A Two-way Randomized Response Technique in Stratification for Tracking HIV Seroprevalence

Usman, A.¹ & Oshungade, I.O.²

1. Department of Maths, Statistics & Computer Science, Kaduna Polytechnic, Nigeria.
2. Department of Statistics, University of Ilorin, Nigeria.

* E-mail of the corresponding author: aliyusman2007@yahoo.com

Abstract

Seroprevalence surveys of HIV pandemic are highly sensitive especially in Africa. The objective of this study is to reach research frontier to devise a two-way randomized response model (RRM) in stratification and use same to estimate HIV seroprevalence rates in a given population and compare results with the existing seroprevalence rates. The randomized response techniques (RRT) guarantees the anonymity of respondents in surveys aimed at determining the frequency of stigmatic, embarrassing or criminal behaviour where direct techniques for data collection may induce respondents to refuse to answer or give false responses. The motivation was to improve upon the existing RRMs as well as to apply them to estimate HIV seroprevalence rates. Warner proposed the pioneering RRM for estimating the proportion of persons bearing a socially disapproved character. Quatember produced unified criteria for all RRTs, Kim and Warde proposed a stratified RRM and so many others. The proposed two-way RRM in stratification for HIV seroprevalence surveys was relatively more efficient than the Kim and Warde stratified estimator for a fixed sample size. The chosen design parameter was 0.7, using the criteria of Quatember who derived the statistical properties of the standardized estimator for general probability sampling and privacy protection. Furthermore, the model was used to estimate the HIV seroprevalence rate in a sampled population of adults 3,740 people aged 18 years and above attending a clinic in Kaduna, Nigeria using a sample size of 550. The findings revealed that HIV seroprevalence rate, as estimated by the Model, stood at 6.1% with a standard error of 0.0082 and a 95% confidence interval of [4.5%, 7.7%]. These results are consistent with that of Nigerian sentinel survey (2003) conducted by NACA, USAID and CDC which estimated the HIV seroprevalence in Kaduna State as 6.0%. Hence, the RRTs herein can serve as new viable methods for HIV seroprevalence surveys.

Key words
Randomized response techniques, two-way randomized response models, seroprevalence rates, design parameter, efficiency, sentinel surveys, stratified random sampling

1. Introduction

Nonresponse in sample surveys may cause a biased estimation of unknown population parameters as well as increase of the variance of the estimates. The randomized response techniques (RRTs) were especially developed to improve the accuracy of answers to sensitive questions. Socially sensitive questions are thought to be threatening to respondents (Lee, 1993). When sensitive topics are studied, respondents often react in ways that negatively affect the validity of the data. Such a threat to the validity of the results is the respondents’ tendency to give socially desirable answers to avoid social embarrassment and to project a positive self-image (Rasinski, 1999). Warner (1965) reasoned that the reluctance of the respondents to reveal sensitive or probably harmful information would diminish when
respondents could be convinced that their anonymity was guaranteed. Hence, Warner (1965) designed the first randomized response model (RRM). The crux of his method and all other RRTs that followed is that the meaning of the respondents’ answers is hidden by a deliberate contamination of the data collection settings.

Studies with RRTs have been conducted in the areas of healthcare (Volicer & Volicer, 1982), on alcohol, drug abuse and sexual behaviour (Jarman, 1997), on child molestation (Fox and Tracy, 1986), on tax evasion (Houston & Tran, 2008), among others. Meta-analysis on 42 comparative studies showed that RRTs resulted in more valid population estimates than direct question–answer techniques (Lensvelt-Mulders et al., 2005). An advantage of using RRT when conducting sensitive research is that, the individual ‘yes’-answer becomes meaningless as it is only a ‘yes-answer’ to the random device (Van der Hout, et al., 2002).

However, the disadvantage of using RR methods is that they are less efficient than direct question designs. Since the RRTs work by adding random noise to the data, they all suffer from larger standard errors, leading to reduced power which makes it necessary to use larger samples than in question–answer designs. Unfortunately, larger samples are associated with prolonged completion time and higher research costs, making RRTs less attractive to applied researchers. This leads to the topic of efficiency versus effectiveness. Effectiveness is related to the validity of research results in the same way that efficiency is related to reliability. The randomized response design is more effective than the direct question-answer design (Lensvelt-Mulders et al., 2005). The loss of efficiency in RR designs could be compensated when the results prove to be more valid (Kuk, 1990). When the loss in efficiency can be kept as small as possible the use of a RR design to study sensitive questions will become more profitable.

2. Methodology

In order apply the two-way RRM; a study was conducted in Gwamna Awan General Hospital, Kaduna, Nigeria in November, 2011. With a carefully coordinated field work and sampling design on a population of 3,740 adults aged 18 years and above attending the Hospital using a sample size of 550. Furthermore, the model was used to estimate the HIV seroprevalence rate in the same population. Quatember (2009) both theoretically and empirically analyzed the effect of different design parameters on the performance of RRTs using different levels of privacy protection. Quatember (2009) suggested that 0.7 approximately works well for most
RRM where the questions are regarded as highly sensitive. Hence, 0.7 is the chosen design parameter and deck of 50 cards as our random device throughout.

2.1 The Proposed HIV Seroprevalence Model

In general, a randomized response model is based on \( m (m \geq 1) \) random devices and a set of rules for determining the communicating the answer. For each random device, the respondent randomly selects one of the \((s_k \geq 1, k = 1, 2, \ldots, m)\) statements and, following the rules, reports ‘yes’ or ‘no’ without revealing which questions he/she is answering. The \( k \)th random device of the RRM \( m \) is described by a vector of \( s_k - 1 \) parameters (probabilities) \( \theta_k = (p_{k1}, \ldots, p_{k(s_k-1)}) \), where \( p_{ki} \in s_{ki} \leq [0,1], s_{ki} \) is the set.

Brookmeyer and Gail (2004) defined HIV seroprevalence as the study of the number of cases where HIV is present in a specific population at a designated time. The presence of HIV in a specific individual is determined by the finding of HIV antibodies in the serum (HIV seropositivity). This study is set to develop an efficient two-way RRM in stratification particularly for HIV seroprevalence surveys and to use the Model for estimating the seroprevalence rate in a given population.

The proposed HIV seroprevalence surveys RRM requires that a sample respondent in stratum \( h \) to answer an innocuous direct question and asked to use the random device \( R_{hi} \) if his/her answer to direct question is “yes”. If answer to the direct question is “no”, he/she is requested to use another random device \( R_{h2} \) twice. Both random devices \( R_{hi} \) and \( R_{h2} \) consist of two statements (i) “I am HIV positive” and (ii) “I am HIV negative”, presented with probabilities \( P_{hi} \) and \( (1 - P_{hi}) \) respectively. Here the random device \( R_{h2} \) would to be answered twice. Hence, we can obtain the estimator of population proportion \( \pi_h \) in \( h \)th stratum based on the responses from \( R_{hi} \) as follows. The probability of a ‘yes’ response from the respondents using \( R_{hi} \) is given by:

\[
\lambda_{hi} = P_{hi} \pi_h^* + (1 - P_{hi}) \pi_{hy} = P_{hi} \pi_h^* + (1 - P_{hi})
\]  

(1)

Also, the probability of a ‘no’ response from the respondents using \( R_{hi} \) is given by:

\[
\lambda'_{hi} = P_{hi} (1 - \pi_h^*) + (1 - P_{hi})(1 - \pi_{hy}) = P_{hi}(1 - \pi_h^*)
\]  

(2)

Since the respondent using \( R_{hi} \) has already answered yes to the direct question, \( \pi_{hy} = 1 \).
Among those that answered ‘yes’ to the innocuous questions in stratum $h$; suppose that $n_{h1}$ report ‘yes’ and $(n_h - n_{h1})$ report ‘no’, the likelihood of the sample in the same stratum is given below:

$$
\xi = \left[P_{h1}\pi_h^* + (1 - P_{h1})\right]^{n_{h1}} \times \left[P_{h1} (1 - \pi_h^*)\right]^{n_h - n_{h1}}
$$

We obtain the maximum likelihood estimate (MLE) of $\pi_h^*$ as follows:

$$
\pi_h^* = \frac{n_h P_{h1} - n_h + n_{h1}}{n_h P_{h1}}
$$

Hence, the unbiased estimators in terms of the responses of the respondents using $R_{h1}$ is given by:

$$
\hat{\pi}_{h1} = \frac{\hat{\lambda}_{h1} - (1 - P_{h1})}{P_{h1}}
$$

Where; the proportion of ‘yes’ answers from $R_{h1}$ in the sample is given as:

$$
\hat{\lambda}_{h1} = \frac{n_{h1}}{n_h}
$$

The variance of is obtained as follows:

$$
Var(\hat{\pi}_{h1}) = \left[\frac{1}{P_{h1}}\right]^2 Var(\hat{\lambda}_{h1})
$$

$$
= \left[\frac{1}{P_{h1}}\right]^2 \left(\frac{\hat{\lambda}_{h1} (1 - \hat{\lambda}_{h1})}{n_{h1}}\right)
$$

$$
\therefore Var(\hat{\pi}_{h1}) = \frac{(1 - \pi_h^*) (P_{h1} \pi_h^* + 1 - P_{h1})}{n_{h1} P_{h1}}
$$

The respondent, in $h$th stratum, giving a “no” answer to the question are to use $R_{h2}$ twice to report two answers, where $R_{h2}$ consists of the two statement of Warner’s RR method. To have the first response reported the probabilities of the two statements are $P_{h2}$ and $(1 - P_{h2})$ whereas to get the second response from the responses these probabilities are $P_{h2}^*$ and $(1 - P_{h2}^*)$. Two unbiased estimators based on the two set of responses from respondents using $R_{h2}$ can be defined as follows:

$$
\pi_{h2}^* = \frac{\hat{\lambda}_{h2} - (1 - P_{h2})}{(2 P_{h2} - 1)}
$$
and \( \pi_{h2} = \frac{\hat{\lambda}_{h2} - (1 - P_{h2}^*)}{(2P_{h2}^* - 1)} \) (8)

where; 
\[ \lambda_{h1} = P_{h2}\pi_h + (1 - P_{h1})(1 - \pi_h) = (2P_{h2} - 1)\pi_h + (1 - P_{h1}) \] (9)

\[ \lambda_{h2}^* = P_{h2}^*\pi_h + (1 - P_{h2}^*)(1 - \pi_h) = (2P_{h2}^* - 1)\pi_h + (1 - P_{h2}^*) \] (10)

Which are the probabilities of “yes” responses for the first and second use of \( R_{h2} \). The variances of the estimators \( \hat{\lambda}_{h12} \) and \( \hat{\lambda}_{h22} \) are given by:

\[ \text{Var}(\hat{\lambda}_{h21}) = \frac{\lambda_{h1}(1 - \lambda_{h1})}{n_{h2}(2P_{h2} - 1)^2} = \frac{\pi_{h1}(1 - \pi_{h1})}{n_{h2}} + \frac{P_{h2}(1 - P_{h2})}{n_{h2}(2P_{h2} - 1)^2} \] (11)

and

\[ \text{Var}(\hat{\lambda}_{h22}) = \frac{\lambda_{h2}^*(1 - \lambda_{h2}^*)}{n_{h2}(2P_{h2}^* - 1)^2} = \frac{\pi_{h2}^*(1 - \pi_{h2}^*)}{n_{h2}} + \frac{P_{h2}^*(1 - P_{h2}^*)}{n_{h2}(2P_{h2}^* - 1)^2} \] (12)

These were obtained from Warner's RR model as given below. The first responses from respondents using \( R_{h2} \) can be defined as follows. The probability of a ‘yes’ response from the respondents using \( R_{h2} \) in the first response is given by:

\[ \lambda_{h12} = P_{h1}\pi_h + (1 - P_{h1})(1 - \pi_h) \] (13)

Also, the probability of a ‘no’ response from the respondents using \( R_{h2} \) in the first response is given by:

\[ \lambda'_{h12} = P_{h1}(1 - \pi_h) + (1 - P_{h1})\pi_h \] (14)

Among those that answered ‘no’ to the innocuous questions in stratum \( h \); suppose that \( n_{h2} \) report ‘yes’ and \( (n_{h} - n_{h2}) \) report ‘no’ in first case, the likelihood of the sample in the same stratum is as follows:

\[ \xi = [P_{h1}\pi_h + (1 - P_{h1})(1 - \pi_h)]^{n_{h2}} \times [P_{h1}(1 - \pi_h) + (1 - P_{h1})\pi_h]^{n_{h} - n_{h2}} \] (15)

We also obtain the MLE of \( \pi_h \), as follows:

\[ \pi_{h12} = \frac{\hat{\lambda}_{h2} - (1 - P_{h2}^*)}{(2P_{h2} - 1)} \] (16)

Where; the proportion of ‘yes’ answers from \( R_{h1} \) in the sample is given as;

\[ \hat{\lambda}_{h2} = \frac{n_{h2}}{n_h} \]

The variance of \( \xi \) is obtained as follows:

\[ \text{Var}(\pi_{h21}) = \left[ \frac{1}{n_{h}(2P_{h2} - 1)} \right]^2 \text{Var}(n_{h2}) \] (17)
Since;
\[ \pi_{h21} = \frac{n_{h2} - (1 - P_{h2})}{(2P_{h2} - 1)} = \frac{n_{h2}}{n_{h}(2P_{h2} - 1)} + \frac{P_{h2} - 1}{2P_{h2} - 1} \]

Then;
\[ Var(\pi_{h21}) = \left[ \frac{1}{n_{h}(2P_{h2} - 1)} \right]^2 Var(n_{h2}) \]
\[ = \frac{Var(X_{i2})}{n_{h}^2(2P_{h2} - 1)^2} \]
\[ = \frac{[P_{h1}\pi_{h} + (1 - P_{h1})(1 - \pi_{h})][P_{h1}(1 - \pi_{h}) + (1 - P_{h1})\pi_{h}]}{n_{h}(2P_{h2} - 1)^2} \]

Hence;
\[ Var(\hat{\pi}_{h21}) = \frac{\pi_{h2} (1 - \pi_{h2})}{n_{h2}} + \frac{P_{h2} (1 - P_{h2})}{n_{h2} (2P_{h2} - 1)} = \frac{\lambda_{h2} (1 - \lambda_{h2})}{n_{h2} (2P_{h2} - 1)} \]  \hspace{1cm} (18)

Where; \( \lambda_{h2} = P_{h2} \pi_{h} + (1 - P_{h2})(1 - \pi_{h}) = (2P_{h2} - 1)\pi_{h} + (1 - P_{h1}) \)

The second response from \( R_{h2} \) have similar parameters; so that we have:
\[ \pi_{h22} = \frac{\hat{\pi}_{h22} - (1 - P_{h2}^*)}{(2P_{h2}^* - 1)} \]

and
\[ Var(\hat{\pi}_{h22}) = \frac{\pi_{h2} (1 - \pi_{h2})}{n_{h2}} + \frac{P_{h2}^* (1 - P_{h2}^*)}{n_{h2} (2P_{h2}^* - 1)} = \frac{\lambda_{h2}^* (1 - \lambda_{h2}^*)}{n_{h2} (2P_{h2}^* - 1)} \]

where; \( \lambda_{h2}^* = P_{h2}^* \pi_{h} + (1 - P_{h2}^*)(1 - \pi_{h}) = (2P_{h2}^* - 1)\pi_{h} + (1 - P_{h2}^*) \)

From Lanke (1976), to provide equal protection in \( R_{h1} \) and \( R_{h2} \) it can be shown that we must have either of the following:
\[ P_{h2} = \frac{1}{2 - P_{h1}} \]

or
\[ P_{h2}^* = \frac{1}{2 - P_{h1}} \]

With this restriction the variance of the estimators \( \hat{\pi}_{h12} \) and \( \hat{\pi}_{h22} \) become same. To estimate \( \pi_{h} \) from the information collected by the double use of \( R_{h2} \), we defined an unbiased estimator as follows:
\[ \hat{\pi}_{h} = \lambda_{1}\hat{\pi}_{h21} + \lambda_{2}\hat{\pi}_{h22} \]

where; \( \lambda_{1} \) and \( \lambda_{2} \) are the weights assuming value 0.5 when Var(\( \hat{\pi}_{h} \)) is optimized.

Thus the \( \hat{\pi}_{h} \) becomes:
\[ \hat{\pi}_{hP} = \frac{\hat{\pi}_{h21} + \hat{\pi}_{h22}}{2} \]  

(19)

Its variance is given by:

\[ \text{Var}(\hat{\pi}_{hP}) = \frac{\text{Var}(\hat{\pi}_{h21})}{2} = \frac{1}{2} \left[ \frac{\pi_h (1 - \pi_h)}{n_{h2}} + \frac{P_{h2} (1 - P_{h2})}{n_{h2} (2P_{h2}^2 - 1)} \right] \]  

(20)

Since, \[ \text{Var}(\hat{\pi}_{h21}) = \text{Var}(\hat{\pi}_{h22}) \]

and \[ P_{h2} = 1 - P_{h1}^* \]

An unbiased estimator in terms of all the information collected by both the random devices \( R_{h1} \) and \( R_{h2} \) in the \( h \)th stratum is defined as follows:

\[ \hat{\pi}_{hP(\text{tot})} = \pi_h = \frac{n_{h1}}{n_h} \hat{\pi}_{h1} + \frac{n_{h2}}{n_h} \hat{\pi}_{hP} \]  

(21)

As both the random devices \( R_{h1} \) and \( R_{h2} \) are independent, the variance of \( \hat{\pi}_{hP(\text{tot})} \) under the restriction by Lanke (1976):

\[ P_{h2} = \frac{1}{2 - P_{h1}} \]

Thus is given by:

\[ \text{Var}(\hat{\pi}_{hP(\text{tot})}) = \left( \frac{\lambda_h + 1}{2} \right) \frac{\pi_h (1 - \pi_h)}{n_h} + \frac{\lambda_h (1 - \pi_h)(1 - P_{h1})}{n_h P_{h1}} + \frac{(1 - \lambda_h)(1 - P_{h1})}{2n_h P_{h1}^2} \]  

(22)

where; \[ \lambda_h = \frac{n_{h1}}{n_h} \]

A stratified proportion estimator of the population proportion of the individuals with sensitive trait is defined as:

where; \[ \hat{\pi}_{\text{Sero}} = \sum_{i=1}^{L} W_h \hat{\pi}_{hP(\text{tot})} \]  

(23)

Its variance is given by:

\[ \text{Var}(\hat{\pi}_{\text{Sero}}) = \sum_{h=1}^{L} W_h^2 \left[ \pi_h (1 - \pi_h) \left( \frac{\lambda_h + 1}{2} \right) + \frac{\lambda_h (1 - \pi_h)(1 - P_{h1})}{P_{h1}^2} + \frac{(1 - \lambda_h)(1 - P_{h1})}{2P_{h1}^2} \right] \]  

(24)

Its variance under the optimum allocation of total sample size into different strata is given by:

\[ \text{Var}(\hat{\pi}_{\text{Sero}}) = \frac{1}{n} \left( \sum_{h=1}^{L} W_h \right) \left[ \left( \frac{\lambda_h + 1}{2} \right) \pi_h (1 - \pi_h) + \frac{\lambda_h (1 - P_{h1})(1 - \pi_h)}{P_{h1}^2} + \frac{(1 - P_{h1})(1 - \lambda_h)}{2P_{h1}^2} \right]^2 \]  

(25)
2.2 Relative Efficiency of the Proposed HIV Seroprevalence Model

One of the most important ways of assessing any sample survey model is through its efficiency relative to the existing models. Again, there is the need to compare the relative efficiency of the proposed two-way RR Model in stratification for HIV seroprevalence tracking with Kim and Warde (2005) stratified estimator. We deduce that the proposed two-way RR Model in stratification for HIV seroprevalence tracking is more efficient for a fixed sample size if and only if:

$$\frac{\text{Var}(\hat{\pi}_{\text{uk}}) - \text{Var}(\hat{\pi}_{\text{sero}})}{\text{Var}(\hat{\pi}_{\text{sero}})} \geq 0$$

(26)

The above inequality will be true if for each stratum $h$, $h=1,2,...,L$ we have the following:

$$\pi_h(1-\pi_h) + \frac{(1-P_{h1})(1-P_{1h})(1-\pi_h) + 1 - \lambda_h}{P_{h1}^2}$$

$$- \left( \frac{\lambda_h + 1}{2} \right) \pi_h(1-\pi_h) - \frac{\lambda_h(1-P_{h1})(1-\pi_h)}{P_{h1}^2} - \frac{(1-P_{h1})(1-\lambda_h)}{2P_{h1}^2} \geq 0$$

Then; $\pi_h(1-\pi_h)(1-\lambda_h) + \frac{(1-P_{h1})(1-\lambda_h)}{P_{h1}^2} \geq 0$ (27)

or $\pi_h(1-\pi_h) + \frac{(1-P_{h1})}{P_{h1}^2} \geq 0$ (28)

The inequality (3.12.19) is always true for every value of $\pi_h$, $P_{h1}$ and $\lambda_h$. Hence the proposed two-way RR Model in stratification for HIV seroprevalence tracking is also more efficient than Kim and Warde (2005) stratified estimator.

3. Results

An unbiased two-way RRM in stratification for HIV seroprevalence rates estimator is given by:
\[ \hat{\pi}_{Sero} = \sum_{i=1}^{L} W_h \hat{\pi}_{P(i\text{tot})} \] where; \( W_h = \frac{N_h}{N} \) for \( h = 1,2,\ldots,L \)

Its variance is given by:

\[ \text{Var}(\hat{\pi}_{Sero}) = \sum_{h=1}^{L} \frac{W_h^2}{n_h} \left[ \pi_h (1-\pi_h) \left( \frac{\lambda_h + 1}{2} \right) + \frac{\lambda_h (1-\pi_h)(1-P_{h1})}{P_{h1}} + \frac{(1-\lambda_h)(1-P_{h1})}{2P_{h1}^2} \right] \]

The computations for the model to estimate HIV seroprevalence rate give the following results:

\[ \phi = \pi_h (1-\pi_h) \left( \frac{\lambda_h + 1}{2} \right) + \frac{\lambda_h (1-\pi_h)(1-P_{h1})}{P_{h1}} + \frac{(1-\lambda_h)(1-P_{h1})}{2P_{h1}^2} \]

\[ \hat{\pi}_{Sero} = \sum_{i=1}^{L} W_h \hat{\pi}_{P(i\text{tot})} = 0.0612 \]

\[ \text{Var}(\hat{\pi}_{Sero}) = 0.000067 \]

\[ \text{SE}(\hat{\pi}_{Sero}) = \sqrt{\text{Var}(\hat{\pi}_{Sero})} = 0.0082 \]

The 95% confidence interval for HIV seroprevalence rate using the two-way RR Model in stratification is given by:

\[ (\hat{\pi}_{Sero}) \pm 1.96 \times \text{SE}(\hat{\pi}_{Sero}) = 0.0612 \pm 1.96 \times 0.0082 = [0.045, 0.077] \]

4. Conclusion

The research herein has dual advantages, modelling and applications. This study was motivated by the fact that conventional data collection techniques usually cause evasive or untruthful responses when people are asked sensitive questions like their HIV serostatus. As a result, it is difficult to make accurate inferences from such unreliable data. Hence a two-way RR Model in stratification was devised using the work of Warner (1965), Arnab (2004), Quatember (2009), among others particularly for HIV seroprevalence surveys. The model was proved to be more efficient than a similar model by Kim and Warde (2005).

Furthermore, the model was used to estimate HIV seroprevalence rate in a small adult population using a sample size of 550 and a design parameter of 0.7. Table 1 describes the strata sizes, the sample sizes, the number of ‘yes’ responses and the strata weights for the three strata. Table 2 gives the proportion of ‘yes’ responses for both random devices 1 and 2 and the estimates of seroprevalence rates for the three strata. Furthermore, Table 3 represents the worksheet for computing the variances of the seroprevalence rate. Table 4 is the summary
depicting the overall HIV seroprevalence rate, its variance and the 95% confidence interval for the estimate.

The result shows that, using the survey data, the model estimated the HIV seroprevalence rate as 6.1% with a standard error of 0.0082 and 95% confidence bands of [4.5%, 7.7%]. These estimates are for adults who are 18 years and above who attend a hospital. These results are consistent with that of Nigerian sentinel survey (2003) conducted by NACA, USAID and CDC which estimated the HIV seroprevalence in Kaduna State as 6.0%. Hence, the RRTs herein can serve as new viable methods for HIV seroprevalence surveys.

References


## Appendix A: Tables

### Table 1: Samples and Strata Sizes

<table>
<thead>
<tr>
<th>Strata</th>
<th>Strata Description</th>
<th>$N_h$</th>
<th>$n_h$</th>
<th>$n_{h1}$</th>
<th>$n_{h21}$</th>
<th>$n_{h22}$</th>
<th>$W_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Married (Men/ Women)</td>
<td>1,285</td>
<td>189</td>
<td>32</td>
<td>42</td>
<td>38</td>
<td>0.344</td>
</tr>
<tr>
<td>2</td>
<td>Unmarried (Men/ Women)</td>
<td>2,020</td>
<td>297</td>
<td>56</td>
<td>55</td>
<td>63</td>
<td>0.540</td>
</tr>
<tr>
<td>3</td>
<td>Divorced/Separated/Widowed</td>
<td>435</td>
<td>64</td>
<td>12</td>
<td>10</td>
<td>11</td>
<td>0.116</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>3,740</td>
<td>550</td>
<td>100</td>
<td>107</td>
<td>112</td>
<td>1.000</td>
</tr>
</tbody>
</table>

### Table 2: Summary of Results of the Random Devices

<table>
<thead>
<tr>
<th>Strata</th>
<th>$\hat{\lambda}_{h1}$</th>
<th>$\hat{\pi}_{h1}$</th>
<th>$V(\hat{\pi}_{h1})$</th>
<th>$\hat{\lambda}_{h21}$</th>
<th>$\hat{\pi}_{h21}$</th>
<th>$V(\hat{\pi}_{h21})$</th>
<th>$\hat{\lambda}_{h22}$</th>
<th>$\hat{\pi}_{h22}$</th>
<th>$V(\hat{\pi}_{h22})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.376</td>
<td>0.109</td>
<td>0.0150</td>
<td>0.402</td>
<td>0.255</td>
<td>0.0358</td>
<td>0.365</td>
<td>0.163</td>
<td>0.0381</td>
</tr>
<tr>
<td>2</td>
<td>0.350</td>
<td>0.071</td>
<td>0.0083</td>
<td>0.401</td>
<td>0.253</td>
<td>0.0273</td>
<td>0.460</td>
<td>0.256</td>
<td>0.0246</td>
</tr>
<tr>
<td>3</td>
<td>0.343</td>
<td>0.061</td>
<td>0.0383</td>
<td>0.345</td>
<td>0.113</td>
<td>0.1412</td>
<td>0.379</td>
<td>0.198</td>
<td>0.0902</td>
</tr>
</tbody>
</table>

### Table 3: Summary of Computations

<table>
<thead>
<tr>
<th>Strata</th>
<th>$n_{h1}$</th>
<th>$n_h$</th>
<th>$\hat{\pi}_{h1}$</th>
<th>$n_{h2}$</th>
<th>$n_h$</th>
<th>$\hat{\pi}_h$</th>
<th>$W_h^2$</th>
<th>$\hat{\pi}_h$</th>
<th>$\hat{\pi}_h(1 - \hat{\pi}_h)$</th>
<th>$\phi$</th>
<th>$\sum_{h=1}^{L} W_h^2 \phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.169</td>
<td>0.209</td>
<td>0.201</td>
<td>0.060</td>
<td>0.0206</td>
<td>0.00063</td>
<td>0.056</td>
<td>0.037</td>
<td>0.000023</td>
<td>0.037</td>
<td>0.000023</td>
</tr>
<tr>
<td>2</td>
<td>0.189</td>
<td>0.255</td>
<td>0.212</td>
<td>0.067</td>
<td>0.0362</td>
<td>0.00098</td>
<td>0.063</td>
<td>0.041</td>
<td>0.000040</td>
<td>0.041</td>
<td>0.000040</td>
</tr>
<tr>
<td>3</td>
<td>0.188</td>
<td>0.156</td>
<td>0.172</td>
<td>0.038</td>
<td>0.0044</td>
<td>0.00021</td>
<td>0.037</td>
<td>0.019</td>
<td>0.000004</td>
<td>0.019</td>
<td>0.000004</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0612</td>
<td></td>
<td>0.000067</td>
</tr>
</tbody>
</table>

### Table 4: Summary of Seroprevalence Results

<table>
<thead>
<tr>
<th>$N$</th>
<th>$n$</th>
<th>$\hat{\pi}_{\text{Sero}}$</th>
<th>$\text{Var}(\hat{\pi}_{\text{Sero}})$</th>
<th>$95%$ confidence interval</th>
<th>Lower limit</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,740</td>
<td>550</td>
<td>0.0610</td>
<td>0.000067</td>
<td>0.045</td>
<td>0.077</td>
<td></td>
</tr>
</tbody>
</table>
This academic article was published by The International Institute for Science, Technology and Education (IISTE). The IISTE is a pioneer in the Open Access Publishing service based in the U.S. and Europe. The aim of the institute is Accelerating Global Knowledge Sharing.

More information about the publisher can be found in the IISTE’s homepage: http://www.iiste.org

The IISTE is currently hosting more than 30 peer-reviewed academic journals and collaborating with academic institutions around the world. Prospective authors of IISTE journals can find the submission instruction on the following page: http://www.iiste.org/Journals/

The IISTE editorial team promises to the review and publish all the qualified submissions in a fast manner. All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Printed version of the journals is also available upon request of readers and authors.

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digital Library, NewJour, Google Scholar