question is being answered. This alleviates the concerns for privacy and reduces the number of refusals to respond to questions with untruthful answers. Here, the probability (say, p) of assigning statement (a) to a respondent by the RD is designed to be known. In other words, the RD is controlled by the interviewer.

2.1.2 Estimation and relative efficiency. Suppose that we want to estimate the proportion π of the population belonging to the sensitive class $\mathcal{A}$. Let n' be the number of yes answers obtained from the n respondents selected by simple random sampling with replacement. Warner (1965) obtained the maximum likelihood estimator (MLE)

$$\hat{\pi}_W = \frac{p - 1 + n'/n}{2p - 1}, \quad p \neq 0.5, \tag{2.1}$$

which is unbiased with variance

$$\mathrm{var}(\hat{\pi}_W) = \mathrm{var}(\hat{\pi}_D) + \frac{p(1-p)}{n(2p-1)^2}, \tag{2.2}$$

where

$$\mathrm{var}(\hat{\pi}_D) = \pi(1-\pi)/n \tag{2.3}$$

denotes the variance of $\hat{\pi}_D$ corresponding to the DDQ. For any fixed π, we note that

$$n\,\mathrm{var}(\hat{\pi}_W) = \pi(1-\pi) + \frac{p(1-p)}{(2p-1)^2} \tag{2.4}$$

is an increasing function of p when $0 < p < 0.5$, which quickly approaches to infinity as $p \to 0.5$, and then becomes a decreasing function of p when $0.5 < p < 1$ (Figure 1).

The most serious limitation of the Warner model is perhaps its inefficiency when compared with the DDQ. Note that the second term of Equation (2.2) is introduced because of the RD. When $p = 0$ or $p = 1$, the Warner model is reduced to the DDQ.

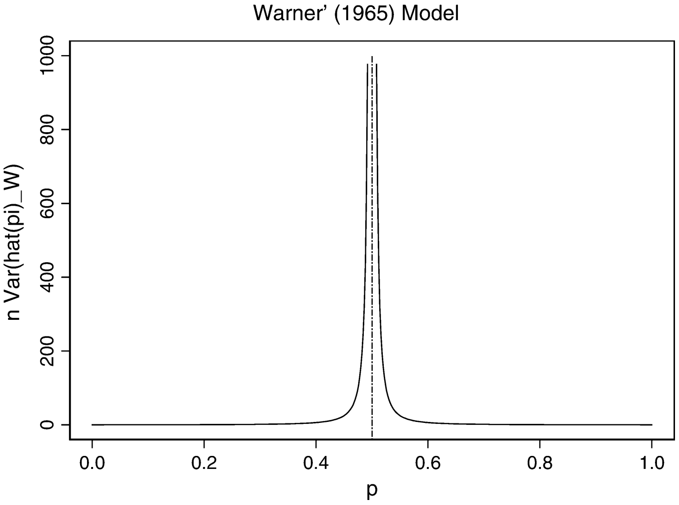

Figure 1. Plot of n var($\hat{\pi}_W$) defined by Equation (2.4) against p for the Warner model.

We define the relative efficiency (RE) of the Warner design to the DDQ by

$$\mathrm{RE}_{W\to D}(\pi, p) \mathrel{\hat{=}} \frac{\mathrm{var}(\hat{\pi}_W)}{\mathrm{var}(\hat{\pi}_D)} = 1 + \frac{p(1-p)/(2p-1)^2}{\pi(1-\pi)}, \tag{2.5}$$

which is independent of the sample size n.

Let n_W and n_D be the sample sizes required for the Warner design and the DDQ, respectively. To achieve the same estimation precision, we want

$$\frac{\pi(1-\pi)}{n_W} + \frac{p(1-p)}{n_W(2p-1)^2} = \frac{\pi(1-\pi)}{n_D}.$$

From Equation (2.5), it follows that $(n_W/n_D) = \mathrm{RE}_{W\to D}(\pi, p)$. Thus the relative efficiency is directly linked to the feasibility as defined by the ratio of the sample sizes of the two designs.

Table 1 reports some $\mathrm{RE}_{W\to D}$ (π, p)'s for various combinations of π and p. For example, when $\pi = 0.3$ and $p = 0.35$, we have $\mathrm{RE}_{W\to D}$ (0.3, 0.35) = 13.037, which implies that the sample size required for the Warner design is about 13 times that required for the DDQ to achieve the same estimation precision. To reduce the variability, we have to increase the sample size n, resulting in a significant increase in cost.

2.1.3 Degree of privacy protection. Intuitively, the optimal DPP is reached at $p = 0.5$, which corresponds to the case of infinite variance (see Fig. 1). When p is either too small or too large, the privacy of respondents cannot be protected sufficiently.

Table 1. Relative efficiency $\mathrm{RE}_{W\to D}(\pi, p)$ for various combinations of π and p

	p				
π	0.25	0.30	0.35	0.40	0.45
0.05	16.789	28.632	54.216	127.31	522.053
0.10	9.3333	15.583	29.086	67.666	276.000
0.20	5.6875	9.2031	16.798	38.500	155.687
0.30	4.5714	7.2500	13.037	29.571	118.857
0.40	4.1250	6.4687	11.532	26.000	104.125
0.50	4.0000	6.2500	11.111	25.000	100.000

Therefore, investigators have to select p within the interval [0.25, 0.45] (because n var$(\hat{\pi}_W)$ is a symmetric function at $p = 0.5$) and thus adopt an uneven (biased coin-based) RD. Because the sensitive information of a respondent regarding his/her membership in sensitive class $\mathcal{A}$ is characterized through Pr $(\mathcal{A}|\text{yes})$ and Pr $(\mathcal{A}|\text{no})$, we define

$$\text{DPP}_{\text{yes}}(\pi, p) \hat{=} \Pr(\mathcal{A}|\text{yes}) = \frac{\pi p}{\pi p + (1-\pi)(1-p)} \quad (2.6)$$

and

$$\text{DPP}_{\text{no}}(\pi, p) \hat{=} \Pr(\mathcal{A}|\text{no}) = \frac{\pi(1-p)}{\pi(1-p) + (1-\pi)p} \quad (2.7)$$

to measure the private information divulged with the Warner model.

Figure 2 shows that for a fixed π, DPP$_{\text{yes}}(\pi, p)$ is a monotonically increasing function of p whereas DPP$_{\text{no}}(\pi, p)$ is a monotonically decreasing function of p. In particular, for any $\pi \in (0, 1)$, when $p = 0.5$, we have DPP$_{\text{yes}}(\pi, 0.5)$ = DPP$_{\text{no}}$ $(\pi, 0.5) = \pi$, implying that in this case, Pr$(\mathcal{A}|\text{yes})$ = Pr$(\mathcal{A}|\text{no})$ = Pr$(\mathcal{A})$. In other words, whether the respondent belongs to the sensitive class $\mathcal{A}$ does not depend on the answer yes or no, thus arriving at the maximum DPP. However, when $p \in [0.25, 0.45]$, we have DPP$_{\text{yes}}(\pi, p) < \pi$ whereas DPP$_{\text{no}}(\pi, p) \geq \pi$.

2.2 Other Randomized Response Models

Since the publication of the Warner RR model in 1965, voluminous enhancements of the RR technique have been proposed during the past 43 years and they roughly fall into four areas: (1) efficiency improvement by refining the Warner RD, (2) extensions to multichotomous categories, (3) introduction of nonsensitive questions into the RR models, and (4) inclusion of multiple sensitive questions. In the first area, for example, Mangat and Singh (1990) suggested a two-stage RR model, which requires the use of two RDs. Mangat (1994) considered another RR model in which each respondent who is selected in the sample is requested to report yes if he or she belongs to the sensitive class $\mathcal{A}$; otherwise, he or she is instructed to use the Warner device. The Mangat (1994) model can be shown to be more efficient than both the Warner (1965) and the Mangat and Singh (1990) models. Chang and Liang (1996) developed a new two-stage unrelated RR procedure based on the model of Mangat and Singh (1990) and the unrelated question model of Horvitz, Shah, and Simmons (1967). Zou (1997) claimed that for both two-stage RR procedures of Mangat and Singh (1990) and Chang and Liang (1996), there was a simpler single-stage procedure that leads to the same distribution of responses. However, this result is not true for all RR models, as shown by Bhargava and Singh (2002). More recently, Gjestvang and Singh (2006) refined the two-stage randomization by adjusting parameters of the RD that results in a RR model more efficient than the models by Warner (1965), Mangat and Singh (1990), and Mangat (1994). However, the refined two-stage RD is convoluted and is too complicated to be practical.

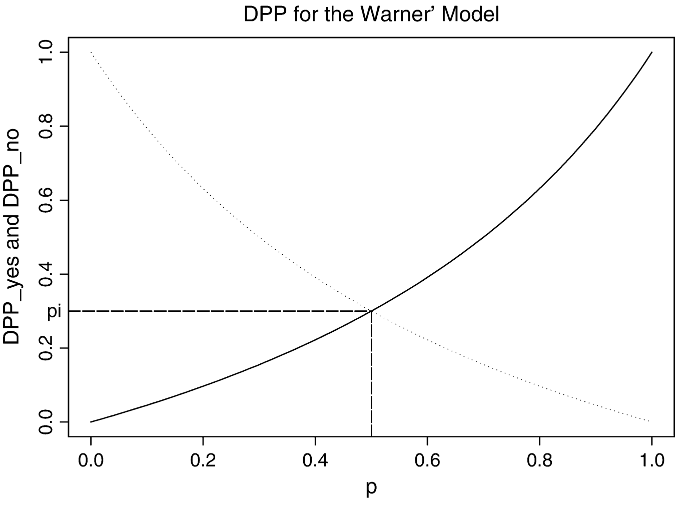

Figure 2. Plots of DPP$_{\text{yes}}(\pi, p)$ (denoted by solid line) and DPP$_{\text{no}}(\pi, p)$ (denoted by dashed line) against p for the Warner model for a given π.

In the second area, for example, Abul-Ela, Greenberg, and Horvitz (1967) and Bourke (1982) extended the Warner model to the trichotomous case to estimate the proportions of three mutually exclusive groups with at least one sensitive group. Eriksson (1973) showed how multinomial proportions can be estimated with only one sample using a different randomizing device. Liu, Chow, and Mosley (1975) developed a new randomizing device that can be used in the multiproportions cases. Franklin (1989) considered a dichotomous population with the use of a randomizing device for continuous distributions. The model was further extended to estimate any m proportions ($m > 3$) when all the m group characteristics are mutually exclusive, with at least one and at most $m - 1$ of them sensitive. Bourke and Dalenius (1973) suggested a Latin square measurement design to extend the Warner design to the multichotomous case.

The third area is the unrelated question RR model, which involves nonsensitive questions (Horvitz, Shah, and Simmons 1967; Greenberg et al. 1969). They suggested that the respondents might be more cooperative if we replace statement (b) in the original Warner model by (c) I was born in the month of April (i.e., I am a member of class U), which is nonsensitive and unrelated to statement (a). His or her privacy can be protected because the RD is operated by the respondent and the interviewer does not know which question has been answered. Recall that p is the probability of selecting statement (a). Let π' denote the proportion of individuals in the population who would answer yes to statement (c). If π' is known, only one sample is required to estimate the proportion π of the population belonging to the sensitive class $\mathcal{A}$. An unbiased estimator of π is

$$\hat{\pi}_U = \frac{n'/n - (1-p)\pi'}{p}$$

with variance

$$\text{var}(\hat{\pi}_U) = \frac{\pi(1-\pi)}{n} + \frac{(1-p)^2\pi'(1-\pi') + p(1-p)(\pi + \pi' - 2\pi\pi')}{np^2}.$$

If π' is unknown, two independent samples of size n_1 and n_2 with different probabilities p_1 and p_2 for two RDs are required. Moors (1971) showed that with an optimal allocation of n_1 and n_2, and $p_2 = 0$, the unrelated question model would be more efficient than the Warner model for $p_1 > 0.5$, regardless of the choice of π'. Dowling and Shachtman (1975) proved that $\text{var}(\hat{\pi}_U)$ is less than $\text{var}(\hat{\pi}_W)$ for all π and π', provided that p (or the $\max(p_1, p_2)$ in the two-sample case) is greater than 0.339. Folsom et al. (1973) evaluated an alternative two-sample design. Eriksson (1973) and Bourke (1974) generalized the unrelated question model to the situation with m mutually exclusive classes with up to $m - 1$ sensitive categories. The basic design uses a deck of cards, with each card containing a number of statements. However, all the aforementioned designs require the use of one or two RDs.

The fourth area is the extensions to multiple sensitive questions. Fox and Tracy (1984) considered the estimation of correlation between two sensitive questions. Lakshmi and Raghavarao (1992) also discussed a two-by-two contingency table based on binary randomized responses. Christofides (2005) presented an RR technique that can be used to estimate the proportion of individuals having two sensitive characteristics at the same time. Kim and Warde (2005) considered a multinomial RR model that can handle situations in which untruthful responses may occur. They also derived the Pearson product moment correlation estimator, which may be used to quantify the linear relationship between two multinomial variables according to an RR procedure.

More comprehensive reviews on the RR-based techniques can be found in, for example, Greenberg, Horvitz, and Abernathy (1974, 1986); Horvitz, Greenberg, and Abernathy (1975, 1976); Greenberg, Abernathy, and Horvitz (1986); Tracy and Mangat (1996); and Franklin, (1998); and in monographs including those by Cochran (1977, pp. 392–395), Chaudhuri and Mukerjee (1988), Hedayat and Sinha (1991, chap. 11), and Chaudhuri and Stenger (1992, Chap. 10); and more recently Kim and Warde (2004), Kim and Elam (2005), and Saha (2006).

2.3 Limitations of the Randomized Response Models

Despite these advances, all RR models have several well-recognized inadequacies that have limited their applications. First, the major undesirable feature of the RR method is its lack of reproducibility. That is, the same respondent may yield different answers depending on the outcome of the randomization, which results in significant loss of efficiency. Second, the uneven RD is controlled by interviewers, which makes it difficult to convince the respondents that their privacy is protected by randomization (Kuk, 1990). Warner (1986), for example, remarked that "interviewees often act as if they believe the RD simply determines whether they are required to reveal or not to reveal a secret." Third, in the Warner model, it is noteworthy that statement (b) is also sensitive because it is simply a complement of statement (a). In this regard, the interviewee still needs to answer a sensitive question no matter which card he or she selects randomly. As pointed out by Franklin (1998), the Warner model implicitly makes an assumption that the respondent is sufficiently cognizant, informed, and educated to recognize and appreciate his or her anonymity. For an audience of lower educational level or of less sophistication, or in a diverse population as in health disparity studies—regardless of the explanation—he or she might elect not to reply or might provide an untruthful answer when he or she is being asked a sensitive question. Fourth, an RD must be provided to the respondent. The device suggested by Warner (1965) is a spinner with an arrow pointer. Greenberg, Abernathy, and Horvitz (1986) pointed out the limitations of this device and other RDs designed by other authors. The extra cost in making such device is a concern. Last, the application of the Warner model has been limited almost exclusively to face-to-face personal interviews. It seems to be unfeasible for a mail questionnaire. For the unrelated question model, the respondents still need to answer a sensitive question with probability p.

3. NONRANDOMIZED RESPONSE MODELS

3.1 The Triangular Model

Let $Y = 1$ denote the class of people who possess a sensitive characteristic (e.g., drug taking) and $Y = 0$ denote the complementary class. Let W be a nonsensitive dichotomous variate and be independent of Y. The interviewer should select a suitable W so that the proportion $p = \Pr(W = 1)$ can be estimated easily. Without loss of generality, p is assumed to be known. For example, we may define $W = 1$ if the respondent was born between August and December, and $W = 0$ otherwise. Hence, it is reasonable to assume that $p \approx 5/12 = 0.41667$. Our aim is to estimate the proportion $\pi = \Pr(Y = 1)$.

3.1.1 Survey design. For a face-to-face interview, the interviewer may use the format at the left side of Table 2. The interviewee is then asked to put a tick in either the open circle or in the triangle formed by the three solid dots in Table 2 according to his or her truthful answer. In this case, $\{Y = 0, W = 0\}$ means that the interviewee was neither a drug user nor born between August and December. That is, $\{Y = 0, W = 0\}$ represents a nonsensitive subclass. On the other hand, a tick in the triangle merely indicates the interviewee was born between August and December regardless of whether he or she is a drug user. Therefore, $\{Y = 1\} \cup \{Y = 0, W = 1\}$ is also a nonsensitive subclass. Such a design encourages respondents not only to participate in the survey, but also to provide their truthful responses.

3.1.2 Alternative formulation. If $W = 1$ represents that a respondent was born between July and December, and $p \approx 1/2$, then we can reformulate the triangular model into a nonsensitive question of the following form:

Table 2. The triangular model and the corresponding cell probabilities

Categories	$W = 0$	$W = 1$	Categories	$W = 0$	$W = 1$	Total
$Y = 0$	○	●	$Y = 0$	$(1 - \pi)(1 - p)$	$(1 - \pi)p$	$1 - \pi$
$Y = 1$	●	●	$Y = 1$	$\pi(1 - p)$	πp	π
			Total	$1 - p$	p	1

Please truthfully put a tick in the circle or in the triangle formed by the three dots.

(a) I was born in the first half year, check here ○.
(b) I was born in the second half year, check here □.

A respondent is asked to put a tick truthfully in the circle or triangle if he or she is not a drug user. Otherwise, he or she is asked to put a tick in the triangle regardless of his or her actual birthday. Obviously, this nonrandomized design encourages cooperation from the respondents, and their sensitive characteristics will not be exposed to others.

In practice, some respondents in the drug-taking category may refuse to provide any answer no matter what survey design is used. One immediate advantage of the triangular model is its robustness to nonresponse in the sense that it allows such nonresponse. For example, for n respondents, we observed n_1 ticks in the circle, n_2 ticks in the triangle, and n_3 nonresponses ($n = n_1 + n_2 + n_3$). Such an observed result is equivalent to the observation $Y_{\text{obs}} = \{n_1, n_2 + n_3\}$. In other words, we observed n_1 ticks in the circle and $n_2 + n_3$ ticks in the triangle under the assumption that a respondent is always willing to answer the question if he or she is not a drug user.

3.1.3 Estimation and relative efficiency. For the triangular design, we define a "hidden" variate Y^{HT} as follows:

$$Y^{\text{HT}} = \begin{cases} 1, \text{with probability } \pi + (1-\pi)p, \text{ if a tick} \\ \quad \text{is put in the triangle,} \\ 0, \text{with probability} (1-\pi)(1-p), \text{if a tick} \\ \quad \text{is put in the circle.} \end{cases} \qquad (3.1)$$

Let $Y_{\text{obs}} = \{y_i^{\text{HT}} : i = 1, \ldots, n\}$ denote the observed data for the n respondents, where $y_i^{\text{HT}} = 1$ if the i-th respondent puts a tick in the triangle; but $y_i^{\text{HT}} = 0$ otherwise. The likelihood function for π is

$$L(\pi|Y_{\text{obs}}) = \prod_{i=1}^{n} [\pi + (1-\pi)p]^{y_i^{\text{HT}}} [(1-\pi)(1-p)]^{1-y_i^{\text{HT}}}.$$

The resulting MLE of π and the corresponding variances is given by

$$\hat{\pi}_T = \frac{\bar{y}^{\text{HT}} - p}{1-p}, \quad \text{var}(\hat{\pi}_T) = \text{var}(\hat{\pi}_D) + \frac{p(1-\pi)}{n(1-p)}, \qquad (3.2)$$

where $\bar{y}^{\text{HT}} = (1/n) \sum_{i=1}^{n} y_i^{\text{HT}}$ and $\text{var}(\hat{\pi}_D)$ is defined by Equation (2.3). For any fixed π, we noted that

$$n \, \text{var}(\hat{\pi}_T) = (1-\pi) \left[\pi + \frac{p}{1-p} \right] \qquad (3.3)$$

is an increasing function of p when $0 < p < 1$, and slowly approaches infinity as $p \to 1$ (Fig. 3).

Similar to Equation (2.5), the RE of the triangular design to the design of direct questioning is

$$\text{RE}_{T\to D}(\pi, p) \hat{=} \frac{\text{var}(\hat{\pi}_T)}{\text{var}(\hat{\pi}_D)} = 1 + \frac{p/(1-p)}{\pi}, \qquad (3.4)$$

which is also free from the sample size n. Table 3 displays some $\text{RE}_{T\to D}(\pi, p)$ for various combinations of π and p. For example, when $\pi = 0.4$ and $p = 5/12$, we have $\text{RE}_{T\to D}(0.4, 5/12) = 2.7857$, indicating that the sample size required for the triangular design is about 2.8 times that required for direct questioning to achieve the same estimation precision.

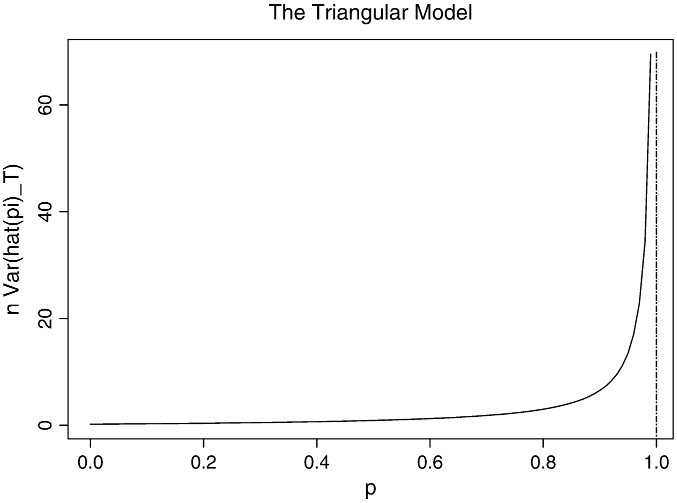

Figure 3. Plot of n var($\hat{\pi}_T$) defined by Equation (3.3) against p for the triangular model.

3.1.4 Degree of privacy protection. Intuitively, for the triangular model, the optimal DPP will be attained at $p \approx 0.5$, when π is very small. Because the privacy information divulged by a respondent regarding his or her membership in the sensitive class $\{Y = 1\}$ is characterized through $\Pr(Y = 1|Y^{\text{HT}} = 0)$ and $\Pr(Y = 1|Y^{\text{HT}} = 1)$, we have

$$\text{DPP}_{\text{C}}(\pi, p) \hat{=} \Pr(Y = 1|Y^{\text{HT}} = 0) = 0 \qquad (3.5)$$

and

$$\text{DPP}_{\text{T}}(\pi, p) \hat{=} \Pr(Y = 1|Y^{\text{HT}} = 1) = \frac{\pi}{\pi + (1-\pi)p}. \qquad (3.6)$$

In particular, when $p = 0$, we have $\text{DPP}_{\text{T}}(\pi, 0) = 1$, which corresponds to the design of direct questioning. When $p = 1$, we obtain $\text{DPP}_{\text{T}}(\pi, 1) = \pi$, which corresponds to the case that $W = 1$ if the respondent has a birthday between January and December. When $p = 0.5$, $\text{DPP}_{\text{T}}(\pi, 0.5) = 2\pi/(\pi + 1) > \pi$. Figure 4 shows that for any fixed π, $\text{DPP}_{\text{T}}(\pi, p)$ is a monotonic decreasing function of p with maximum 1 and minimum π.

3.2 The Hidden Sensitivity Model

This nonrandomized hidden sensitivity model was proposed recently by Tian, Tang, and Geng (2007) for surveys with

Table 3. Relative efficiency $\text{RE}_{T\to D}(\pi, p)$ for various combinations of π and p

	p						
π	0.25	0.30	0.35	0.40	0.45	5/12	0.50
0.05	7.66667	9.57143	11.7692	14.3333	17.363	15.285	21.0000
0.10	4.33333	5.28571	6.38462	7.66667	9.1818	8.1429	11.0000
0.20	2.66667	3.14286	3.69231	4.33333	5.0909	4.5714	6.00000
0.30	2.11111	2.42857	2.79487	3.22222	3.7273	3.3810	4.33333
0.40	1.83333	2.07143	2.34615	2.66667	3.0455	2.7857	3.50000
0.50	1.66667	1.85714	2.07692	2.33333	2.6364	2.4286	3.00000

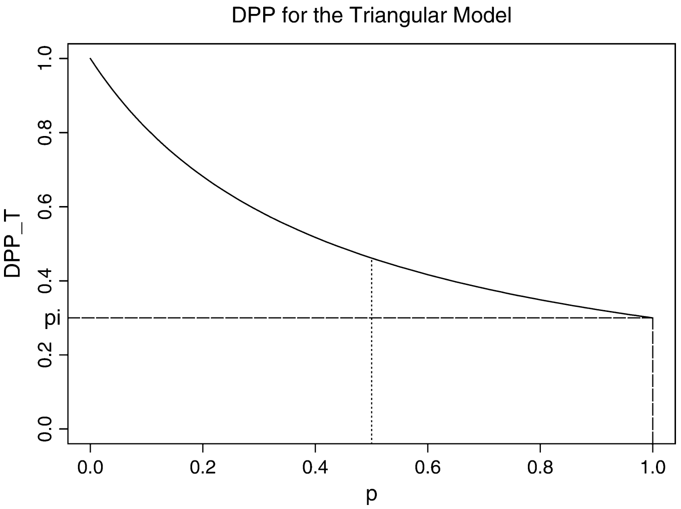

Figure 4. Plot of $\text{DPP}_{\text{T}}(\pi, p)$ defined by Equation (3.6) against p for the triangular model for a given π.

sensitive questions to assess the association of two binary sensitive characteristics. Using the hidden sensitivity model, respondents only need to answer a nonsensitive question instead of the original two sensitive questions. As a result, it can protect a respondent's privacy, avoid the usage of any RD, and is applicable to both face-to-face interviews and mail questionnaires. They adopted the Expectation-Maximization (EM) algorithm to obtain the constrained MLEs of the cell probabilities and the odds ratio, and the bootstrap approach to estimate their standard errors. Likelihood ratio and chi-squared tests are developed for testing associations between the two binary variables. Simulations are performed to evaluate the empirical Type I error rates and powers for the two tests.

Note that the triangular model is designed to handle dichotomous (yes/no) responses to sensitive questions such as: Have you ever used illegal drugs? Similar to Tian, Tang, and Geng (2007), we can extend the triangular model to a nonrandomized multichotomous model for a single sensitive variate (denoted by Y) to treat multiproportional responses to sensitive questions such as: How many days did you use illegal drugs last month? For example, possible responses for this question might be 0, 1, 2 or ≥ 3. Obviously, the nonrandomized multichotomous model can be applied to sensitive questions regarding number of abortions, number of sex partners, or incidence of tax evasion.

Table 4. The crosswise model

Categories	$W=0$	$W=1$	Categories	$W=0$	$W=1$	Total
$Y=0$	○	●	$Y=0$	$(1-\pi)(1-p)$	$(1-\pi)p$	$1-\pi$
$Y=1$	●	○	$Y=1$	$\pi(1-p)$	πp	π
			Total	$1-p$	p	1

Please truthfully put a tick in the diagonal with the two circles or the off-diagonal with the two dots.
NOTE: From Yu, Tian, and Tang (2008).

4. COMPARISONS AMONG RANDOMIZED RESPONSE MODELS AND NONRANDOMIZED RESPONSE MODELS

4.1 A Nonrandomized Version for the Warner Model

To set the comparisons of different models on the same footing, we introduce a nonrandomized version for the Warner model. Let Y and W be defined as in Section 3.1, and $p = \Pr(W = 1)$ and $\pi = \Pr(Y = 1)$. The interviewer may also design the questionnaire in the format as shown in the left side of Table 4, and may ask the interviewee to put a tick truthfully in either the diagonal with the two circles or the off-diagonal with the two dots. Note that both $\{Y = 0, W = 0\} \cup \{Y = 1, W = 1\}$ and $\{Y = 0, W = 1\} \cup \{Y = 1, W = 0\}$ are nonsensitive. Thus, whether the interviewee belongs to the sensitive class is not on record. We call this the "crosswise model."

Let $Y_{\text{obs}} = \{y_i^{\text{HW}} : i = 1, \ldots, n\}$ denote the observed data for the n respondents, where $y_i^{\text{HW}} = 1$ if the ith respondent puts a tick in the main diagonal with the two circles or $y_i^{\text{HW}} = 0$ otherwise. The likelihood function for π is

$$L(\pi|Y_{\text{obs}}) = \prod_{i=1}^{n} [(1-\pi)(1-p) + \pi p]^{y_i^{\text{HW}}} \\ \times [\pi(1-p) + (1-\pi)p]^{1-y_i^{\text{HW}}}.$$

Let $n' = \sum_{i=1}^{n} y_i^{\text{HW}}$. The resulting MLE of π and the corresponding variance are exactly that given by Equations (2.1) and (2.2) for the Warner model. Hence, we consider the crosswise model a nonrandomized version of the Warner model.

In comparison with the original Warner model, this nonrandomized crosswise model has the following advantages: (1) it does not require any RD, thus reducing study costs; (2) it is easy to operate for both interviewer and interviewee; (3) the interviewee does not need to face any sensitive question directly; and (4) it can be applied to both face-to-face personal interviews and mail questionnaires.

Table 5. Relative efficiency $\text{RE}_{W\to T}(\pi, p)$ for various combinations of π and p

	p													
π	0.25	0.30	0.35	0.40	0.45	0.49	0.51	0.54	0.56	0.58	7/12	0.60	0.63	0.66
0.05	2.19	2.99	4.61	8.88	30.06	650.66	602.93	33.42	13.65	7.04	6.38	4.11	2.10	1.18
0.10	2.15	2.94	4.55	8.82	30.05	654.48	608.57	33.93	13.92	7.20	6.54	4.22	2.18	1.24
0.20	2.13	2.92	4.54	8.88	30.58	672.94	629.54	35.45	14.65	7.65	6.96	4.52	2.37	1.37
0.30	2.16	2.98	4.66	9.17	31.89	708.13	665.86	37.82	15.73	8.26	7.53	4.92	2.60	1.53
0.40	2.25	3.12	4.91	9.75	34.19	765.48	722.96	41.35	17.28	9.12	8.32	5.47	2.92	1.73
0.50	2.40	3.36	5.34	10.71	37.93	855.70	811.26	46.67	19.587	10.38	9.47	6.25	3.35	2.00

Table 6. Comparison of degrees of privacy protection for various combinations of π and p

	Warner model						Triangular model			
	$p = 0.30$		$p = 0.35$		$p = 0.40$		$p = 0.50$		$p = 7/12$	
π	$\mathrm{DPP}_{\mathrm{yes}}$	$\mathrm{DPP}_{\mathrm{no}}$	$\mathrm{DPP}_{\mathrm{yes}}$	$\mathrm{DPP}_{\mathrm{no}}$	$\mathrm{DPP}_{\mathrm{yes}}$	$\mathrm{DPP}_{\mathrm{no}}$	$\mathrm{DPP}_{\mathrm{C}}$	$\mathrm{DPP}_{\mathrm{T}}$	$\mathrm{DPP}_{\mathrm{C}}$	$\mathrm{DPP}_{\mathrm{T}}$
0.05	0.022	0.109	0.027	0.089	0.033	0.073	0	0.095	0	0.082
0.10	0.045	0.205	0.056	0.171	0.068	0.142	0	0.181	0	0.160
0.20	0.096	0.368	0.118	0.317	0.142	0.272	0	0.333	0	0.300
0.30	0.155	0.500	0.187	0.443	0.222	0.391	0	0.461	0	0.423
0.40	0.222	0.608	0.264	0.553	0.307	0.500	0	0.571	0	0.533
0.50	0.300	0.700	0.350	0.650	0.400	0.600	0	0.666	0	0.631

NOTE: $\mathrm{DPP}_{\mathrm{yes}}$, $\mathrm{DPP}_{\mathrm{no}}$, $\mathrm{DPP}_{\mathrm{C}}$, and $\mathrm{DPP}_{\mathrm{T}}$ are defined by Equations (2.6), (2.7), (3.5), and (3.6), respectively.

4.2 Relative Efficiency of the Warner Model to the Triangular Model

Equivalence between the randomized Warner model and the nonrandomized crosswise model in efficiency provides a basis for comparing the Warner model with the nonrandomized triangular model. Using the variance criterion, we obtain from Equations (2.2) and (3.2)

$$\mathrm{var}(\hat{\pi}_W) - \mathrm{var}(\hat{\pi}_T) = \frac{p}{n(1-p)(2p-1)^2} \cdot [(1-p)^2 - (1-\pi) \times (2p-1)^2].$$

We have the following result.

Theorem 1. The nonrandomized triangular model is always more efficient than the randomized Warner model for any $\pi \in (0, 1)$ and any $p < (2/3)$. Specifically, we have

$$\mathrm{var}(\hat{\pi}_W) \geq \mathrm{var}(\hat{\pi}_T), \quad \text{for } 0 < p \leq p_\pi \; (p \neq 0.5), \qquad (4.1)$$

where $(2/3) \leq p_\pi = (2\pi - 1 - \sqrt{1-\pi})/(4\pi - 3)$ is an increasing function of π.

Furthermore, we consider the RE of the Warner model to the triangular model,

$$\mathrm{RE}_{W \to T}(\pi, p) = \frac{\mathrm{var}(\hat{\pi}_W)}{\mathrm{var}(\hat{\pi}_T)} = \frac{\pi + p(1-p)/[(2p-1)^2(1-\pi)]}{\pi + p/(1-p)}, (p \neq 0.5), \qquad (4.2)$$

which is independent of the sample size n and depends only on the parameters π and p. We note that interviewers usually select p between [0.25, 0.45] in the Warner model. From Table 5, when $0.25 \leq p \leq 0.45$, the efficiency of the triangular strategy is about two to 37 times that of the Warner model. In particular, when $0.49 \leq p \leq 0.51$ (which is the optimal range for which the privacy of respondents is protected), the efficiency of the triangular strategy is about 600 to 855 times that of the Warner model.

4.3 Degree of Privacy Protection

To compare the DPP of the Warner model with that of the triangular model, we consider three values (namely, $p = 0.3$, 0.35, and 0.4), which are some practical choices for the Warner model; and two values (namely, $p = 0.5$ and 7/12), which are some optimal choices for the triangular model. Table 6 gives DPPs for various combinations of π and p. From Table 6, when comparing the Warner model with $p = 0.35$ with the triangular model with $p = 7/12$, we have $\mathrm{DPP}_{\mathrm{yes}} > \mathrm{DPP}_C$ and $\mathrm{DPP}_{\mathrm{no}} > \mathrm{DPP}_T$ for all π in the table. Therefore, the triangular model with $p = 7/12$ has better DPPs than the Warner model with $p = 0.35$. We reach a similar conclusion when comparing the triangular model with $p = 0.5$ with the Warner model with $p = 0.35$.

5. SUMMARY

We have shown that the nonrandomized approach is not limited by the factors that commonly limit the randomized approach. The triangular design is consistently more efficient than the Warner design, thus providing a promising alternative to RR designs. The main advantages of the nonrandomized methods are (1) their reproducibility (i.e., the same respondent is expected to give the same answer by the design if the survey is conducted again) by eliminating randomization, (2) better protection of privacy (with the sensitive question being replaced by a nonsensitive question (see Section 3.1.2) and the nonsensitive variate W (see Section 3) being uniform ($p = 0.5$)), (3) better robustness of the triangular model to the nonresponse (see Section 3.1.2), (4) better cooperation (with the nonsensitive question not being controlled by interviewers, and the respondents being likely to trust the interviewers), (5) significant cost reduction by getting rid of the RD and an increase in efficiency over the RR (e.g., by 2- to 37-fold, see Section 4.2), and (6) applicability to different types of surveys (i.e., the NRR design is applicable to both the face-to-face interview and other forms of surveys). Last, we realize that social and behavioral aspects of the NRR designs are yet to be fully explored in actual surveys. Further development of this promising approach both in survey statistical methodology and in actual survey studies involving sensitive questions would broaden its scope of application and facilitate the acquirement of useful information needed for policy and decision making without invading privacy.

[Received September 2007. Revised September 2008.]

REFERENCES

Abul-Ela, A. A., Greenberg, B. G., and Horvitz, D. G. (1967), “A Multiproportions Randomized Response Model,” *Journal of the American Statistical Association,* 62, 990–1008.

Bhargava, M., and Singh, R. (2002), “On the Efficiency Comparison of Certain Randomized Response Strategies,” *Metrika,* 55, 191–197.

Bourke, P. D. (1974). “Multi-proportions Randomized Response Using the Unrelated Question Technique.” Report no. 74 of the Errors in Surveys Research Project. Institute of Statistics, University of Stockholm (Mimeo).

——. (1982), “Randomized Response Multivariate Designs for Categorical Data,” *Communications in Statistics Theory and Methods,* 11, 2889–2901.

Bourke, P. D., and Dalenius, T. (1973). “Multi-proportions Randomized Response Using a Single Sample.” Report no. 68 of the Errors in Surveys Research Project. Institute of Statistics, University of Stockholm (Mimeo).

Chang, H. J., and Liang, D. H. (1996), “A Two-Stage Unrelated Randomized Response Procedure,” *The Australian Journal of Statistics,* 38, 43–51.

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

Chaudhuri, A., and Stenger, H. (1992). *Survey Sampling: Theory and Methods*. New York: Marcel Dekker.

Christofides, T. C. (2005), “Randomized Response Technique for Two Sensitive Characteristics at the Same Time,” *Metrika,* 62, 53–63.

Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.), New York: Wiley.

Dowling, T. A., and Shachtman, R. H. (1975), “On the Relative Efficiency of Randomized Response Models,” *Journal of the American Statistical Association,* 70, 84–87.

Eriksson, S. A. (1973), “A New Model for Randomizing Response,” *International Statistical Review. Revue Internationale de Statistique,* 41, 101–113.

Folsom, R. E., Greenberg, B. G., Horvitz, D. G., and Abernathy, J. R. (1973), “The Two Alternate Questions Randomized Response Model for Human Surveys,” *Journal of the American Statistical Association,* 68, 525–530.

Fox, J., and Tracy, P. (1984), “Measuring Associations With Randomized Response,” *Social Science Research,* 13, 188–197.

Franklin, L. A. (1989), “Randomized Response Sampling From Dichotomous Populations With Continuous Randomization,” *Survey Methodology,* 15, 225–235.

——. (1998), “Randomized Response Techniques,” in *Encyclopedia of Biostatistics*, ed. P. Armitage and T. Colton, New York: Wiley, pp. 3696–3703.

Gjestvang, C. R., and Singh, S. (2006), “A New Randomized Response Model,” *Journal of the Royal Statistical Society* Ser. B, 68, 523–530.

Greenberg, B. G., Abernathy, J. R., and Horvitz, D. G. (1986), “Randomized Response,” in *Encyclopedia of Statistical Sciences* (Vol. 7), ed. S. Kotz and N.L. Johnson, New York: Wiley, pp. 540–548.

Greenberg, B. G., Abul-Ela, A. A., Simmons, W. R., and Horvitz, D. G. (1969), “The Unrelated Question Randomized Response Model: Theoretical Framework,” *Journal of the American Statistical Association,* 64, 520–539.

Greenberg, B. G., Horvitz, D. G., and Abernathy, J. R. (1974), “Comparison of Randomized Response Designs,” in *Reliability and Biometry: Statistical Analysis of Life Length*, ed. F. Prochan and R.J. Serfling, Philadelphia: SIAM, pp. 787–815.

Hedayat, A. S., and Sinha, B. K. (1991), *Design and Inference in Finite Population Sampling*. New York: Wiley.

Horvitz, D. G., Greenberg, B. G., and Abernathy, J. R. (1975), “Recent Developments in Randomized Designs,” in *A Survey of Statistical Design and Linear Models*, ed. J. N. Srivastava, New York: North Holland/ American Elsevier Publishing, pp. 271–285.

——. “Randomized Response: A Data Gathering Device for Sensitive Questions,” *International Statistical Review. Revue Internationale de Statistique,* 44, 181–196.

Horvitz, D. G., Shah, B. V., and Simmons, W. R. (1967), “The Unrelated Question Randomized Response Model,” in *1967 Proceedings of the Social Statistics Section*, American Statistical Association, pp. 65–72.

Kim, J. M., and Elam, M. E. (2005), “A Two-Stage Stratified Warner’s Randomized Response Model Using Optimal Allocation,” *Metrika,* 61, 1–7.

Kim, J. M., and Warde, W. D. (2004), “A Stratified Warner’s Randomized Response Model,” *Journal of Statistical Planning and Inference,* 120, 155–165.

——. (2005), “Some New Results on the Multinomial Randomized Response Model,” *Communications in Statistics Theory and Methods,* 34, 847–856.

Kuk, A. Y. C. (1990), “Asking Sensitive Questions Indirectly,” *Biometrika,* 77, 436–438.

Lakshmi, D. V., and Raghavarao, D. (1992), “A Test for Detecting Untruthful Answering in Randomized Response Procedures,” *Journal of Statistical Planning and Inference,* 31, 387–390.

Liu, P. T., Chow, L. P., and Mosley, W. H. (1975), “Use of the Randomized Response Technique With a New Randomizing Device,” *Journal of the American Statistical Association,* 70, 329–332.

Mangat, N. S. (1994), “An Improved Randomized Response Strategy,” *Journal of the Royal Statistical Society,* Ser. B, 56, 93–95.

Mangat, N. S., and Singh, R. (1990), “An Alternative Randomized Response Procedure,” *Biometrika,* 77, 439–442.

Moors, J. J. A. (1971), “Optimization of the Unrelated Question Randomized Response Model,” *Journal of the American Statistical Association,* 66, 627–629.

Saha, A. (2006), “Optimal Randomized Response in Stratified Unequal Probability Sampling: A Simulation Based Numerical Study With Kuk’s Method,” *Test,* 16, 346–354.

Tian, G. L., Yu, J. W., Tang, M. L., and Geng, Z. (2007), “A New Non-randomized Model for Analyzing Sensitive Questions With Binary Outcomes,” *Statistics in Medicine,* 26, 4238–4252.

Tracy, D. S., and Mangat, N. S. (1996), “Some Developments in Randomized Response Sampling During the Last Decade: A Follow up of Review by Chaudhuri and Mukerjee,” *Journal of Applied Statistical Science,* 4, 147–159.

Warner, S. L. (1965), “Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias,” *Journal of the American Statistical Association,* 60, 63–69.

——. (1986), “The Omitted Digit Randomized Response Model for Telephone Applications,” in *Proc. Survey Res. Meth. Sect.*, American Statistical Association, pp. 441–443.

Yu, J. W., Tian, G. L., and Tang, M. L. (2008), “Two New Models for Survey Sampling With Sensitive Characteristic: Design and Analysis,” *Metrika,* 67, 251–263.

Zou, G. H. (1997), “Two-Stage Randomized Response Procedures as Single Stage Procedures,” *The Australian Journal of Statistics,* 39, 235–236.