

On the Estimation of Population Mean and Sensitivity in Two-Stage Optional Randomized Response Model

Sat Gupta and Javid Shabbir¹

Department of Mathematics and Statistics, University of North Carolina, Greensboro, USA

SUMMARY

The paper discusses a two-stage optional randomized response model and presents estimators for the mean and sensitivity level of a sensitive question. A simulation study is used to assess the validity of the proposed estimators.

Key words : Two-stage randomized response models, Efficiency comparison, Optional scrambling, Sensitivity estimation, Simulation study, Unbiased estimation.

1. INTRODUCTION

Randomized response techniques (RRT) have been widely used for personal interview surveys ever since the pioneering work of Warner (1965). Umesh and Peterson (1991), Scheers (1992) and Hosseini and Armacost (1993), among others, have shown that RRT methods do, in fact, lead to more valid answers and prove effective in circumventing social desirability bias. Several randomized response models have been developed by researchers for collecting data on both the qualitative and quantitative variables. Greenberg *et al.* (1971) have proposed the unrelated question model for estimating the mean and the variance of a sensitive quantitative variable. Mangat and Singh (1990) have introduced a two-stage RRT model for estimating the prevalence of a sensitive trait in a binary population and showed that an improvement over the Warner (1965) model is possible. Eichhorn and Hayre (1983) discussed a multiplicative RRT model for quantitative responses. Gupta *et al.* (2002) modified the Eichhorn and Hayre (1983) model and introduced an optional RRT model for estimating simultaneously the mean as well as the sensitivity level of a sensitive variable. They defined sensitivity to be the proportion of subjects in a population who consider a given question to be of sensitive nature. The argument

put forward by Gupta *et al.* (2002) was that a question may be sensitive to one person but may not be sensitive to another.

Ryu *et al.* (2006) have recently attempted to combine the models introduced by Mangat and Singh (1990) and by Gupta *et al.* (2002) to introduce an estimator of the mean of a quantitative sensitive variable and show that their estimator of the mean is more efficient than that of the estimator by Gupta *et al.* (2002). However, the model by Ryu *et al.* (2006) is not an optional RRT model and does not estimate the sensitivity level of the sensitive question, unlike Gupta *et al.* (2002) whose model estimated simultaneously both mean and sensitivity of the sensitive variable.

The focus of this paper is on introducing a truly optional two-stage RRT model and on comparing it with the Ryu *et al.* (2006) model.

2. THEORETICAL FRAMEWORK

We first discuss a few relevant RRT models for quantitative response.

Eichhorn and Hayre Model

Let X be the true response and S be some scrambling variable, independent of X , with mean θ and standard deviation σ_S . The respondent is asked to report the response Z as given by

¹ Department of Statistics, Quaid-I-Azam University, Islamabad, Pakistan

