methods, data, analyses | Vol. 13(2), 2019, pp. 307-320 DOI: 10.12758/mda.2019.03

Sensitive Question Techniques and Careless Responding: Adjusting the Crosswise Model for Random Answers

Patrick Schnapp

Abstract

The crosswise model is a popular sensitive question technique often considered more accurate than direct questioning. When this technique is used, a sensitive question is paired with a nonsensitive question that has a known prevalence and respondents are asked to give a joint answer to the pair of questions. Recent research has shown that prevalence estimates based on the crosswise model are biased towards 50% when respondents answer randomly, and that random answers are frequent. I develop methods to adjust the crosswise model for self-reported random answers. Results from an exploratory online survey (n = 103) show that (i) fewer respondents report random answers than might be expected given unadjusted results, (ii) results differ considerably between questions, and (iii) one of three questions yields an estimate that is substantially and significantly above the true value even after adjusting for random answers.

Keywords: Crosswise model; randomized response technique; sensitive question techniques; socially desirable responding; careless responding; random answers; satisficing

Survey researchers are often interested in answers to sensitive questions, i.e., "questions about violations of social norms by the respondent's behavior" (Näher & Krumpal, 2012, p. 1603).[1] When answering such questions, some respondents exhibit socially desirable responding, "the tendency to give overly positive self-descriptions" (Paulhus, 2002, p. 50). For example, many respondents who have engaged in drunk driving will not admit this in a survey (Locander, Sudman, & Bradburn, 1976). This entails two problems. First, researchers who take respondents' answers at face value will underestimate the prevalence of undesirable characteristics and overestimate the prevalence of desirable characteristics. Second, researchers' ability to estimate associations between the characteristic in question and other variables will be impeded, as the respondents' answers are influenced by both their true status on the characteristic and their tendency to engage in socially desirable responding (Wolter & Preisendörfer, 2013).

One approach to this problem is the use of sensitive question techniques. These allow the respondent to hide his or her answer to the sensitive question, but also allow the researcher to estimate the prevalence of the sensitive characteristic in the sample as a whole, as well as associations between the sensitive and other characteristics.

A technique recently popular is the crosswise model (CM), a variant of the randomized response technique (Warner, 1965) introduced by Yu, Tian, and Tang (2008). In the CM, respondents are asked to reply to a combination of two yes/no questions. One is the sensitive question, the other is nonsensitive. For example, the respondent may be asked (i) whether he or she has ever engaged in drunk driving and (ii) whether his or her mother was born in January, February or March. The respondent is then asked whether (A) the answer to the two questions is the same (both "yes" or both "no") or (B) the answers to the questions are different. The prevalence of the sensitive characteristic (drunk driving) can be estimated because the prevalence of the nonsensitive item (mother's month of birth) is known (approx. .25).

Prevalence estimates on the basis of the CM are often significantly higher than those derived from direct questions (Enzmann, 2017; Hoffmann et al., 2015; Hoffmann & Musch, 2016, 2018; Höglinger & Jann, 2018; Höglinger, Jann, &

1 The term "sensitive question" is also used for related but different concepts (for in-depth discussions, Krumpal, 2013; Tourangeau & Yan, 2007).

Acknowledgments

I thank Thomas Ullmann of maQ for technical support, all survey respondents for participating, and two reviewers for helpful comments.

Direct correspondence to

Patrick Schnapp

E-mail: p_schnapp@gmx.de

Diekmann, 2016; Hopp & Speil, 2019; Jann, Jerke, & Krumpal, 2012; Korndörfer, Krumpal, & Schmukle, 2014; Kundt, 2014; Waubert de Puiseau, Hoffmann, & Musch, 2017). Many researchers interpret this as evidence for the superiority of the CM (e.g., Hopp & Speil, 2019; Kazemzadeh et al., 2016; Kundt et al., 2017; Waubert de Puiseau et al., 2017). These authors rely on the more-is-better assumption, according to which techniques that yield higher prevalence estimates of sensitive characteristics are more valid. The assumption is unwarranted in the case of the CM. This is because the more respondents choose an answer at random, the more the prevalence estimate will be biased towards 50% (Enzmann, 2017). Random answers will hence bias estimates downwards when the true prevalence is above 50% and upwards when the true prevalence is below 50%, as is often the case with sensitive characteristics (Höglinger & Diekmann, 2017). While the same is true of direct questions (Hemenway, 1997), the fact that CM questions are more complex makes it more likely that respondents will answer randomly because they are unwilling or unable to put in the cognitive effort necessary for choosing the correct response. Thus, the more-is-better assumption may lead to the conclusion that the CM produces more valid results than direct questions, when in fact the higher estimates are a consequence of random responding (Höglinger & Diekmann, 2017).

Three recent studies shine a light on this issue. Höglinger and Diekmann (2017) asked respondents whether they had ever received a donated organ and whether they had ever suffered from Chagas disease. The prevalence of both characteristics was assumed to be zero. Under the assumption of no other causes for bias, the rate of random answers is twice the false positive rate. After removing cases based on apparently problematic nonsensitive items, prevalence estimates were 6% (organ) and 1% (Chagas), implying random answer rates of 12% and 2%, respectively. Höglinger and Jann (2018), studying cheating in games, estimated that the CM misclassified 14% of respondents, implying 28% answered randomly under the same assumption. Enzmann (2017) reported results from a survey asking students to report illegal behavior. In a follow-up to one of the CM questions, respondents were asked how they had answered the question, with 13% stating they had answered randomly.

There also are studies showing that standard CM prevalence estimates are close to the true values (Hoffmann et al., 2015; Hoffman & Musch, 2016, 2018). It is important to appreciate what this does and does not show. These studies are evidence in favor of the accuracy of the CM as a measure of prevalence. It is unknown, however, whether individual respondents in these studies answered truthfully or not. In aggregate estimates, incorrect answers in both directions may even each other out to produce an estimate close to the true value (Höglinger & Diekmann, 2017). Incorrect answers at the individual level impede researchers' ability to correctly estimate the association between the sensitive characteristic and other variables (Enzmann, 2017), even if aggregate prevalence estimates are correct. The appeal of Höglinger & Diekmann's (2017) validation strategy of using zero-prevalence items is that it allows the researcher to estimate the rate of false positives even if the distribution of the sensitive characteristic cannot be measured.

This body of research suggests a potential remedy to the problem of bias due to random answers, and a way of testing its validity. In a survey, CM questions may be followed by direct questions asking whether the respondent answered the CM question randomly. Adjusted prevalence estimates can be calculated. These estimates can be compared to the unadjusted estimates and known true values. The present paper reports the results of a small, exploratory study to demonstrate the application of the technique and present first results.

The remainder of this article is organized as follows. In the theoretical part of the paper, I review formulae for the standard crosswise model and variants that feature adjustments. In the empirical part, I report results of the exploratory survey, showing estimates on the basis of different versions of the CM. Finally, results are discussed.

Crosswise Estimation

The Standard Crosswise Model

Suppose we are interested in the prevalence of a sensitive behavior, such as drunk driving. We could present respondents with the question about the respondents' drunk driving and couple it with a question about a nonsensitive matter, such as whether the respondent's mother's birthday is between January and March. We then ask whether (A) the answer to the two questions is the same (both "yes" or both "no") or (B) the answers to the two questions is different. When the standard CM is applied, the estimator of the prevalence of the sensitive characteristic is (Yu et al., 2008, notation altered)

$$\hat{\pi}_{SCM} = \frac{\hat{\lambda} + p - 1}{2p - 1}; p \neq 0.5 \tag{1}$$

where $\hat{\pi}_{SCM}$ is the estimate of the proportion of respondents carrying the sensitive characteristic estimated by the standard CM, $\hat{\lambda}$ is the estimate of the proportion of respondents whose true answer is "A" ("the same"), and p is the proportion of respondents for whom the true answer to the nonsensitive question is "yes". The proportion of respondents for whom the true answer is "A" is estimated as (Yu et al., 2008, notation altered)

$$\hat{\lambda} = \frac{n_A}{n} \tag{2}$$

where n_A is the sum of “A” answers and n is the sample size.

An unbiased estimator of the variance of $\hat{\pi}_{SCM}$ is (Yu et al., 2008)

$$\bar{\mathbb{V}}\left[\hat{\pi}_{SCM}\right]=\frac{\hat{\lambda}\left(1-\hat{\lambda}\right)}{(n-1)(2p-1)^2}=\frac{\hat{\pi}_{SCM}\left(1-\hat{\pi}_{SCM}\right)}{(n-1)}+\frac{p(1-p)}{(n-1)(2p-1)^2}\,;\, p\neq 0.5 \tag{3}$$

The right-hand side of the equation shows the decomposition of the variance into the sampling part and an additional term due to the uncertainty introduced by the use of the nonsensitive item (Kundt, 2014).

Adjusting the Crosswise Model for Random Answers

Some respondents may answer CM questions randomly (choosing either answer with equal probability). As can be seen from formulae (2) and (1), this biases $\hat{\lambda}$. and hence $\hat{\pi}_{SCM}$ toward 0.5. However, we may be able to estimate the proportion of CM questions that were answered randomly (e.g., on the basis of follow-up questions). Then estimates adjusted for random answers may be derived by (Enzmann, 2017, notation altered)

$$\hat{\pi}_{CMR-S}=\frac{\hat{\pi}_{SCM}-0.5r}{1-r}\,;\, r\neq 1 \tag{4}$$

where *CMR-S* stands for “crosswise model adjusted for random answers at the level of the sample” and r is the proportion of random answers. When $r = 1$, the result is undefined. This is as it should be; if all respondents answer randomly, $\hat{\pi}_{SCM}$ carries no information about the true value of π.

Equation (4) implies at the variance may be calculated as

$$\bar{\mathbb{V}}\hat{\pi}_{CMR-S}=\frac{\bar{\mathbb{V}}\left[\hat{\pi}_{SCM}\right]}{(1-r)^2}=\left[\frac{\hat{\pi}_{SCM}\left(1-\hat{\pi}_{SCM}\right)}{(n-1)}+\frac{p(1-p)}{(n-1)(2p-1)^2}\right]/(1-r)^2\,;r\neq 1 \tag{5}$$

This variance is hence larger than the variance of the standard crosswise model if $r > 0$, reflecting the added uncertainty introduced by random answers.

However, if information about random answering can be linked to individual respondents’ answers to individual items, this can be taken into account more directly in what I call the CMR-I (crosswise model adjusted for random answers at the individual level). If the respondent stated that he answered randomly, then the value for A should be set to 0.5; if he did not, the value should remain unchanged (i.e., 1 for an “A” answer and 0 for a “B” answer). Formally,

$$A_{adj} = 0.5R + (Y = A)(1 - R) \tag{6}$$

where A_{adj} is the adjusted value for the A variable at the individual level, R is an indicator variable taking the value 1 if the respondent answered the question randomly and $Y = A$ is the unadjusted value for the answer to the CM question, taking the value 1 for an "A" answer and 0 for a "B" answer. As can be seen from the formula, both "A" answers ($A = 1$) and "B" answers ($A = 0$) are converted to 0.5 if the respondent answered the question randomly ($R = 1$); otherwise ($R = 0$), they remain unchanged. $n_{A_{adj}}$ (the sum of the A_{adj} variable) can then be used instead of n_A (the sum of the unadjusted variable) in (2). In the crosswise model adjusted for random answers at the individual level, the estimate is hence given by (1), (6) and

$$\hat{\lambda} = \frac{n_{A_{adj}}}{n} \tag{7}$$

The variance is given by (3) rather than (5) under the assumption that information on random answering is correct.

Data and Methods

An online survey was conducted. The German-language questionnaire was designed to produce a sufficient number of random answers to test whether adjusting for them leads to improvements in the estimate, but no attempt was made to actively confuse participants. After an introductory page, participants received instructions on how to answer CM questions (but no practice examples). This was followed by the main part of the questionnaire. Following Höglinger and Diekmann (2017), lifetime prevalences of three rare diseases were chosen as zero-prevalence items (Castleman disease, Chagas disease, Barth syndrome). Similar to Diekmann (2012) and Kundt (2014), one nonsensitive item asked about the respondent's house number; the other two are standard nonsensitive questions in the CM literature (concerning mother's and father's month of birth being January or February). Each crosswise question was followed by a companion question on the next page. Respondents were informed that many participants find these types of questions hard to answer and asked whether they had "just chosen an answer at random" (answers: yes/no). English translations of all CM and companion questions are displayed in the appendix. Sociodemographic information was also collected. The last question asked whether respondents had answered this survey before; this was accompanied by the information that their answer would have no influence on their obtaining the code they needed to gain points (see below). CM and companion questions were obligatory, other questions were optional. Respondents had to click through to the last page of the questionnaire to obtain the code.

The survey was programmed in maQ (Ullmann, 2004) and posted on two sites, Survey Circle and Poll Pool. On both, members can fill in surveys to gain points. The more points a member has, the more points other members can gain when filling in his or her survey. The questionnaire's last page contained codes necessary to obtain points. The questionnaire was advertised as a "Short profile on health", open to all participants who were at least 18 years old and had passed their last school exam in Germany. Answers were collected in October and November 2018.

Data on months of birth in Germany were obtained directly from the Federal Statistical Office and date back to 1948. The distribution of first digits of house numbers was taken from Kundt (2014). Age was approximated by subtracting the year of birth from 2018. Answers to CM questions were set to missing if their companion questions had not been answered.

Results

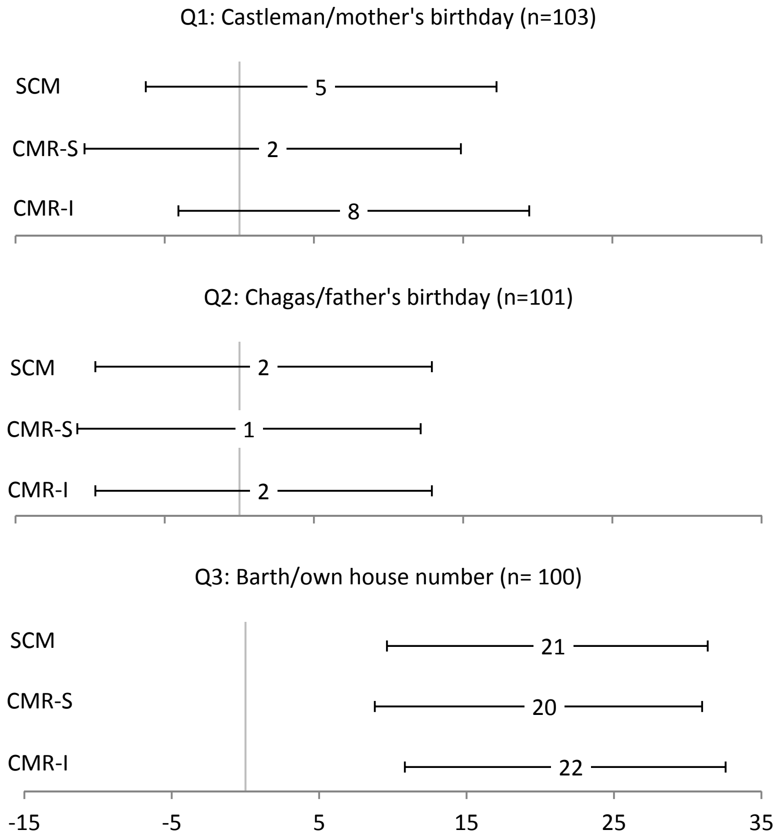
One hundred and nine respondents answered at least one combination of CM and companion question. Six respondents were excluded because they stated they had answered the survey before or were not sure. The sample size is hence 103, but there is some missing data. Descriptive statistics are shown in Table 1. The youngest respondents were born in 1998. Under the assumption that parents would have been at least 20 years old when the respondents were born, the proportion of parents born in January or February was calculated for the years 1948 to 1978; the result is approx. 16.7 percent. The proportion of German house numbers starting with the digit 8 or 9 is approx. 8.8% (Kundt, 2014).

Table 1 Descriptive statistics.

	Mean / Proportion	*SD*	Minimum	Maximum	*n*
Gender (1=male)	0.34	0.05	0	1	98
Passed "Abitur" exam	0.91	0.03	0	1	100
Year of Birth	1991.64	5.79	1962	1998	99
Age	26.36	5.79	18	54	99
Answered randomly					
Q1	0.07	0.02	0	1	103
Q2	0.02	0.01	0	1	101
Q3	0.02	0.01	0	1	100

Abbreviations. SD: standard deviation; Q: question

Figure 1 Estimates of the percentages of respondents carrying the sensitive characteristic



Abbreviations. SCM: standard crosswise model; CMR-S: crosswise model adjusted for random answers at the sample level; CMR-I: crosswise model adjusted for random answers at the individual level.

The proportions of self-reported random answers are low and differ considerably between questions. They are 6.8% for Q3 (question 3), 2.0% for Q2, and 2.0% for Q3.

Figure 1 presents point estimates and 95% confidence intervals for all combinations of questions and types of CM. In all cases, an unbiased point estimate would be zero. All point estimates are above zero, but the size of the bias differs considerably between questions. Most strikingly, answers to Q3 depart substantially and significantly from the true value. Under the assumption of no other causes for bias, this implies that 42 percent of respondents answered Q3 randomly.

These question effects are larger overall than the effects of the estimation method. The CMR-I improves on the standard estimates in no case and does worse in two. In contrast, the CMR-S fares a little better than both the standard method and the CMR-I in all cases. These differences are far from significant, however.

Discussion

Results from this small study show that point estimates based on the standard CM are higher than the true value; in one case, the difference is very large and statistically significant. This result adds to validation studies showing that the more-is-better assumption is invalid in the case of the CM. It also casts doubt on results from the literature in which the CM was applied. If readers consult such studies, they could apply a mental correction for random responding using the formulae given above and reasonable assumptions about the proportion of responses that are random. In this context, note that the present results were obtained from respondents who were extrinsically motivated to participate. A substantial proportion of respondents probably participated despite low intrinsic interest, and Brower (2018) reports negative associations between respondent interest and measures of careless responding. It hence seems likely that the bias observed in this study, as well as other CM studies using incentives for participation, is higher than it would have been if respondents had been intrinsically motivated to participate.

One may wonder why the results are so different for Q3. Possible reasons include (i) the position of the question, (ii) the content of the sensitive item (perhaps respondents mistook Barth syndrome for something else), (iii) the prevalence of the nonsensitive item; (iv) the person the nonsensitive question referred to (self), and (v) the topic of the nonsensitive question (house number).

The position of the sensitive question may seem unlikely to have played a large role given that Höglinger and Diekmann (2017) found no substantial or significant positional effects. However, it is possible that by the time they reached Q3, some respondents were sufficiently disappointed by the contents of the survey to start giving random answers. Respondents who start a survey advertised as a "Short profile on health" may expect questions concerning their exercise and eating habits rather that questions about rare diseases in an unusual question format. Such an effect could be particularly strong if respondents who live a healthy lifestyle self-selected into the survey because they were looking forward to presenting themselves in a favorable light, an opportunity not given by the questionnaire. Some of these respondents might have combined random answers to Q3 with an untruthful answer to the companion question.

Concerning the content of the sensitive item, it is unclear how the respondents might have misunderstood what “Barth syndrome” refers to.

While the prevalence of the nonsensitive item in Q3 is low and hence accords the respondent little protection of his privacy, Diekmann (2012) showed that the students in his sample overestimated the proportion of house numbers starting in high digits. However, it is conceivable that respondents were unwilling to answer truthfully to a question involving their own house number in a time in which data security had been a prominent topic in the German media due to the introduction of the new General Data Protection Regulation.

There is no definitive answer to the question why Q3 performed so much worse than the other two. To avoid such uncertainty, future researchers wishing to test the CMR-I may prefer to vary features of sensitive and nonsensitive questions randomly.

The main result is that adjusting for random answering, as implemented in this study, does little to remove bias from CM results. A number of potential explanations present themselves. First, the companion question did not ask about deliberately false answers. Hence, the study was not designed to remove bias from this source, if any. Second, the companion questions themselves were sensitive, as answering randomly in a survey violates the norm of honesty. This may lead to socially desirable responding to these questions, impeding sufficient adjustments. Third, a yes/no question is too crude to measure random answering. A scale with more than two points may be preferable, as such scales generally exhibit better psychometric properties than binary scales (Krosnick & Presser, 2010; Markon, Chmielewski, & Miller, 2011) and can measure degrees of certainty that the correct answer was given. Ideally, such a question would also capture the fact that some respondents intentionally try to give the wrong answer, but may not be sure whether they succeeded in doing so. Authors who would like to pursue this avenue of research may also want to test whether it is really helpful to ask a companion question after every CM question – a design feature that seems impractical for applied surveys. Results may be equally good or better if the questionnaire presents a battery of CM questions followed by a summary companion question asking respondents what proportion of CM questions they answered randomly.

These may be fruitful avenues for future research. While the results of this and other studies suggest that the crosswise model has shortcomings, the problem of socially desirable responding is too serious to give up on techniques that may lead to viable solutions after all.

References

Brower, C. K. (2018). *Too long and too boring: The effects of survey length and interest on careless responding* (Master's Thesis). Fairborn: Wayne State University.

Diekmann, A. (2012). Making use of ''Benford's law'' for the randomized response technique. *Sociological Methods and Research*, *41*(2), 325-334. https://doi.org/10.1177/0049124112452525

Enzmann, D. (2017). Die Anwendbarkeit des Crosswise-Modells zur Prüfung kultureller Unterschiede sozial erwünschten Antwortverhaltens: Implikationen für seinen Einsatz in internationalen Studien zu selbstberichteter Delinquenz. In S. Eifler & F. Faulbaum (Eds.), *Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung* (pp. 231-269). Wiesbaden: VS.

Hemenway, D. (1997). The myth of millions of annual self-defense gun uses: A case study of survey overestimates of rare events. *Chance, 10*(3), 6-10. https://doi.org/10.1080/09332480.1997.10542033

Hoffmann, A., Diedenhofen, B., Verschuere, B., & Musch, J. (2015). A strong validation of the crosswise model using experimentally-induced cheating behavior. *Experimental Psychology, 40*(6), 403-414. https://doi.org/10.1027/1618-3169/a000304

Hoffmann, A., & Musch, J. (2016). Assessing the validity of two indirect questioning techniques: A stochastic lie detector versus the crosswise model. *Behavior Research Methods, 48*(3), 1032-1046. https://doi.org/10.3758/s13428-015-0628-6

Hoffmann, A., & Musch, J. (2018). Prejudice against women leaders: Insights from an indirect questioning approach. *Sex Roles*. Advance online publication. https://doi.org/10.1007.s11199-018-0969-6

Höglinger, M., & Diekmann, A. (2017). Uncovering a blind spot in sensitive question research: False positives undermine the crosswise-model RRT. *Political Analysis, 25*(1), 131-137. https://doi.org/10.1017/pan.2016.5

Höglinger, M., & Jann, B. (2018). More is not always better: An experimental individual-level validation of the randomized response technique and the crosswise model. *PLoS ONE, 13*(8), e0201770. https://doi.org/10.1371/journal.pone.0201770

Höglinger, M., Jann, B., & Diekmann, A. (2016). Sensitive questions in online surveys: An experimental evaluation of different implementations of the randomized response technique and the crosswise model. *Survey Research Methods, 10*(3), 171-187. https://doi.org/10.18148/srm/2016.v10i3.6703

Hopp, C., & Speil, A. (2019). Estimating the extent of deceitful behaviour using crosswise elicitation models. *Applied Economics Letters, 26*(5), 396-400. https://doi.org/10.1080/13504851.2018.1486007

Jann, B., Jerke, J., & Krumpal, I. (2012). Asking sensitive questions using the crosswise model: An experimental survey measuring plagiarism. *Public Opinion Quarterly, 76*(1), 32-49. https://doi.org/10.1093/poq/nfr036

Kazemzadeh, Y., Shokoohi, M., Baneshi, M. R., & Haghdoost, A. A. (2016). The frequency of high-risk behaviors among Iranian college students using indirect methods: Network scale-up and crosswise model. *International Journal of High Risk Behaviors & Addiction, 5*(3), e25130. https://doi.org/10.5812/ijhrba.25130

Korndörfer, M., Krumpal, I., & Schmukle, S. C. (2014). Measuring and explaining tax evasion: Improving self-reports using the crosswise model. *Journal of Economic Psychology, 45*(1), 18-32. https://doi.org/10.1016/j.joep.2014.08.001

Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wright (eds.), *Handbook of Survey Research* (pp. 263-313). Bingley: Emerald.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality and Quantity, 47*(4), 2025–2047. https://doi.org/10.1007/s11135-011-9640-9

Kundt, T. (2014). *Applying "Benford's law" to the crosswise model: Findings from an online survey on tax evasion* (Working Paper Series No. 148). Hamburg: Helmut Schmidt Universität.

Kundt, T. C., Misch, F., & Nerré, B. (2017). Re-assessing the merits of measuring tax evasion through business surveys: An application of the crosswise model. *International Tax and Public Finance, 24*(1), 112-133. https://doi.org/10.1007/s10797-015-9373-0

Locander, W., Sudman, S., & Bradburn, N. (1976). An investigation of interview method, threat and response distortion. *Journal of the American Statistical Association, 71*(354), 269-275. https://doi.org/10.1080/01621459.1976.10480332

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin, 137*(4), 856-879. https://doi.org/10.1037/a0023678

Näher, A.-F., & Krumpal, I. (2012). Asking sensitive questions: The impact of forgiving wording and question context on social desirability bias. *Quality & Quantity, 46*(5), 1601-1616. https://doi.org/10.1007/s11135-011-9469-2

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*(5), 859-883. https://doi.org/10.1037/0033-2909.133.5.859

Ullmann, T. D. (2004). *maQ-Fragebogengenerator: Make a questionnaire*. Retrieved from http://maq-online.de

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association, 60*(309), 63-69. https://doi.org/10.1080/01621459.1965.10480775

Waubert de Puiseau, B., Hoffmann, A., & Musch, J. (2017). How indirect questioning techniques may promote democracy: A preelection polling experiment. *Basic and Applied Social Psychology, 39*(4), 209-217. https://doi.org/10.1080/01973533.2017.1331351

Wolter, F., & Preisendörfer, P. (2013). Asking sensitive questions: An evaluation of the randomized response technique versus direct questioning using individual validation data. *Sociological Methods & Research, 42*(3), 321-353. https://doi.org/10.1177/0049124113500474

Yu, J. W., Tian, G. L., & Tang, M. L. (2008). Two new models for survey sampling with sensitive characteristic: Design and analysis. *Metrika, 67*(3), 251-263. https://doi.org/10.1007/s00184-007-0131-x

Appendix A

Question wordings and answer options for the crosswise model

Page 3

Here are two questions:

A: Was your **mother** born in January or February?

B: Have you ever been diagnosed with Castleman disease?

Page 5

Here are two questions:

A: Was your **father** born in January or February?

B: Have you ever been diagnosed with Chagas disease (a.k.a. American trypanosomiasis)?

Page 7

Here are two questions:

A: Please think of your **main residence**. Is the **first digit** of your house number 8 or 9?

B: Have you ever been diagnosed with Barth syndrome?

Item on pages 3, 5, 7

Which of the following statements is true?

(Answers: The answer to both questions is the same (twice yes or twice no) / The answers to the two questions are different (once yes and once no, irrespective of the order))

Pages 4. 6, 8 (companion question)

Many respondents find he type of question you have just answered hard.

Is the following statement correct?

Answering the question on the previous page, I just chose an answer at random. (Answers: yes / no)