Sociological Methods & Research

<http://smr.sagepub.com>

Log-Linear Randomized-Response Models Taking Self-Protective Response Behavior Into Account

Maarten J. L. F. Cruyff, Ardo van den Hout, Peter G. M. van der Heijden and
Ulf Böckenholt

Sociological Methods Research 2007; 36; 266

DOI: 10.1177/0049124107301944

The online version of this article can be found at:
<http://smr.sagepub.com/cgi/content/abstract/36/2/266>

Published by:

SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *Sociological Methods & Research* can be found at:

Email Alerts: <http://smr.sagepub.com/cgi/alerts>

Subscriptions: <http://smr.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 12 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
<http://smr.sagepub.com/cgi/content/refs/36/2/266>

Log-Linear Randomized-Response Models Taking Self-Protective Response Behavior Into Account

Maarten J. L. F. Cruyff

Utrecht University, Netherlands

Ardo van den Hout

MRC Biostatistic Unit, Cambridge, United Kingdom

Peter G. M. van der Heijden

Utrecht University, Netherlands

Ulf Böckenholt

McGill University, Montreal, Canada

Randomized response (RR) is an interview technique designed to eliminate response bias when sensitive questions are asked. In RR the answer depends partly on the true status of the respondent and partly on the outcome of a randomizing device. Although RR elicits more honest answers than direct questions do, it is susceptible to self-protective response behavior; that is, the respondent gives an evasive answer irrespective of the outcome of the randomizing device. The authors present a log-linear RR model that accounts for this kind of self-protection (SP). The main results of this SP model are estimates of (1) the probability of SP, (2) the log-linear parameters describing the associations between the sensitive characteristics, and (3) the prevalence of the sensitive characteristics that are corrected for SP. The model is illustrated with two examples from a Dutch survey measuring noncompliance with social welfare rules.

Keywords: *randomized response; log-linear model; self-protective response behavior; regulatory noncompliance*

Since most people are reluctant to answer questions about sensitive topics such as the use of drugs or alcohol, sexuality, or antisocial behavior, sensitive characteristics are often underreported in surveys and

questionnaires. Randomized response (RR) is an interview technique that is especially designed to eliminate evasive response bias (Chaudhuri and Mukerjee 1988; Warner 1965). In the RR design, the answer is to a certain extent determined by the outcome of a randomizing device, for example, a pair of dice or the draw of a card. Since the outcome is known only to the respondent, confidentiality is guaranteed. A meta-analysis shows that RR yields more valid prevalence estimates than direct-questioning designs (Lensvelt-Mulders et al. 2005).

Although the respondents' privacy is protected, RR does not completely eliminate evasive response bias. Several studies show that some respondents do not always give the affirmative answer when this is required by the RR design. In line with Böckenholt and van der Heijden (2004, forthcoming), we refer to this answer strategy as self-protection (SP). Edgell, Himmelfarb, and Duncan (1982) showed the presence of SP in an experimental study in which the outcomes of the randomizing device were fixed in advance. To a question about having experiences with homosexuality, 25 percent of the respondents who had to answer *yes* by design gave an SP *no* response. In another study, van der Heijden and colleagues (2000) applied different interview techniques to participants identified as having committed social welfare fraud. Although the RR condition elicited more admission of fraud than direct questioning or computer-assisted self-interviews, a substantial percentage of the participants still denied having committed fraud. In a study by Boeije and Lensvelt-Mulders (2002), most of the respondents who participated in a computer-assisted RR survey found it difficult to give a false *yes* response and some of them admitted that they had answered *no*.

Some studies have recently focused on the detection and estimation of SP in RR designs. Clark and Desharnais (1998), who used the term *cheating* to denote SP, proposed to split the sample in two groups and assign different randomization probabilities to each group. They showed that significant cheating can be detected if cheating behavior and randomization probability are assumed to be independent. A multivariate approach was taken by Böckenholt and van der Heijden (2004), who assumed an underlying noncompliance scale for a set of RR variables and estimated SP using an item-response model.

In this report, we present a log-linear modeling approach to account for SP in an RR design. This SP model is derived from the log-linear RR

Authors' Note: Please address correspondence to Maarten J. L. F. Cruyff, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, Netherlands; e-mail: m.cruyff@uu.nl.

(LLRR) model (Chen 1989) by the introduction of an SP parameter. The three main results of the SP model are (1) an estimate of the probability of SP, (2) log-linear parameter estimates describing the associations between RR variables, and (3) prevalence estimates of the sensitive behavior corrected for SP. The model is illustrated with two examples from the 2000 Social Welfare Survey conducted in the Netherlands (Van Gils, van der Heijden, and Rosebeek 2001; see also Lensvelt-Mulders et al. 2006).

In the remainder of this article, we present the questions and the RR design used in the Social Welfare Survey. We introduce the general RR model and show that identification problems arise when an SP parameter is included. We present the SP model as an extension of the LLRR model. We provide two examples from the Social Welfare Survey and investigate the robustness of the parameter estimates against violations of model assumptions. We close with our conclusions.

The Social Welfare Survey

Employees in the Netherlands are insured under various social welfare acts against the loss of income due to redundancy, disability, or sickness. Social benefit recipients have to comply with the rules and regulations of these acts. Noncompliance with the rules is considered fraud and can have serious repercussions. In 2000, 2002, and 2004, the Dutch Department of Social Affairs conducted a nationwide survey to monitor the degree of noncompliance with respect to these rules.

We present two examples from the 2000 Social Welfare Survey. The sample consists of 1,308 persons who receive benefits within the framework of the Disability Benefit Act (DBA). The DBA offers financial benefits to employees who, due to sickness or an accident, have been unable to work for a period longer than 1 year. The amount depends on the degree of disablement, with a maximum of 70 percent of the last earned wage. To be eligible for benefits, beneficiaries are required to report all additional income from work and improvements in their health status. A detailed description of the sampling procedures used in the Social Welfare Survey can be found in Lensvelt-Mulders et al. (2006).

The examples consist of one set of three work-related questions and one set of four health-related questions. The work-related questions are

1. Have you recently done any small jobs for or via friends or acquaintances, for instance, in the past year or done any work for payments

of any size without reporting it to the Department of Social Services?
(This pertains only to monetary payments.)

2. Have you ever in the past 12 months had a job or worked for an employment agency in addition to your disability benefit without informing the Department of Social Services?
3. Have you worked off the books in the past 12 months in addition to your disability benefit?

Let the variables A^* , B^* , and C^* denote the answers to these questions for $a^*, b^*, c^* \in \{1 \equiv \text{yes}, 2 \equiv \text{no}\}$. The observed-response profile frequencies 111, 112, ..., 222 are given by $\mathbf{n}^*(66, 67, 68, 169, 52, 95, 123, 668)^t$.

The health-related questions are

4. Has a doctor or specialist ever told you that the symptoms that your disability classification is based on have decreased without you informing the Department of Social Services of this change?
5. At a Social Services checkup, have you ever acted as if you were sicker or less able to work than you actually are?
6. Have you yourself ever noticed an improvement in the symptoms causing your disability, for example, in your present job, in volunteer work, or the chores you do at home, without informing the Department of Social Services of this change?
7. For periods of any length at all, do you ever feel stronger, healthier, and able to work more hours without informing the Department of Social Services of this change?

Let the variables D^* , E^* , F^* , and G^* analogously denote the answers to Questions 4 through 7. The observed-response profile frequencies 1111, 1112, ..., 2222 are given by $\mathbf{n}^* = (43, 22, 10, 34, 20, 31, 40, 93, 30, 29, 40, 91, 60, 86, 146, 533)^t$.

The questions were all answered according to the Kuk design (Kuk 1990; van der Heijden et al. 2000). In this RR design, the respondent is given two decks with red and black playing cards. One deck contains 80 percent red cards and 20 percent black cards and is called the *yes* deck. The other deck contains 80 percent black cards and 20 percent red cards and is called the *no* deck. For each sensitive question, the respondent draws one card from both decks and answers the question by naming the color of the card from the deck corresponding to the true answer. So if the true answer is *yes*, the respondent names the color of the card from the *yes* deck, and if the true answer is *no*, the respondent names the color of the card from the *no* deck.

The General RR Model

Consider a multivariate RR design with K dichotomous sensitive questions. The true responses are denoted by the random variables $A, B, \dots$, and the random variable X denotes the $D = 2^K$ true-response profiles $A = a, B = b, \dots$. Analogously, define the variables $A^*, B^* \dots$ for the observed responses and X^* for the observed-response profiles. Let $\mathbf{P}_K$ be a $D \times D$ dimensional transition matrix, with elements (i, j) given by the conditional misclassification probabilities $p_{ij} = IP(X^* = i | X = j)$ for $i, j \in \{1, \dots, D\}$. For the univariate Kuk design, the transition matrix is given by

$$\mathbf{P}_K = \mathbf{P}_1 = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} 8/10 & 2/10 \\ 2/10 & 8/10 \end{pmatrix}. \quad (1)$$

In a multivariate design, the transition matrix $\mathbf{P}_K$ is found by taking the Kronecker product of the univariate transition matrices. For $K = 3$, the multivariate transition matrix $\mathbf{P}_3$ is found by taking the Kronecker product $\mathbf{P}_1 \otimes \mathbf{P}_1 \otimes \mathbf{P}_1$, where

$$\mathbf{P}_1 \otimes \mathbf{P}_1 = \begin{pmatrix} p_{11}\mathbf{P}_1 & p_{12}\mathbf{P}_1 \\ p_{21}\mathbf{P}_1 & p_{22}\mathbf{P}_1 \end{pmatrix},$$

is a 4×4 transition matrix.

The general RR model is given by

$$\boldsymbol{\pi}^* = \mathbf{P}_K \boldsymbol{\pi}, \quad (2)$$

where $\boldsymbol{\pi} = (\pi_1, \dots, \pi_D)^t$ is a vector denoting the true-response profile probabilities and $\boldsymbol{\pi}^* = (\pi_1^*, \dots, \pi_D^*)^t$ is a vector denoting the observed-response profile probabilities. Model (2) is estimated by maximization of the kernel of the log likelihood

$$\ln \ell(\boldsymbol{\pi} | \boldsymbol{\pi}^*, \mathbf{P}_K) = \sum_{i=1}^D n_i^* \ln \pi_i^* = \sum_{i=1}^D n_i^* \ln \left(\sum_{j=1}^D p_{ij} \pi_j \right), \quad (3)$$

for $\pi_1, \dots, \pi_D \in (0, 1)$.

Boundary Solutions, SP, and Identification

The general RR model sometimes exhibits a lack of fit. In case the general RR model lacks fit, a boundary solution is obtained that is characterized

by probability estimates on the boundary of the parameter space (van den Hout and van der Heijden 2002). A lack of fit is a somewhat unexpected result because the general RR model is a saturated model in the sense that the number of independent parameters equals the number of independent observed-response frequencies. There are two potential reasons for boundary solutions to occur.

Boundary solutions occur if a relative observed-response frequency is below (or above) chance level, with chance level defined as the probability of observing a *yes* response given a true *yes*-response probability of zero (or equivalently as the probability of observing a *no* response given a true *no*-response probability of one). We illustrate this for one variable by writing out the probability of π_1^* in (1) and (2), with subscript 1 $\equiv$ *yes*. Since in the univariate case $\pi_2 = 1 - \pi_1$, it follows that

$$\pi_1^* = p_{11}\pi_1 + p_{12}\pi_2 = 0.2 + 0.6\pi_1.$$

Solving this equation for π_1 yields the moment estimator

$$\hat{\pi}_1 = \frac{\hat{\pi}_1^* - 0.2}{0.6}, \quad (4)$$

with $\hat{\pi}_1^*$ estimated by the relative observed-response frequency n_1^*/n . If π_1^* is smaller than the chance level of .2, a negative moment estimate of π_1 is obtained. It follows that in this case π_2^* is greater than the change level of .8 and the moment estimate $\hat{\pi}_2 > 1$. Since the probability estimates obtained by maximizing log likelihood (3) are constrained to be the interval (0, 1), the model will not exhibit a perfect fit.

One potential reason for boundary solutions is RR sampling variation. By this we mean the sampling fluctuation in the frequency of red cards, given the true-response frequencies. If the number of red cards drawn in the sample is less than expected on the basis of the randomization probabilities, the percentage of observed *yes* responses might fall below chance level, especially when the frequency of the true *yes* responses is near zero. The other potential reason for a boundary solution is SP, which has a similar effect on the frequency of the observed *yes* responses as RR sampling variation. If respondents answer *no* when the answer required by the randomizing device is *yes*, the percentage of the observed *yes* responses may also be below the chance level.

In the univariate setting, the effects of SP and RR sampling variation on the observed-response frequencies are confounded. The effect of sample proportions of red cards larger than the corresponding conditional misclassification probabilities p_{11} or p_{12} described in (1) cancel out the effect

of SP, whereas smaller sample proportions reinforce the effect of SP. In a multivariate setting, the situation is more complicated because the effect of RR sampling variation on the sample proportion of red cards is different for each variable.

In this report, we define SP respondents as persons who answer *no* to every question, regardless of their true status or the outcome of the randomizing device. Given this definition, we account for SP by introducing an SP parameter θ in the general RR model, such that

$$\pi^* = (1 - \theta)P_K \pi + \theta \nu, \quad (5)$$

where θ denotes the probability of SP, and ν is the D -dimensional vector $(0, \dots, 0, 1)^t$. Notice that model (5) implies that SP can result only in the observed-response profile consisting of only *no* responses and that all true-response profiles are equally likely to be subject to SP. The model can also be rewritten as

$$\pi^* = Q_K \pi, \quad (6)$$

where the transition matrix Q_K has elements

$$q_{ij} = \begin{cases} (1 - \theta)p_{ij} & \text{for } i \neq D, j \in \{1, \dots, D\} \\ (1 - \theta)p_{ij} + \theta & \text{for } i = D, j \in \{1, \dots, D\}. \end{cases} \quad (7)$$

Model (6) is not identified. We illustrate this with the work-related questions of the Social Welfare Survey. We estimated the true-response probabilities by fitting models to the respective observed-response (profile) frequencies $\mathbf{n}^* = (309, 999)$ of variable C^* , $\mathbf{n}^* = (118, 162, 191, 873)$ of the variables B^* and C^* , and $\mathbf{n}^* = (66, 67, 68, 169, 52, 95, 123, 668)$ of the variables A^* , B^* , and C^* . The models were estimated by maximizing the kernel of the log likelihood

$$\ln \ell(\pi | \mathbf{n}^*, P_K, \theta) = \sum_{i=1}^D n_i^* \ln \left(\sum_{j=1}^D q_{ij} \pi_j \right) \quad (8)$$

for fixed values of θ in the interval $(0, 1)$. Figure 1 shows the likelihood-ratio statistic L^2 of the models as a function of the value of θ .

In case of variable C (solid line), there is a serious identification problem since the model exhibits a perfect fit for all $\theta \in (0, .72)$. The interval of θ for which the model fits perfectly is reduced to $(.2, .6)$ when the variable B is added to the model (dashed line). If the model is estimated for all three variables A , B , and C simultaneously (dotted line), the interval of θ for which a perfect fit is obtained is further reduced to $(.25, .32)$. In the

Figure 1
Likelihood-Ratio Statistic When Fitting the General
Randomized-Response Model to Uni- and Multivariate
Data With Fixed Values of θ

next section, we show how the identification problem can be overcome by using a log-linear model.

The LLRR and SP Models

The LLRR model is presented by Chen (1989) in the context of misclassification of categorical data and is further developed by van den Hout and van der Heijden (2004). In this section, we briefly review the theory of this model and then introduce the SP model.

Consider the true-response variables A, B , and C , with the true-response profiles abc , for $a, b, c \in \{1, 2\}$. For $j \in \{1, \dots, D\}$, let π_j denote the probabilities of the respective true-response profiles 111, 112, $\dots$, 222. Then the saturated LLRR model $[ABC]$ is given by

$$\pi_j = \exp(\lambda_0 + \lambda_a^A + \lambda_b^B + \lambda_c^C + \lambda_{ab}^{AB} + \lambda_{ac}^{AC} + \lambda_{bc}^{BC} + \lambda_{abc}^{ABC}), \quad (9)$$

where the λ terms are constrained to sum to zero over any subscript. The kernel of the log likelihood of the LLRR model,

$$\ln \ell(\lambda | \mathbf{n}^*, \mathbf{P}_K) = \sum_{i=1}^D n_i^* \ln \left(\sum_{j=1}^D p_{ij} \pi_j \right), \quad (10)$$

is identical to the kernel of the log likelihood (3) of the general RR model, except that log likelihood (10) is maximized as a function of the log-linear parameters. Constrained LLRR models are formulated by setting log-linear parameters in (9) to zero or by imposing equality constraints. For a more detailed discussion of the LLRR model, we refer to Chen (1989) and van den Hout and van der Heijden (2004).

The LLRR model can be adapted to accommodate SP by replacing the elements p_{ij} of transition matrix $\mathbf{P}_K$ in the log-likelihood function (10) by the elements q_{ij} of transition matrix $\mathbf{Q}_K$ defined in (7). Since the matrix $\mathbf{Q}$ contains the SP parameter θ , this results in an overparameterized model. We solve this problem by constraining the highest order interaction parameter of the log-linear model to zero. In a design with K variables, constraining the K -factor interaction parameter preserves the hierarchical structure of the model. The saturated SP model is the model θ , $[AB, AC, BC]$ that is given by

$$\pi_j = \exp(\lambda_0 + \lambda_a^A + \lambda_b^B + \lambda_c^C + \lambda_{ab}^{AB} + \lambda_{ac}^{AC} + \lambda_{bc}^{BC}), \quad (11)$$

where the term *saturated* is used in the sense that the number of free parameters in the model equals the number of independent observed-response frequencies. As with the LLRR model, constrained SP models are formulated by imposing restrictions on the log-linear parameters. The kernel of the log likelihood of the SP model is given by

$$\ln \ell(\lambda, \theta | \mathbf{n}^*, \mathbf{P}_K) = \sum_{i=1}^D n_i^* \ln \left(\sum_{j=1}^D q_{ij} \pi_j \right). \quad (12)$$

The SP model is estimated by maximizing log likelihood (12) as a function of the SP parameter θ in (7) and the log-linear parameters in (11). The estimation can be performed with standard optimization routines. A code written for the statistical program Gauss can be found on the Web site www.randomizedresponse.nl.

Examples

Table 1 presents the model selection results for the work- and health-related questions. The table reports the likelihood-ratio statistics L^2 obtained from fitting various LLRR models and SP models by maximizing the respective log likelihoods (10) and (12). The table also presents the estimates of θ for the SP models.

Table 1
Model Selection and Cheating Parameter Estimates

Data	Model	$\hat{\theta}$	L_2	df
Work	W0: $[ABC]$	—	41.6	0
	W1: $[AB, AC, BC]$	—	42.6	1
	W2: $\emptyset, [AB, AC, BC]$	.25 (.04)	0.0	0
	W3: $\emptyset, [AB, AC, BC]^a$	.26 (.04)	5.2	2
	W4: $\emptyset, [AB, BC]^b$	.25 (.04)	0.2	2
	W5: $\emptyset, [A, B, C]$	.36 (.02)	18.4	3
Health	H0: $[DEFG]$	—	37.3	0
	H1: $[DEF, DEG, DFG, EFG]$	—	38.9	1
	H2: $\emptyset, [DEF, DEG, DFG, EFG]$	.15 (.03)	7.1	0
	H3: $\emptyset, [DE, DF, DG, EF, EG, FG]$	.13 (.05)	7.1	4
	H4: $\emptyset, [DE, DF, DG, EF, EG, FG]^a$	.13 (.05)	36.2	9
	H5: $\emptyset, [DE, EF, FG]^c$	.15 (.03)	8.4	8
	H6: $\emptyset, [D, E, F, G]$	.27 (.02)	82.9	10

a. Equality constraints on all interaction parameters.

b. Equality constraints $\lambda_{ab}^{AB} = \lambda_{bc}^{BC}$.

c. Equality constraints $\lambda_{de}^{DE} = \lambda_{ef}^{EF}$.

The models W0, H0, W1, and H1 are LLRR models. The saturated LLRR models W0 and H0 both fit poorly. In the models W1 and H1, the highest order interaction parameters λ_{abc}^{ABC} and λ_{defg}^{DEFG} are constrained to zero. The slight deterioration in fit suggests that no substantial K -factor interaction is present in the data when SP is not taken into account.

The results are reported for four SP models (W2 to W5) of the work example. Model W2 fits perfectly, with an estimated SP probability of .25. The most parsimonious model is W4, with the parameters λ_{ab}^{AB} and λ_{bc}^{BC} constrained to be equal. The deterioration of fit for the models W3 with equality constraints on all interaction parameters and W5 with only main effects illustrates that no further restrictions on the parameters are feasible.

For the health example, the results are shown for five SP models (H2 to H6). Elimination of all three-factor interaction parameters in model H3 does not affect the fit. Model H5, with equality of the interaction parameters λ_{de}^{DE} and λ_{ef}^{EF} , is the most parsimonious model. In this model the estimated probability of SP is .15. Models H4 and H6 illustrate that the fit deteriorates if more constraints are imposed.

Table 2 reports the estimated odds ratios and interaction parameters. The results for the work example suggest that the status on variable A (small jobs for friends) is positively associated to the status on variable B (job or

Table 2
Estimated Two-Way Interactions

Model	Interaction	Odds Ratio	$\hat{\lambda}$
W4	$AB = BC$	26.5	.82 (.22)
H5	$DE = EF$	181.3	1.30 (.30)
	FG	29.2	.84 (.23)

Table 3
True-Response Probability Estimates

		$B = y$	$C = n$	$B = n$	$C = n$
W4		$C = y$		$C = y$	
	$A = y$	.092	.027	.017	.160
	$A = n$	.019	.006	.065	.615
		$F = y$	$G = n$	$F = n$	$G = n$
H5		$G = y$		$G = y$	
$D = y$	$E = y$	.068	.014	.001	.005
	$E = n$	.001	.000	.002	.013
$D = n$	$E = y$	.017	.004	.000	.001
	$E = n$	.038	.008	.120	.708

employment agency). The odds of having the same status on both variables are estimated to be 26 to 1. As indicated by the equality constraint $AB = BC$, the positive association between the status on the variables B and C (working off the books) is roughly equally strong. Furthermore, given the status on variable B , there is no evidence for a significant association between the variables A and C . Similar association patterns are found in the health example. The estimated odds ratios of 181 imply a high probability of the same status on the variables E (pretending to be sick at the checkup) and D (withholding the doctor's information about symptom improvements) and on the variables E and F (not reporting symptom improvements noticed by the respondent himself or herself). The results also show a positive, although somewhat less strong, association between the status on variables F and G (not reporting feeling stronger and more able to work).

The estimated true-response profile probabilities $\pi_1, \dots, \pi_D$ are shown in Table 3. The large odds ratio estimates in Table 2 turn out to be caused by probability estimates that are close to their boundary values. In the work example, the response profile nyn has an estimated probability

Table 4
Estimated Noncompliance Probabilities
and 95 Percent Confidence Intervals

Model Work	<i>A</i>	<i>B</i>	<i>C</i>	
W0	.14 (.10, .18)	.09 (.06, .13)	.09 (.07, .13)	
W4	.30 (.23, .38)	.14 (.10, .20)	.19 (.13, .27)	
Health	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>
H0	.07 (.05, .11)	.08 (.06, .12)	.11 (.09, .15)	.16 (.12, .20)
H5	.10 (.07, .17)	.11 (.08, .17)	.15 (.11, .21)	.25 (.20, .32)

smaller than .01. In the health example, more than half of the response profiles have an estimated probability smaller than .01.

Table 4 reports the univariate noncompliance estimates with corresponding confidence intervals, obtained with the parametric bootstrap method. A comparison of the results of the LLRR and SP models shows that the correction for SP has a substantial effect on the estimated noncompliance probabilities.

Robustness Against Model Violations

In this section, we evaluate the robustness of the SP model against model violations. First, we examine the robustness of the SP parameter and the univariate prevalence estimates against violations of the assumption that the K -factor interaction is zero. Second, we investigate the extent to which the SP parameter captures the effects of RR sampling variation. Last, we generate the sampling distribution of the likelihood-ratio statistic for models W4 and H5 and infer the critical value.

We evaluate the robustness of the SP parameter and univariate prevalence estimates of the models W4 and H5 against nonzero K -factor interaction λ_k^K by fitting the SP models θ , $[AB, AC, BC]$ and θ , $[DEF, DEG, DFG, EFG]$ to three manipulated data sets $\mathbf{n}_{(\lambda_k^K)}^*$ for $K \in \{3, 4\}$. The data sets are computed for different log-linear parameter vectors λ , which consist of the log-linear parameter estimates $\hat{\lambda}$ of the models W4 and H5, extended with the K -factor interaction parameter λ_k^K , for $\lambda_k^K \in \{-1, 0, 1\}$. The data sets are computed using the equations $\ln(\pi_{(\lambda_k^K)}) = M\lambda$ and $\mathbf{n}_{(\lambda_k^K)}^* = n\mathbf{Q}_K\pi_{(\lambda_k^K)}$,

Table 5
Bias in Estimated Parameters of the Saturated SP Model
as a Function of Ignored K -Factor Interaction

Model	Parameter	$\lambda_k^K = 0$	$\lambda_k^K = -1$		$\lambda_k^K = 1$	
		True = Est.	True	Est.	True	Est.
W4	θ	.249	.249	.300	.249	.235
	$\pi_1 (A)$	.295	.113	.145	.614	.597
	$\pi_1 (B)$	.142	.081	.110	.250	.239
	$\pi_1 (C)$	.192	.073	.102	.400	.386
H5	θ	.146	.146	.142	.146	.149
	$\pi_1 (D)$	.104	.135	.133	.094	.095
	$\pi_1 (E)$	.110	.151	.149	.096	.097
	$\pi_1 (F)$	.150	.189	.187	.137	.138
	$\pi_1 (G)$	.249	.539	.535	.151	.153

Note: SP = self-protection.

with $n = 1,308$. In the latter equation, the transition matrix $\mathbf{Q}_K$ is based on the estimated values $\hat{\theta} = .249$ for model W4 and $\hat{\theta} = .146$ for model H5. Since the expectation of $\mathbf{Q}_K$ is used to compute the observed-response frequencies, the data are not affected by RR sampling variation.

The results are shown in Table 5. The “True” columns refer to the parameter values used to construct the data, and the columns labeled “Est.” refer to the estimates of the saturated SP models. The upper panel of Table 5 shows that in the event of negative K -factor interaction, the SP parameter and univariate noncompliance probabilities are overestimated. The effects are reversed if the K -factor interaction is positive. In the lower panel, the effects of the K -factor interaction are opposite to those in the upper panel in both conditions and for all parameters. In comparing the true values and the estimates, the results show that given the absence of RR sampling variation, the SP model is unbiased when the K -factor interaction is zero and that otherwise the bias in the SP parameter and univariate probability estimates is relatively small.

We perform a parametric bootstrap to examine the bias in the SP parameter estimate resulting from RR sampling variation. We draw two sets of 1,000 random samples from the fitted vectors $\hat{\mathbf{n}}^*$ of the models W4 and H5, and we fit the SP models θ , $[AB, AC, BC]$ and θ , $[DEF, DEG, DFG, EFG]$ to the respective bootstrap samples. We subtract the fitted values $\hat{\theta} = .249$ of model W4 and $\hat{\theta} = .146$ of model H5 from the respective SP parameter averages in the bootstrap. Table 6 shows that the SP parameters

Table 6
Parametric Bootstrap of Models W4 and H5

Bootstrap Model	Fitted Model	Bias in $\hat{\theta}$	$L^2_{95\%}$
W4	θ , [AB, AC, BC]	.003	1.4
H5	θ , [DEF, DEG, DFG, EFG]	.008	11.1

are overestimated by .003 for model W4 and by .008 for model H5. These results suggest that the SP parameter estimate is not substantially affected by the effects of RR sampling variation.

Last, the parametric bootstraps are used to generate the distribution of the likelihood-ratio statistic for the models W4 and H5. We find an average value of 0.3 for the samples based on model W4 and of 4.4 for the samples based on model H5. The fact that these averages do not equal zero shows that the SP parameter cannot account entirely for the lack of fit resulting from RR sampling variation. It also shows that even though the SP model is correctly specified, it may not always fit perfectly. To find the rejection area of the saturated SP models, we determined the 95th percentile value $L^2_{95\%}$ of the likelihood-ratio statistic in the parametric bootstrap. These are shown for models W4 and H5 in the last column of Table 6. The likelihood-ratio statistic of 7.1 of the saturated SP model (H2) in Table 1 does not exceed the critical value of 11.1 obtained in the bootstrap. The result suggests that lack of fit is attributable to RR sampling variation and that therefore the model need not be rejected.

Conclusion

The SP model is a useful tool to analyze RR data that are potentially affected by self-protective response bias. The two applications presented in this report show that the SP model fits significantly better than models that do not take SP into account. The SP model is unbiased if the assumption of zero K -factor interaction is fulfilled and RR sampling variation is absent. Given that RR sampling variation is present, the SP parameter and univariate prevalence estimates are slightly positively biased. If K -factor interaction is present in the data, the bias in the SP parameter and univariate prevalence estimates is relatively small. Furthermore, in real data the highest order interaction parameter is usually not significant unless the sample size is large relative to the number of variables. The costs of a priori setting this parameter to zero thus seem to be low.

In this article, we restrict ourselves to the assumption that SP always results in the observed-response profile with only *no* responses, regardless of the outcome of the randomizing device or the true status of the respondent. This assumption implies that SP is independent of the true-response profile. Therefore, the prevalence estimates of the model are unbiased if SP and noncompliance are independent. However, if SP correlates positively with noncompliance, the prevalence of noncompliance is underestimated. Similarly, the SP model will overestimate the prevalence if SP correlates negatively with noncompliance. Different assumptions about SP are possible, for example, that the probability of SP depends on the true-response profile or on person characteristics. However, the new identifiability problems that arise when SP is assumed to depend on the true-response profile are beyond the scope of this report. An interesting question is to what extent SP depends on person characteristics. For example, if SP is due to a lack of trust of the RR design, improved instructions might reduce the probability of SP. The development of regression models in which the SP parameter is defined as a function of covariates is an interesting topic for future research.

If the number of variables in the RR design is large and the variables are strongly associated, the response profile data can rapidly become sparse. In this case it would be interesting to compare the SP model to an approach proposed by Gilula and Haberman (2001) that combines log-linear modeling and a summarization of the true-response profile data that is obtained after correcting the observed-response profile for RR. The methodology of Gilula and Haberman seems especially suited when the number of variables is large and SP is absent. However, it is less obvious how their methodology can be applied when SP responses are present and the probability of observing an SP response has to be estimated from the data.

The SP model is estimated by maximizing the log-likelihood function. It would also be interesting to model SP within a Bayesian framework. An advantage of the Bayesian approach is the possibility of using an informative prior for the SP parameter. In this way, knowledge of the prevalence of SP from other RR research can be taken into account. Within the Bayesian framework, it is also possible to use fully specified distributions of the SP parameter in a sensitivity analysis. If the distribution of the SP parameter is specified, there is no identification problem. By choosing different distributions, one can study the effect of these distributions on the estimated log-linear parameters and the univariate prevalence estimates of the sensitive characteristics.

References

- Böckenholt, Ulf and Peter G. M. van der Heijden. 2004. "Measuring Noncompliance in Insurance Benefit Regulations With Randomized Response Methods for Multiple Items." Pp. 106-10, in *Proceedings of the 19th International Workshop on Statistical Modelling*, edited by A. Biggeri et al. Florence, Italy: University of Florence.
- . Forthcoming. "Item Randomized-Response Models for Measuring Noncompliance: Risk-Return Perceptions, Social Influences, and Self-Protective Responses." *Psychometrika*.
- Boeije, Henny and Gerty Lensvelt-Mulders. 2002. "Honest by Chance: A Qualitative Interview Study to Clarify Respondents' (Non)-Compliance With Computer-Assisted Randomized Response." *Bulletin de Methodologie Sociologique* 75:24-39.
- Chaudhuri, Arijit and Rahul Mukerjee. 1988. *Randomized Response: Theory and Techniques*. New York: Marcel Dekker.
- Chen, Timothy T. 1989. "A Review of Methods for Misclassified Categorical Data in Epidemiology." *Statistics in Medicine* 8:1095-106.
- Clark, Stephen J. and Robert A. Desharnais. 1998. "Honest Answers to Embarrassing Questions: Detecting Cheating in the Randomized Response Model." *Psychological Methods* 3:160-8.
- Edgell, Stephen E., Samuel Himmelfarb, and Karen L. Duncan. 1982. "Validity of Forced Response in a Randomized Response Model." *Sociological Methods & Research* 11:89-110.
- Gilula, Zvi and Shelby J. Haberman. 2001. "Analysis of Categorical Response Profiles by Informative Summaries." *Sociological Methodology* 31:129-87.
- Kuk, Anthony Y. C. 1990. "Asking Sensitive Questions Indirectly." *Biometrika* 77:436-8.
- Lensvelt-Mulders, Gerty J. L. M., Joop J. Hox, Peter G. M. van der Heijden, and Cora Maas. 2005. "Meta-Analysis of Randomized Response Research, Thirty-Five Years of Validation." *Sociological Methods & Research* 33:319-48.
- Lensvelt-Mulders, Gerty, Peter G. M. van der Heijden, Olav Laudy, and Ger van Gils. 2006. "A Validation of a Computer-Assisted Randomized Response Survey to Estimate the Prevalence of Fraud in Social Security." *Journal of the Royal Statistical Association, Series A: Statistics in Society* 169:305-18.
- van den Hout, Ardo and Peter G. M. van der Heijden. 2002. "Randomized Response, Statistical Disclosure Control and Misclassification: A Review." *International Statistical Review* 70:269-88.
- . 2004. "The Analysis of Multivariate Misclassified Data With Special Attention to Randomized Response." *Sociological Methods & Research* 32:384-410.
- van der Heijden, Peter G. M., G. Van Gils, J. Bouts, and J. Hox. 2000. "A Comparison of Randomized Response, Computer-Assisted Self-Interview and Face-To-Face Direct-Questioning." *Sociological Methods & Research* 28:505-37.
- van Gils, Ger, Peter G. M. van der Heijden, and Annemarie Rosebeek. 2001. *Randomized Response: Onderzoek naar Regelovertreding* [in Dutch]. Amsterdam, Netherlands: NIPO.
- Warner, Stanley L. 1965. "Randomized Response: A Survey Technique for Eliminating Answer Bias." *Journal of the American Statistical Association* 60:63-9.

Maarten J. L. F. Cruyff has the position of researcher/teacher at Utrecht University. His research interests include randomized-response and capture-recapture models.

Ardo van den Hout works on multistate models in survival analysis at the MRC Biostatistics Unit. Before that, he worked on randomized response and statistical disclosure control problems at Utrecht University.

Peter G. M. van der Heijden is chairman of the Department of Methodology and Statistics at Utrecht University. He teaches multivariate techniques and categorical data analysis. His research interests include randomized response, capture-recapture, and categorical data analysis.

Ulf Böckenholt teaches undergraduate and graduate courses, including consumer behavior and relationship marketing, behavioral economics, decision theory, and market research. His research activities focus on quantitative aspects of marketing and psychological models of consumer behavior. He is a past editor of *Psychometrika* and a past president of the Psychometric Society.