A Survey of Tax Evasion Using the Randomized Response Technique

Jodie Houston and Alfred Tran

The Australian National University

Corresponding author:

Alfred Tran
School of Business and Information Management
The Australian National University
Canberra ACT 0200
AUSTRALIA
E-mail: Alfred.Tran@anu.edu.au
Phone: +61 2 6125 5118
Fax: +61 2 6125 5005

A Survey of Tax Evasion Using the Randomized Response Technique

ABSTRACT

We conducted a mail questionnaire survey using both the randomized response (RR) technique and the direct questioning (DQ) technique to directly estimate the prevalence and type of income tax evasion. We also assessed the effectiveness of the RR technique in reducing response and non-response biases and examined the relationship between tax evasion and key demographic variables.

Of the respondents completing the RR survey instrument, 5.5% admitted tax evasion by under-reporting income, and 6.5% admitted tax evasion by over-claiming deductions. The corresponding proportions obtained from the DQ survey instrument were 1.7% and 4.2% respectively. The RR technique was ineffective in reducing non-response bias, but the estimated proportions of tax evasion obtained by the RR technique are higher than those obtained by the DQ technique. A relationship was found between the demographic variables examined and tax evasion. However, interpretation of the results was restricted by the lack of statistical significance of the differences.

INTRODUCTION

Despite interest in tax evasion, very little research has been carried out in Australia and not much is known about the extent of the problem. Empirical investigation into tax evasion is needed to understand the extent and the cause of the problem, but current research is hampered by difficulties in obtaining sensitive information about tax evasion from individuals. One of the main limitations facing researchers investigating tax evasion is the inability to directly observe individual evasion behavior. As such, most empirical evidence is based on individuals' self-reports (*i.e.,* surveys) to describe evasion behavior.[1]

Surveys of tax evasion are complicated by the sensitive nature of the topic. In general, tax evasion is perceived to be an illegal and socially undesirable behavior. Individuals are reluctant to admit to having evaded tax. The threat of penalties, prosecution and stigmatization can induce individuals either to lie about their tax evasion behavior (response bias), or to refuse to take part in the study because they wish to avoid answering sensitive questions (non-response bias).[2] Response and non-response biases in a survey affect the validity and the generalizability of the results, making reliable estimates of tax evasion difficult to obtain.

The problem facing researchers is how to encourage participants to respond, and then to provide truthful response in surveys. A suggested solution is the Randomized Response (RR) technique first developed by Stanley Warner (1965). The RR technique was designed to reduce both response bias and non-response bias in surveys which ask sensitive questions. It uses probability theory to protect the privacy of an individual's response and has been used successfully in several sensitive research areas, such as abortion, drugs and assault.

There have been calls from the tax research community to use the RR technique to investigate tax evasion (see, for example, Elffers *et al* 1988; Roth *et al* 1989; Harwood *et al* 1993), but a review of literature shows that little has been done. Motivated by the need to gather more reliable and meaningful data on tax evasion and to improve the research methods, we conducted a survey designed to achieve the following objectives:

1. to directly estimate the proportion and type of tax evasion (*i.e.,* under-reporting income, over-claiming deductions, and over-claiming rebates and credits) of individuals in Australia;
2. to assess the effectiveness of the RR technique in reducing response and non-response biases in surveys asking sensitive questions; and
3. to examine the relationship between tax evasion and key demographic variables.

The rest of this article is organized as follows. The next section briefly reviews the development of the RR technique and its applications in tax research. This is followed by a description of the hypotheses developed and tested in the study. Next, the research design is described, and then the results are reported. The article concludes with a discussion of the limitations of the study, and some improvements in designing surveys using the RR technique are suggested.

LITERATURE REVIEW

The Warner Model

The RR technique was proposed originally by Warner (1965). The innovative approach was designed to protect the privacy of survey respondents when they were asked sensitive questions. In the Warner design, the respondents are given two logically opposite questions and are instructed to answer one or the other depending on the outcome of a randomizing device. For example, suppose the sensitive characteristic is tax evasion. The respondent may be asked to toss a dice, and the outcome determines which question they answer:

1, 2, 3 or 4:	Question 1: I have evaded tax.
5 or 6:	Question 2: I have never evaded tax.
Answer:	True or False

When the respondent answers 'true' or 'false,' the researcher does not know whether the respondent is answering Question 1 or Question 2. Thus the privacy of the respondent is protected. The use of probability theory allows the researcher to estimate the proportion of affirmative responses to Question 1 (π) and the associated sampling variance using the following equations:

$$P(\text{True}) = P(\text{Question 1})\, P(\text{True} \mid \text{Question 1}) + P(\text{Question 2})\, P(\text{True} \mid \text{Question 2}) \quad (1)$$

or
$$\lambda = p\pi + (1 - p)(1 - \pi) \quad (2)$$

Thus,
$$\boxed{x} = (\lambda + p - 1) / (2p - 1) \qquad (p \neq .5) \quad (3)$$

and
$$\text{Var}(\boxed{x}) = [\pi(1 - \pi) / n] + [p(1 - p) / n(2p - 1)^2] \quad (4)$$

where: $\boxed{x}$ = the estimated proportion of 'true' responses to Question 1;

$\hat{\lambda}$ = the observed proportion of 'true' responses;

p = the probability of answering Question 1; and

n = the sample size.

Warner (1965) claimed that the RR technique had the potential to reduce both response bias and non-response bias resulting from sensitive survey questions. The main drawback was the increase in variance of the estimator due to the introduction of the randomizing procedure into the design. Because of this inflated variance, Warner stressed the importance of using the RR technique only for sensitive issues, so as to offset the increased variance of the estimate with the lower mean square error produced by more truthful reporting.

The Unrelated Question Design

An important improvement to the Warner model was proposed by Horvitz *et al* (1967) who suggested the use of an unrelated question with a non-sensitive characteristic.[3] For example:

Question 1: Did you cheat on your tax return last year?

Question 2: Did you watch the 6:00 pm news yesterday?

This unrelated question approach requires two independent samples with different selection probability ($p_1 \neq p_2$) to estimate two parameters: π_x for the sensitive behavior, and π_y for the non-sensitive behavior. It has the improvement of reducing the sensitivity of the design, as only one of the questions relates to the sensitive topic. However, the samples also are used to estimate the distribution of the unrelated question (*i.e.,* watching news) which may not be of interest to the researcher. This technique is referred to as the *unrelated question design with*

an unknown distribution.[4] The design was expected to further reduce response bias and improve the efficiency of the estimate. The estimated proportion of affirmative responses to the sensitive question and the associated variance are calculated using the following equations:

$$\hat{x}_x = [\hat{\lambda}_1(1 - p_2) - \hat{\lambda}_2(1 - p_1)] / (p_1 - p_2) \quad (5)$$

$$\text{Var}(\hat{x}_x) = [1 / (p_1 - p_2)] [\lambda_1(1 - \lambda_1)(1 - p_2)^2 / n_1 + \lambda_2(1 - \lambda_2)(1 - p_1)^2 / n_2] \quad (6)$$

where: $\hat{x}_x$ = the estimated proportion of 'yes' responses to the sensitive question;

$\hat{\lambda}_1, \hat{\lambda}_2$ = the proportion of 'yes' responses for samples 1 and 2 respectively;

p_1, p_2 = the probability of answering the sensitive question for samples 1 and 2 respectively; and

n_1, n_2 = the size of samples 1 and 2 respectively.

The unrelated question design was further improved by Greenberg *et al* (1969). They examined a similar design where the distribution of the non-sensitive question is known in advance. Knowing the distribution of the non-sensitive question offers a substantial improvement in the precision of the estimate of the sensitive characteristic and reduces the number of samples to one, as there is now only one parameter to estimate. The *unrelated question design with a known distribution* uses a simplified version of the original unrelated question equations to estimate the proportion of the sensitive characteristic and the sampling variance:

$$\hat{x}_x = [\hat{\lambda} - (1 - p)\,\pi_y] / p \quad (7)$$

$$\text{Var}(\hat{x}_x) = \lambda(1 - \lambda) / np^2 \quad (8)$$

where: $\hat{x}_x$ = the estimated proportion of 'yes' responses to the sensitive question;

π_y = the known proportion of 'yes' responses to the non-sensitive question;

$\hat{\lambda}$ = the observed proportion of 'yes' responses;

p = the probability of answering the sensitive question; and

n = the sample size.

Although the concept of the unrelated question design was introduced by Horvitz *et al* (1967),[5] Greenberg *et al* (1969) offered a more comprehensive treatment of the refined RR technique and a theoretical proof that it provided a significant improvement on the Warner design by increasing the precision and efficiency of the sensitive estimate, especially when the distribution of the non-sensitive question is known.

Further extensions of the RR technique include the use of polychotomous measures and quantitative measures.[6] Despite the variety of extensions and variations that have emerged since Warner's original design, the unrelated question design developed by Greenberg *et al* (1969)[7] has remained one of the most popular RR techniques used by researchers investigating sensitive issues.

The RR technique has been used in a few studies in the tax area. They are briefly reviewed below.

The Use of RR Technique in Tax Research

An early taxpayer compliance study employing the RR technique was carried out by Aitken and Bonneville (1980)[8] who compared the RR technique to a locked-box. The results appeared to be promising: the admission of tax cheating was higher for the RR technique

sample than the locked-box sample. These results were later compared to a study done by Yankelovich *et al* (1984) who did not use the RR technique. Aitken and Bonneville (1980) found a significantly greater number of respondents admitting to cheating on their taxes than in Yankelovich *et al* (1984). However, little reliance could be placed on the comparison as there were numerous differences between the two studies, such as different surveying techniques, questions and time frames, and a lack of statistical testing.

Despite the encouragement offered by Aitken and Bonneville (1980), the randomized response technique did not appear to have been widely used in tax evasion research. In 1993, Harwood *et al* (1993) published a research note on the potential benefits of using RR technique in tax compliance research. Their study did not explicitly deal with taxpayer compliance but examined the relationship between the income level of paid tax preparers[9] and non-compliance. They briefly reviewed different RR designs and their application in other areas to illustrate the use of the RR technique for sensitive issues. Harwood *et al* (1993) offered an adequate description and overview to draw the tax research community's attention to the potential usefulness of the RR technique.

The RR technique was used again in the tax area by Larkins *et al* (1997), who surveyed tax practitioners. They investigated the ethical issues involved in the tax practice, and more importantly, compared the effectiveness of the RR technique to direct questioning (DQ). The results indicated that the RR technique did not reduce response bias and non-response bias compared to DQ. Larkins *et al* (1997) concluded that the RR technique was unsuitable for surveying tax practitioners about their ethical behavior because the existence of an affinity

between CPAs might have resulted in a desire to protect the reputation of the profession. A more likely reason might be that the questions asked, which addressed the compliance of CPAs with the Statements of Responsibilities in Tax Practice, might not have been sufficiently sensitive to warrant the use of the RR technique as the Statement only provided advisory guidelines. Despite the conclusion, Larkins *et al* (1997) urged the use of RR techniques in tax research, and emphasized the importance of comparing and validating the RR technique in other tax settings.

Overall there has been very little work done using the RR technique in the tax area, with only one study directly addressing the issue of taxpayer compliance (Aitken and Bonneville 1980). This lack of research reinforces the importance of first establishing the validity of the RR technique in tax research and what improvement, if any, it can add to the traditional DQ technique.

HYPOTHESES DEVELOPMENT

We conducted a mail questionnaire survey of Australian individual taxpayers using both RR and DQ techniques to directly estimate the proportion and the types of income tax evasion (under-reporting income, over-claiming deductions, and over-claiming rebates and credits) for the 1997/98 tax year.

We designed two survey instruments: one used the RR technique to ask sensitive questions, and the other used the traditional DQ technique. We tested the following two hypotheses to assess whether the RR technique was effective in reducing non-response and response biases in surveys asking sensitive questions:

H_1: The response rate will be higher for individuals receiving the RR survey instrument than for those receiving the DQ survey instrument.

H_2: The proportion of individuals admitting to tax evasion will be higher for those completing the RR survey instrument than for those completing the DQ survey instrument.

The first hypothesis tests whether the RR technique reduces non-response bias in the survey. The second hypothesis tests whether response bias is reduced by use of the RR technique and is based on the assumption that a higher proportion of respondents admitting evasion indicates more truthful reporting. [10]

Six more hypotheses also have been developed to investigate the relationship between tax evasion and six taxpayer demographic variables. Table 1 presents a summary of the empirical results from previous studies which examined the relationship between taxpayer demographic variables and tax evasion.

TABLE 1 ABOUT HERE

Care should be taken in interpreting the empirical findings summarized in Table 1 because there are factors which could restrict the comparability of results between studies and contribute to the observed inconsistency, such as those described below.

1. Different research methods have been employed to collect tax evasion data and this may account in part for the observed inconsistency in results between the studies.[11]
2. Different target populations have been used; for example, convenience samples of students, telephone listings, tax agents and other specific subsets of the population.
3. Not all the studies use the same definition of tax evasion, many use the term 'non-compliance' which is a broader term, encompassing both intentional and unintentional behavior.
4. The main dependent variable in all these studies is tax evasion (or non-compliance) but only a few have directly estimated tax evasion. Many use hypothetical scenarios or taxpayers' attitudes and opinions, then adopt the assumption that intentional behavior is the same as actual behavior.[12]
5. There are differences between studies in the number and type of variables that are associated with tax evasion. The absence of relevant variables can produce a confounding effect if they are not properly controlled.

Guided by the results of previous studies, we hypothesized the following relationships between the demographic attributes of taxpayers and their propensity to evade tax.

> **H_3:** Taxpayers in a lower income bracket will have a higher proportion of evasion than taxpayers in a higher income bracket.

H_4: The proportion of evasion will be higher for younger taxpayers than for older taxpayers.

H_5: Taxpayers with a higher level of education will exhibit a higher proportion of evasion than taxpayers with a lower level of education.

H_6: Self-employed taxpayers will exhibit a higher proportion of tax evasion than employees.

H_7: Men will display a higher proportion of evasion than women.

The tax preparer (or tax agent in Australian terminology) variable has recently been included in tax evasion studies.[13] Relevant findings from the tax preparer literature suggest that many tax preparers display aggressive behavior against the tax authorities, aiding taxpayers in avoidance and non-compliance. We therefore hypothesized that:

H_8: Taxpayers using a tax agent will display a higher proportion of evasion than taxpayers who do not use a tax agent.

RESEARCH DESIGN

We designed two survey instruments: one instrument used the RR technique to ask the sensitive questions about tax evasion (a copy is included as an appendix), while the other used the traditional DQ technique. Otherwise, the two instruments were identical.[14] Both survey instruments offered respondents the protection of anonymity. Respondents of the RR

instrument had the added protection of the randomizing procedure. We used the unrelated question design with a known distribution in the RR instrument. We made a few decisions to select the most suitable RR design.

Randomizing Procedure

The randomizing procedure is vitally important to the success of the RR technique, as it shows respondents that their answers are being protected by probability theory and that the researcher cannot know which question they have answered. The most common randomizing devices used in mail RR surveys are bank notes supplied by the respondent, and random number charts supplied by the researcher. The use of bank notes has been popular in accounting and tax research (see Berry *et al* 1987; Harwood *et al* 1993; Larkins *et al* 1997) as the serial number on a bank note is considered sufficiently random for the purposes of the RR technique and easily accessible to most individuals. We used the last three digits of the serial number on a bank note of the respondent. The only problem was the potential lack of a bank note,[15] so respondents were instructed to use the last three digits of their telephone number (which are sufficiently random) if a bank note was not handy.

An important parameter to determine is the probability of answering the sensitive question, p, which has an impact on the variance of the estimate. The smaller the level of p (*i.e.,* the fewer respondents who are instructed by the randomizing device to respond to the sensitive question), the greater the protection offered to the respondents. However, this also means that the sampling variance of the estimator (refer to equation (8)) will increase. Thus, there is a trade-off between respondent jeopardy and estimation efficiency. For example, other things

being equal, using a probability of p = 0.5 will increase the variance by a factor of 4, whereas using a probability of p = 0.7 will only increase the variance by a factor of 2. This means that a RR design using a probability of 0.7 would require a sample only about half as large as one using a probability of 0.5 to achieve the same level of efficiency. Because of this, researchers using the RR technique are encouraged to make p as large as they dare (see Lanke 1975), especially for small samples where the level of sampling error is higher. Soecken and Macready (1982) recommended that p be chosen between 0.7 and 0.85 to obtain sufficient efficiency in the design and still protect the privacy of respondents. We struck a balance between respondent jeopardy and estimation efficiency and chose p = 0.7 in the study.

Non-sensitive Question and Known Distribution

In choosing a non-sensitive question and a known distribution, previous studies have used known demographic distributions for certain populations,[16] or have asked respondents if they were born in a certain month,[17] but these measures can be unreliable, with problems of memory recall, respondent knowledge and the validity of the demographic statistics used. Because of these limitations, we followed the practice of several more recent studies (Berry *et al* 1987; Larkins *et al* 1997) and used the serial number on a bank note to create a known distribution.

The choice of the known distribution will have an impact on the variance of the sensitive estimate. The probability of getting either type of 'yes' response is given by:

$$\lambda = p\pi_x + (1 - p)\,\pi_y \qquad (9)$$

where: π_x = the true proportion of respondents with the sensitive behavior; and

π_y = the proportion of 'yes' response to the non-sensitive question.

The observed proportion of 'yes' responses (λ) increases as the known distribution (π_y) increases. This leads to a larger numerator in equation (8), and results in a higher variance and a less efficient estimate of the prevalence of the sensitive behavior. Setting π_y to zero is in effect direct questioning, as any 'yes' response obtained only could refer to the sensitive question. A smaller π_y leads to a smaller variance of the sensitive estimate. On the other hand, a larger π_y provides more protection to respondents as there is a greater likelihood of more respondents answering 'yes.' Again, a trade-off exists between respondent protection and estimation efficiency.

In the RR survey instrument, we asked the respondents to use a digit in the bank note's serial number as a randomizing device, directing them to answer either the sensitive or non-sensitive question. The non-sensitive question also uses the same digit to create a known distribution of answering 'yes,' which has a probability of 1/3 or 33.3%. Ideally, the chosen known distribution for the non-sensitive question should be as close as possible to the sensitive attribute being estimated. With the benefits of hindsight, we found that the known probability for the non-sensitive question we used was too high. The large π_y we used has substantially inflated the sampling variance of the estimator for the sensitive attribute.

Survey Procedure

A mail questionnaire survey was used because the use of RR technique required larger samples for effective data analysis. We also wanted to survey across all of Australia in order to obtain a representative sample of Australian individuals.

The target population for this study was members of the Australian public who derive income and are subject to income tax. Two random samples were drawn using *Australia on Disc* (May 1999 version) which is a CD-ROM containing all the latest residential telephone directories across Australia and includes a software to draw random samples from the database. The sample size was 500 for the DQ survey instrument, and 1,500 for the RR survey instrument. The larger sample for the RR instrument was meant to compensate for the inflated sampling variances caused by the randomizing procedure. We used a screening question at the beginning of the survey instruments to make sure the respondents were taxpayers who had lodged a tax return for the 1997/98 tax year.

An advance letter was mailed prior to the survey instruments to explain the purpose of the study and the procedure, inform respondents of the survey and encourage participation. Ten days after the survey instruments were dispatched, a follow-up letter was posted to thank those who had responded and to remind those who had not to complete and return the questionnaire using the reply paid envelope.

Statistical Procedures

Z-tests were used in hypotheses testing. All comparisons involving RR data used the estimated proportion of evasion and the sampling variance based on equations (7) and (8) to calculate the z-score, using the standard formula:

$$z = (\boxed{x}_1 - \boxed{x}_2) / [\mathrm{Var}(\boxed{x}_1) + \mathrm{Var}(\boxed{x}_2)]^{1/2} \quad (10)$$

where: $\boxed{x}$ = estimated proportion of the respondents admitting tax evasion; and

$\mathrm{Var}(\boxed{x})$ = variance of the estimated proportion.

To calculate the proportion of respondents admitting to two types of evasion (under-reporting income and over-claiming deductions), the joint distribution of two evasion types was estimated using the following formula:[18]

$$\pi_{x1x2} = [\lambda - p_1(1 - p_2)\pi_{x1}\pi_{y2} - (1 - p_1)p_2\pi_{y1}\pi_{x2} - (1 - p_1)(1 - p_2)\pi_{y1}\pi_{y2}] / (p_1p_2) \quad (11)$$

where: π_{x1x2} = the joint probability of the two sensitive characteristics (evasion types);

λ = the proportion of respondents answering 'yes' to both questions;

p_1, p_2 = the probability of answering two sensitive questions 1 and 2 respectively;

π_x = the estimated proportion of 'yes' responses to the sensitive question 1 or 2; and

π_y = the known proportion of 'yes' responses for the non-sensitive question 1 or 2.

RESULTS

Non-Response Bias

The effectiveness of the RR technique in reducing non-response bias was tested in H_1. The results are summarized in Table 2.

TABLE 2 ABOUT HERE

There are statistically significant differences (at the 0.05 level) between the *gross* response rate of the RR instrument (27.8%) and that of the DQ instrument (34.6%), and between the *useable* response rates of the RR (22.3%) and DQ (27.9%) instruments. The null hypothesis that the two instruments had the same response rates is rejected, but the direction of rejection is opposite to the one hypothesized in H_1, with the DQ instrument having a significantly higher response rate than the RR instrument. Thus, the RR technique has failed to reduce non-response bias.

The following explanations could account for this result:

1. The survey instrument using the RR technique was much more time consuming to complete and involved carrying out relatively complex instructions and procedures to complete it.
2. Some individuals receiving the RR instrument did not believe that the survey results could be useful because of the randomizing procedure. This is evidenced by one letter and a few telephone calls in which the recipients expressed this concern to us.
3. A randomizing device was not handy so the recipients did not bother to complete the survey.[19]
4. There might be a general distrust about the RR technique, as it is a relatively unknown technique and respondents may have felt their privacy was not protected sufficiently.

Response Bias

The second hypothesis examined the effectiveness of the RR technique to reduce response bias. The results summarized in Table 3 show that the estimated proportions of admitted evasion in the survey using the RR technique are higher than those in the survey using the DQ technique for the evasion types 'under-reporting income' (RR 5.5%; DQ 1.7%) and 'over-claiming deductions' (RR 6.5%; DQ 4.2%). However, the differences between the two techniques are not statistically significant at the 0.05 level. The higher estimated prevalence of tax evasion obtained using the RR technique may suggest that the use of the RR technique has reduced response bias. On the other hand, the higher estimated proportions also may be attributed to the randomizing procedure and the responses to the non-sensitive questions. The estimated proportion of taxpayers admitting 'over-claiming tax rebates and tax credits' is zero for both RR[20] and DQ instruments. Therefore, H_2 is rejected. The effectiveness of the RR technique in reducing response bias cannot be established statistically.

TABLE 3 ABOUT HERE

Since the proportions of admitted tax evasion in the survey using the DQ technique are very low (*e.g.,* 1.7%, or 2 out of 121 valid responses, admitted evasion by under-reporting income), testing the association between demographic variables and tax evasion has little meaning. As such, the results of hypotheses testing reported below are related to the RR survey only.

Income Level

Consistent with the hypothesized direction in H_3, the results in Table 4 indicate that respondents in the lower income group tended to have a higher proportion of tax evasion by under-reporting income and by over-claiming deductions (6.9% and 9.3% respectively), compared to respondents in the higher income group (3.6% and 2.2% respectively). However, the differences are not statistically significant and H_3 is rejected.

TABLE 4 ABOUT HERE

Age

H_4 examines the relationship between age and tax evasion. The results in Table 5 confirm that the younger age group was less compliant than the older age group, displaying a higher proportion of evasion by under-reporting income (6.3% versus 5.0%) and by over-claiming deductions (10.2% versus 4.1%), but none of the differences proved to be statistically significant. Thus, H_4 is rejected.

TABLE 5 ABOUT HERE

Education Level

H_5 tests whether taxpayers with a higher level of education exhibit a higher proportion of evasion than taxpayers with a lower level of education. Contrary to the hypothesized direction, a negative relationship between education and tax evasion was found. Taxpayers without tertiary education tended to have higher proportions of tax evasion (6.4% by under-

reporting income and 9.4% by over-claiming deductions) than taxpayer with tertiary education (4.8% by under-reporting income and 3.8% by over-claiming deductions), but none of the differences are statistically significant, so H_5 is rejected.

TABLE 6 ABOUT HERE

Occupation

In H_6, we hypothesized that self-employed taxpayers would exhibit a higher proportion of tax evasion than employees. A statistically significant difference was found between the self-employed and employee groups. As shown in Table 7, self-employed respondents admitted a significantly higher proportion of tax evasion by under-reporting income (16.5%) than respondents who were employees (3.6%). H_6 is supported in the hypothesized direction. The difference in the proportions of over-claiming deductions also is in the hypothesized direction (9.2% for self-employed versus 7.5% for employee) but is not statistically significant.

TABLE 7 ABOUT HERE

Gender

H_7 hypothesizes that men would display a higher proportion of evasion than women. The results in Table 8 indicate a higher proportion of evasion occurred among women (8.3% under-reporting income, and 11.2% over-claiming deductions) than men (4.0% under-reporting

income, and 3.9% over-claiming deductions), pointing to a different direction from the hypothesized one. However, the differences are not statistically significant. H_7 is rejected.

TABLE 8 ABOUT HERE

The Influence of Tax Agents

The final variable examined was the influence tax agents had on individuals' evasion behavior. The results presented in Table 9 confirm that higher proportions of evasion are present among taxpayers who used the services of a tax agent to prepare their 1997/98 tax returns (6.2% under-reporting income and 7.4% over-claiming deductions) compared to those who did not use tax agents (0.8% under-reporting income and 0.8% over-claiming deductions). This is consistent with the hypothesized direction of H_8. However, the differences are not statistically significant, so H_8 is rejected.

TABLE 9 ABOUT HERE

SUMMARY AND CONCLUSION

Of the respondents completing the RR survey instrument, 5.5% admitted tax evasion by under-reporting income and 6.5% admitted evasion by over-claiming deductions. The corresponding proportions obtained from the DQ survey instrument were 1.7% and 4.2%, respectively. The RR technique was ineffective in reducing non-response bias, and its effectiveness in reducing response bias could not be established statistically.

Some relationships were found between the demographic variables examined and tax evasion. The association of the demographic variables and tax evasion confirmed some new trends found in tax evasion research, such as women evading more than men and the relevance of the tax preparer variable. However, interpretation of the survey results was restricted by the lack of statistical significance of the differences in hypotheses testing.[21] The only statistically significant result was for the occupation variable, where a significantly higher proportion of tax evasion by under-reporting income was found among respondents who were self-employed compared to respondents who were employees.

There are several ways to improve the efficiency of the design and hence the significance of the results. The first consideration is the type of RR technique chosen, as some RR techniques are more efficient than others. The RR technique used in this study (unrelated question with a known distribution) is one of the most efficient RR designs, as it uses a known distribution for the non-sensitive question, leaving only the sensitive attribute to be estimated.[22]

The second consideration is the choice of parameters used in operationalizing the RR technique since this can affect the efficiency of the estimates, particularly the choice of p (the probability of answering the sensitive question) and π_y (the known distribution for the non-sensitive question). The choice of parameters and their relationship to the sampling variance of the estimator are discussed in the research design section. The general rule is that the closer p is to 1 and π_y is to zero, the greater the efficiency of the design.[23] However, the jeopardy to the respondents also must be taken into consideration. The jeopardy level of the RR design

cannot be ignored for the sake of efficiency, as this would be a refutation of whole reasoning behind using a RR technique. The tradeoff between efficiency and jeopardy is the dilemma of using the RR technique and presents no easy solution. Researchers must either deal with the increased inefficiency of the estimates affecting the significance of the results or, if choosing an overly efficient RR design, run the risk of respondents refusing to participate due to high levels of respondent jeopardy. As noted earlier, with the benefits of hindsight, we found that the known probability for the non-sensitive question we used was too high. The large π_y we used substantially inflated the sampling variances of the sensitive estimators.

Another consideration is the sample size and the response rate. Sample size is constrained by the availability of resources. Higher response rates produce smaller variances of the estimates, and this will increase the efficiency. Low response rates are a concern for most surveys addressing sensitive issues. We expected to increase response rate by using the RR technique. Unfortunately, we found that ordinary people receiving a RR technique survey instrument might find the instructions difficult to comprehend, and even when the instructions were comprehensible, they still might have difficulties in appreciating the usefulness of the survey results. Furthermore, when the RR technique is used, the respondents inevitably have to spend more time to read and follow the instructions. All these factors contribute to the low response rate, which is contrary to expectations. Thus, it is a challenge to researchers to write clear RR instructions which are easy to comprehend and follow, and are able to convince the respondents that the procedure will protect their privacy, yet also provide useful data for the researchers.

In this study, the data analysis was restricted to univariate analysis. The extension of the RR technique to multivariate analysis, such as multiple regression and analysis of variance, has been discussed in the RR technique literature but none of the suggested theoretical models[24] for adjusting the RR data for multivariate analysis have been attempted by tax researchers. In a recent study, van der Heijden *et al* (1998) successfully used the RR technique and regression analysis to investigate social security fraud. Future tax research using the RR technique may consider using multivariate analyses to test more sophisticated theories of tax evasion.

ACKNOWLEDGMENT

The authors appreciate the helpful comments of the participants at the 2000 Tax Research Network Conference in Birmingham, the United Kingdom. The financial support of the Australian Research Council's Small ARC Grants is gratefully acknowledged.

APPENDIX _ Randomized Response Survey Instrument

THE AUSTRALIAN NATIONAL UNIVERSITY

Survey of Taxpayers

All responses are anonymous and will be kept strictly confidential.

If you have not lodged an income tax return for the 1997/98 tax year (the year ended 30 June 1998), please pass on this questionnaire to someone else in your household who did lodge a 1997/98 tax return for completion. If no one in your household has lodged a tax return for 1997/98, please tick 'No' in question 1, then answer questions 5 to 14.

1. Did you lodge a tax return for the 1997/98 tax year? *(Please tick one box.)*

   Yes ☐ *Please answer all the questions in Sections I, II and III.*

   No ☐ *Please answer all the questions in Sections II and III only.*

SECTION I *In this section we would like to ask you some potentially sensitive questions about your 1997/98 tax return, but we don't want to put you on the spot, so we are using a procedure that makes it safe for you to respond truthfully to each question without anyone ever knowing which question you actually answered.*

First, take a bank note from your wallet or purse and look at the LAST three digits of the serial number on the bank note. (If you don't have a bank note handy, please use the LAST three digits of your telephone number.) Do not make any note of these numbers on the questionnaire.

*We will ask three pairs of questions on the next page. You answer only one question in each pair, depending on the serial number on the bank note, **which only you know**. We will not know which question in the pair you answered; we will only be able to statistically draw some conclusions about all the respondents **as a group**.*

2. If the LAST digit of your bank note's serial number is 1, 2, or 3, answer question 2A. Otherwise, answer question 2B.

2A. Is the LAST digit of the serial number an even number?	**Your answer to 2A or 2B is:** *(Please tick one box.)*
2B. In filing your 1997/98 tax return, did you intentionally omit some of your **income** that was subject to tax?	☐ Yes ☐ No

Now go to question 3 on the next page.

3. If the SECOND LAST digit of your bank note's serial number is 1, 2, or 3, answer question 3A. Otherwise answer question 3B.

3A. Is the SECOND LAST digit of the serial number an even number? 3B. In filing your 1997/98 tax return did you intentionally claim **deductions** (expenses and losses) that you were not entitled to?	**Your answer to 3A or 3B is:** *(Please tick one box.)* ☐ Yes ☐ No

4. If the THIRD LAST digit of your bank note's serial number is 1, 2, or 3, answer question 4A. Otherwise answer question 4B.

4A. Is the THIRD LAST digit of the serial number an even number? 4B. In filing your 1997/98 tax return did you intentionally claim **tax rebates or tax credits** that you were not entitled to?	**Your answer to 4A or 4B is:** *(Please tick one box.)* ☐ Yes ☐ No

SECTION II *Please complete the following background information which will help us prepare a profile of the respondents.*

5. How old are you? *(Please tick one box.)*

☐ Under 18 years of age
☐ 18 - 30
☐ 31 - 45
☐ 46 - 60
☐ Over 60 years of age

6. Are you male or female? *(Please tick one box.)*

☐ Male
☐ Female

7. Which of the following best describes your occupation during the 1997/98 tax year? *(Please tick one box.)*

☐ Self-employed
☐ Employee
☐ Unemployed
☐ Retired

8. Was your 1997/98 tax return prepared by a tax agent? *(Please tick one box.)*

- ☐ Yes
- ☐ No

9. Which is the highest level of education that you have achieved? *(Please tick one box.)*

- ☐ No formal education
- ☐ Completed primary school
- ☐ Completed high school (year 10)
- ☐ Completed college (year 12)
- ☐ Completed tertiary education other than university (*eg* TAFE)
- ☐ Completed university

10. What was your **taxable income** (net of deductions) for the 1997/98 tax year? *(Please tick one box.)*

- ☐ Below $5,400
- ☐ $5,401 - $20,700
- ☐ $20,701 - $38,000
- ☐ $38,001 - $50,000
- ☐ Over $50,000

SECTION III *Please indicate the extent of your agreement or disagreement with the following statements by **CIRCLING a number** to help us evaluate the questionnaire we use.*

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
11. All of the questions and instructions were clear in their meaning.	1	2	3	4	5
12. I felt that my privacy was protected by the anonymity of my response.	1	2	3	4	5
13. I understood the method of selecting which question in a pair to answer using the serial number on a bank note.	1	2	3	4	5
14. I felt that my privacy was further protected by the procedure used in questions 2 to 4.	1	2	3	4	5

Thank you very much for your time. Please return the questionnaire in the enclosed reply paid envelope.

REFERENCES

Aitken, S.S. and Bonneville, L. (1980). *A general taxpayer opinion survey.* Prepared for Office of Planning and Research, Internal Revenue Service. Washington: CSR, Inc.

Berry, L.E., Harwood, G.B. and Katz, J.L. (1987). Performance of auditing procedures by governmental auditors: some preliminary evidence. *The Accounting Review 62,* 14-28.

Bradburn. N., Sudman, S. and Associates (1979). *Improving interview method and questionnaire design.* San Francisco: Jossey-Bass Publishers.

Collins, J., Milliron, V. and Toy, D. (1992). Determinants of tax compliance: a contingency approach. *The Journal of the American Taxation Association 14,* 1-29.

Elffers, H., Weigel, R.H. and Hessing, D.J. (1987). The consequences of different strategies for measuring tax evasion behavior. *Journal of Economic Psychology 8,* 311-337.

Elffers, H., Robben, H.S.J. and Hessing, D.J. (1992). On measuring tax evasion. *Journal of Economic Psychology 13,* 545-567.

Eriksson, S. (1976). Regressions analysis of data from randomized interviews. *Confidentiality in surveys. Report No. 17.* Department of Statistics, University of Stockholm.

Fox, J.A. and Tracy, P.E. (1986). *Randomized response: a method for sensitive surveys.* Beverly Hills, CA: Sage Publications Inc.

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 (June),* 520-539.

Harwood, G.B., Larkins, E.R. and Martinez-Vazquez, J. (1993). Using a randomized response methodology to collect data for tax compliance research. *The Journal of the American Taxation Association 15(3),* 79-92.

Hite, P.A. (1988). An examination of the impact of subject selection on hypothetical and self-reported taxpayer noncompliance. *Journal of Economic Psychology 9,* 445-466.

Horvitz, D.G., Shah, B.U. and Simmons, W.R. (1967). The unrelated question randomized response model. *Proceedings of the Social Statistics Section*. American Statistical Association.

Lanke, J. (1975). On the choice of the unrelated question in Simmons version of randomized response. *Journal of the American Statistical Association 70,* 80-83.

Larkins, E.R., Hume, E.C. and Garcha, B.S. (1997). The validity of the randomized response method in tax ethics research. *Journal of Applied Business Research 13(3),* 25-32.

Mason, R. and Calvin, L. (1978). A study of admitted income tax evasion. *Law and Society Review (Fall),* 73-89.

Mason, R. and Calvin, L. (1984). Public confidence and admitted tax evasion. *National Tax Journal 37(4),* 489-496.

Porcano, T.M. (1988). Correlates of tax evasion. *Journal of Economic Psychology 9,* 47-67.

Rosenberg, M.J. (1979). *Multivariable analysis by a randomized response technique for disclosure control*. Unpublished doctoral dissertation, University of Michigan.

Roth, J.A., Scholz, J.T. and Witte, A.D. (Eds.) (1989). *Taxpayer compliance - Volume 1: An agenda for research.* Philadelphia: University of Pennsylvania Press.

Soeken, K.L. and Macready, G.B. (1982). Respondents' perceived protection when using randomized response. *Psychological Bulletin 92 (September),* 487-489.

Van der Heijden, P.G.M., van Gils, G., Bouts, J. and Hox, J. (1998). A comparison of randomised response, CASAQ, and direct questioning; eliciting sensitive information in the context of social security fraud. *Kwantitative Methoden 19.* 15-34.

Vogel, J. (1974). Taxation and public opinion in Sweden: an interpretation of recent survey data. *National Tax Journal (December),* 499-513.

Wahlund, R. (1992). Tax changes and economic behavior: the case of tax evasion. *Journal of Economic Psychology 13,* 657-677.

Warner, S. (1965). Randomized response: a survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association 60,* 63-69.

Yankelovich, Skelly and White, Inc. (1984). *Taxpayers Attitudes Study: Final Report.* Prepared for the Internal Revenue Service. Washington: Government Printing Office.

Table 1 _ Association of Tax Evasion and Demographic Variables in Prior Studies

Author, Date	Income Level	Age	Occupation *	Education	Gender
Vogel, 1974**		−	+	+	−
Mason & Calvin, 1978	−	−	0		−
Mason & Calvin, 1984	−				
Porcano, 1988			+		−
Collins, Milliron & Toy, 1992					
First model***	−	−		−	+
Contingency models****	−	+/–		−	+
Wahlund, 1992		−	+		

+ denotes a positive association of the demographic variable with tax evasion.
– denotes a negative association of the demographic variable with tax evasion.
0 denotes that the association is indeterminate.
*Occupation has been defined in these studies in two ways: (a) self-employed versus employee, and (b) blue-collar versus white-collar.
**The findings of Vogel (1974) were actually based on measures of taxpayer attitudes: taxpayers were asked if they considered themselves the type of person to evade tax. Vogel (1974) is included here as the measures found are close to those from a direct assessment of evasion.
***These are the results from the first model developed by Collins *et al* (1992) following traditional guidelines.
****These are the results from the contingency models. +/– indicates that both positive and negative associations were found for different contingencies (*i.e.,* using a tax preparer and the individuals preparing the return themselves.)

Table 2 _ Response Rates and Non-response Bias

	Randomized Response Instrument	%	Direct Questioning Instrument	%	z-score
Number originally sent	1,500		500		
Number returned undelivered	225		66		
Effective mail-out	1,275	100%	434	100%	
Responses received*	354	27.8%	150	34.6%	-2.61#
Invalid responses**	26	2.0%	9	2.1%	
1997/98 tax return not yet lodged***	44	3.5%	20	4.6%	
Usable responses*	284	22.3%	121	27.9%	-2.29#

*The response rates and usable response rates were calculated based on the effective mail-out.
**Returned survey instruments were classified as invalid and were not included in the data set if the entire questionnaire was left blank, or the majority of the questionnaire had not been completed, or the respondents clearly did not follow the RR technique properly.
***Survey instruments returned by respondents who had not yet lodged a return for the 1997/98 tax year were also excluded, as the tax evasion questions referred to the 1997/98 tax year.
Statistically significant at the 0.05 level.

Table 3 _ Prevalence of Tax Evasion and Response Bias

Type of Evasion	Prevalence RR Instrument	Prevalence DQ Instrument	z-score
Under-reporting income (Q2)	5.5%	1.7%	1.22
Over-claiming deductions (Q3)	6.5%	4.2%	0.66
Over-claiming tax rebates or tax credits (Q4)	-1.6%*	0.0%	n/a
Under-reporting income and/or over-claiming deductions	7.1%**	5.0%	n/a

*A negative estimated proportion is possible using equation (7) when the realized π_x is close to zero and the realized π_y is smaller than its theoretical value (1/3).
**This proportion is computed using $\pi_{x1} + \pi_{x2} - \pi_{x1x2}$ and equation (11).

Table 4 _ Income Level and Tax Evasion (RR Instrument)

Type of Evasion	Level of Taxable Income	No. of Responses	Proportion of Evasion	z-score
Under-reporting income	Up to $38,000	175	6.9%	
	$38,001 and above	104	3.6%	0.56
Over-claiming deductions	Up to $38,000	176	9.3%	
	$38,001 and above	104	2.2%	1.18

Table 5 _ Age and Tax Evasion (RR Instrument)

Type of Evasion	Age Group	No. of Responses	Proportion of Evasion	z-score
Under-reporting income	18 - 45 years	111	6.3%	
	46 and above	170	5.0%	0.21
Over-claiming deductions	18 - 45 years	111	10.2%	
	46 and above	171	4.1%	0.97

Table 6 _ Education and Tax Evasion (RR Instrument)

Type of Evasion	Education Group*	No. of Responses	Proportion of Evasion	z-score
Under-reporting income	Non-tertiary	138	6.4%	
	Tertiary	142	4.8%	0.27
Over-claiming deductions	Non-tertiary	139	9.4%	
	Tertiary	142	3.8%	0.92

*Non-tertiary education includes taxpayers with no formal education and those that have completed education up to year 12. Tertiary education includes taxpayers who have a tertiary education other than university and those who have completed university education.

Table 7 _ Occupation and Tax Evasion (RR Instrument)

Type of Evasion	Occupation	No. of Responses	Proportion of Evasion	z-score
Under-reporting income	Self-employed	79	16.5%	
	Employee	144	3.6%	1.68*
Over-claiming deductions	Self-employed	79	9.2%	
	Employee	144	7.5%	0.23

* Statistically significant at the .05 level (one-tailed test).

Table 8 _ Gender and Tax Evasion (RR Instrument)

Type of Evasion	Gender	No. of Responses	Proportion of Evasion	z-score
Under-reporting income	Female	101	8.3%	
	Male	180	4.0%	0.70
Over-claiming deductions	Female	101	11.2%	
	Male	181	3.9%	1.13

Table 9 _ Tax Agent Use and Tax Evasion (RR Instrument)

Type of Evasion	Return Prepared by Tax Agent	No. of Responses	Proportion of Evasion	z-score
Under-reporting income	Yes	223	6.2%	
	No	57	0.8%	0.82
Over-claiming deductions	Yes	224	7.4%	
	No	57	0.8%	0.99

[1] Tax audits carried out by the Australian Taxation Office focus more on measuring non-compliance, and can only infer evasion behavior, rather than directly measure it. Experimental designs cannot directly measure tax evasion either, as many rely on the assumption that intended behavior, and behavior occurring under laboratory conditions, is equivalent to actual behavior.

[2] Response bias refers to any systematic distortion of the respondents' true answer and this endangers the validity of the survey measurements. Non-response bias arises when the respondents answering the survey are not representative of the sample, and there are systematic differences between the respondents who do and do not respond to the survey questions.

[3] This suggestion is attributed to Simmons and the unrelated question design is often referred to as Simmon's model.

[4] The unknown distribution refers to the distribution of answers to the unrelated question which is estimated as part of the design.

[5] The unrelated question design with a known distribution was mentioned only briefly in the Horvitz *et al* (1967) article.

[6] See Fox and Tracy (1986) for a review.

[7] Although the unrelated question was proposed by Simmons and first published by Horvitz *et al* (1967) it is referred to throughout this paper as being developed by Greenberg *et al* (1969) as this was the more popular and detailed paper outlining the technique, and has been used as the basis for most studies.

[8] This article was unavailable. The reported results were taken from Roth *et al* (1989:343).

[9] In Australia tax preparers are referred to as tax agents.

[10] This is a common assumption made in comparison studies involving the RR technique (for example, see Larkins et al 1997).

[11] Elffers *et al* (1992) compared three different measures of tax evasion. No correlation was found to exist between self-reports and the experiment data, and the authors suggested that the lack of correlation might have been caused by the different methodologies measuring different aspects of tax evasion.

[12] There is some doubt as to the validity of this assumption (Hite 1988).

[13] For example, see Collins *et al* (1992) who examine tax evasion using a contingency approach, with one of the contingencies being whether a taxpayer uses the services of a tax preparer to prepare their return.

[14] The only other difference between the two surveys is the omission in the DQ instrument of two questions in section III that evaluate the RR technique used.

[15] Lack of a randomizing device can lower the response rate and reduce the effectiveness of the RR technique (Larkins *et al* 1997).

[16] For example, targeting respondents with a prison record would allow researchers to use reliable demographic data as the details may be available from their prison record.

[17] See Bradburn and Sudman (1979) who used birth months; though this study was flawed through over use of birth months creating respondent suspicion.

[18] See Fox and Tracy (1986:52).

[19] Respondents were given the option of using the last three digits of their telephone number, but this may have been forgotten as it was only mentioned once in the instructions, and the questions all refer to the serial number on a bank note.

[20] A negative estimated proportion is possible using equation (7) when π_x was close to zero and the realized π_y was smaller than its theoretical value (1/3).

[21] Other studies in the accounting and tax areas using the RR technique also encountered similar problem of obtaining results in hypotheses testing which are not statistically significant (Berry *et al* 1987; Larkins *et al* 1997).

[22] The relative efficiency of the unrelated question approach with a known distribution compared to Warner's model was illustrated in Section II.

[23] Recall that setting p equal to 1 results in direct questioning, as all the respondents are directed to answer the sensitive question, and setting π_y equal to zero means that all 'yes' responses can only refer to the sensitive question.

[24] See Rosenberg (1979) and Eriksson (1976).