Journal of Official Statistics
Vol 4. No. 4, 1988, pp. 375–384
© Statistics Sweden

Interviewer Variance in a Telephone Survey

Jeroen Pannekoek[1]

Abstract: This article investigates the interviewer variance component for data obtained from a computer-assisted telephone survey. A maximum likelihood procedure based on a beta-binomial model is proposed to estimate this variance component for proportions. It is found that in 9 out of 21 categories the interviewer variance component is significant. The largest values of this variance component were found for the "refusal" and "don't know" categories. Comparing the results of this study with the literature on interviewing variance, it is found that the interviewer variance component in our example is within the range of values common for telephone interviewing. In general these values seem to be smaller than the values obtained for face to face interviewing.

Key words: Interviewer variance; binary response; beta-binomial distribution; maximum likelihood.

1. Introduction

Some of the questions in a survey, like those on age and gender, are very easy to answer. In these cases respondents are asked to report information that is readily accessible and non-threatening. But answering other questions, for instance, questions concerning the respondent's opinions, income, or questions that rely on the respondent's memory may be a more difficult task for the respondent or may even cause some embarrassment. As a consequence, respondents often ask for explanations or hesitate to answer and the interviewer has to respond in some way to keep the interview going. The manner and the extent of the interviewer's response to the respondent seems to vary systematically from interviewer to interviewer (see, e.g., Cannel, Miller, and Oksenberg (1981)). The most important thing is, however, that the interviewers's behaviour may affect the respondent's answer and to the extent that different interviewers systematically have different effects on the answers, the variances of sample means and totals are increased. Variance estimators that do not take these interviewer effects into account will have a downward bias. For a long time, survey statisticians have been interested in measuring the increase in variance that can be attributed to interviewer effects.

Over the years, a great number of estimates of the magnitude of the interviewer variance component for face to face interviews have been published (see, e.g., Hanson and Marks (1958), Bailar and Dalenius (1969), Kish (1962), Fellegi (1964), Bailar, Bailey, and Stevens (1977), and Hagenaars and Heinen (1982)). More recently, results for telephone surveys have appeared in the literature (see,

[1] Department of Statistical Methods, Netherlands Central Bureau of Statistics, P.O. Box 959, 2270 AZ Voorburg, The Netherlands.

The views expressed in this paper are those of the author and do not necessarily reflect the policies of the Netherlands Central Bureau of Statistics. The author thanks the referees for their helpful comments.

e.g., Tucker (1983), Groves and Magilavy (1986)). The style of communication differs between telephone and face to face interviews and this difference may have an effect on the interviewer variance. Miller and Cannel (1982) review several findings concerning this difference. They conclude that: "There is a consistent tendency for subjects to be less confident of their judgments in no-vision conditions, and to express a preference for face to face contact." And also: "The pace of interaction ... may be faster on the telephone, leading to hurried and perhaps less thoughtful responses." Such findings suggest that for telephone interviewing subjects are more susceptible to interviewer effects. But, at the same time and for much the same reason (the limited communication possibilities), there is less leeway for different styles of interviewing. Therefore it is not clear whether to expect more or less interviewer variance in a telephone survey than in a face to face survey. Empirical results can help to settle this question.

All the investigators cited above apply a one-way random effect analysis of variance model (see, e.g., Searle (1971)) to determine the effect of the variable "interviewer." This model is appropriate for continuous response variables but problems occur when the model is applied to categorical variables (see, e.g., Stokes and Mulry (1987)). One problem is that to use the ANOVA method we must assume that the conditional variance of the response variable is the same for each interviewer. This homogeneity of variance assumption is violated if the response variable is a proportion. The variance of a proportion is a function of the mean and, except in the case that there is no interviewer variance, the mean values vary among interviewers. Another problem with the ANOVA method is that the usual confidence interval for the estimated interviewer variance component and its associated significance test are based on the assumption that both the interviewer effects and the response variable are normally distributed. These assumptions are not met if the response variable is a proportion. For these reasons Stokes and Hill (1985) and Anderson and Aitkin (1985) have proposed alternative methods for estimating the interviewer variance component for proportions. The method used by Stokes and Hill is based on the same model as is used in this paper. They do, however, use a different method for estimating the interviewer variance. Anderson and Aitkin use a different model that is computationally more difficult than the method proposed here.

In this article a model is proposed that is suitable for the analysis of interviewer variance for proportions. This model will be used to estimate the interviewer variance component from data obtained by computer-assisted telephone interviewing (CATI). The model and the consequences of interviewer effects for the sample variances of category totals and proportions are described in Sections 2 to 5. The application follows in Section 6. In Section 7 some conclusions are drawn and the results are compared with results from the literature on both CATI and face to face interviewing.

2. The Beta-binomial Model for Interviewer Variance

If there are no interviewer effects (or other non-sampling errors) to be considered, the observations in a category of a discrete variable are usually considered to come from a homogeneous population, that is to say the probability of obtaining an observation in that category (the probability of success) is the same at each draw from that population. For n mutually independent observations with the same probability of success, p, the total number of successes, F, is binomially distributed with parameters n and p and the expectation and variance of F are given by

$$\mathrm{E}(F) = np \tag{1}$$

and

$$\mathrm{var}(F) = np(1 - p). \tag{2}$$

When interviewer effects do occur, the probability of success is not constant at each draw but varies between interviewers. In this case, F is not binomially distributed. An alternative for the binomial distribution in the case of varying success probabilities is the Pólya distribution (see, e.g., Feller (1968, p. 119), Johnson and Kotz (1969, p. 231), and Paul and Plackett (1978). One way of describing the genesis of this distribution is by assuming that the parameter p of a binomial distribution is itself a stochastic variable with a beta(a,b) distribution. For this reason the Pólya distribution is also known as the beta-binomial distribution. The choice of the beta distribution for modeling variation in p is to a large extent arbitrary. Other distributions on the [0,1] interval could be used as well. The advantage of the beta distribution is that an explicit expression for the marginal distribution of F is available in the form of the beta-binomial distribution. For this reason the choice of the beta distribution is common to almost all applications where models are used to describe binomial distributions with varying probabilities of success. In terms of interviewer effects, this beta-binomial model entails that the probability of success for a randomly chosen interviewer i is a stochastic variable, P_i. If the number of successes of interviewer i is F_i then, according to this model, we have

$$F_i \mid P_i = p_i \sim \mathrm{binomial}(n_i, p_i)$$

and

$$P_i \sim \mathrm{beta(a,b)}\ , \tag{3}$$

where n_i is the number of respondents of interviewer i. The expectation and variance of P_i are

$$\mathrm{E}(P_i) = a / (a + b) = \pi \tag{4}$$

and

$$\mathrm{var}(P_i) = \pi(1 - \pi) / (a + b + 1). \tag{5}$$

For a given value of π, the variance of P_i approaches its maximum value of $\pi(1 - \pi)$ when both the parameters a and b approach their limiting value of zero.

Formula (3) gives the distribution of P_i and the conditional distribution of F_i given a realization p_i of P_i. The simultaneous distribution of F_i and P_i is the product of these two distributions. Our interest is, however, in the marginal distribution of F_i and this distribution is obtained by integrating the simultaneous distribution over P_i. The distribution thus obtained is the beta-binomial distribution mentioned above. This distribution can be written in terms of the parameters n_i, a, and b but also in terms of the parameters n_i, π, and $\alpha = 1 / (a + b)$ (see Johnson and Kotz (1969, pp. 79 and 230)). In this last parameterization the beta-binomial distribution can be written as

$$\Pr(F_i = f_i) = \frac{\binom{n_i}{f_i} \prod_{j=0}^{f_i-1} (\pi + j\alpha) \prod_{j=0}^{n_i-f_i-1} (1 - \pi + j\alpha)}{\prod_{j=0}^{n_i-1} (1 + j\alpha)} \tag{6}$$

with expectation and variance

$$\mathrm{E}(F_i) = n_i\pi \tag{7}$$

and

$$\mathrm{var}(F_i) = n_i\pi(1 - \pi)\left(1 + \frac{\alpha(n_i - 1)}{\alpha + 1}\right). \tag{8}$$

The expectation of F_i equals the expectation of a binomial (n_i, π) variable. The variance of F_i, however, exceeds the variance of this binomial variable by a factor $1 + \alpha(n_i - 1) / (\alpha + 1)$. This heterogeneity or overdispersion factor (with respect to the binomial model) approaches its maximum value of n_i as α approaches infinity. In the limit for $\alpha \rightarrow 0$, the beta-binomial distribution approaches the binomial distribution and the heterogeneity factor approaches 1.

3. The Intra-class Correlation and the Effect of Interviewer Variance

A standard measure of the magnitude of a variance component is the intra-class correlation coefficient. Especially in research on interviewer variance, the results are almost always expressed in terms of this coefficient. The reason is that the intra-class correlation coefficient expresses the interviewer variance component as a proportion of the total variance and this makes it easier to compare the magnitude of the interviewer variance for different questions. In this section the relation between the intra-class correlation coefficient and the parameters of the beta-binomial model will be shown. Also, the relation between the interviewer variance and the variance of a proportion will be expressed in terms of the intra-class correlation coefficient. The intra-class correlation coefficient is defined as (see, e.g., Winer (1971, p. 184))

$$\varrho_I = \frac{\operatorname{cov}(X_{ij}, X_{ij'})}{\sqrt{\operatorname{var}(X_{ij})\operatorname{var}(X_{ij'})}} = \frac{\operatorname{cov}(X_{ij}, X_{ij'})}{\operatorname{var}(X_{ij})}, \tag{9}$$

where X_{ij} and $X_{ij'}$ are two observations on respondents j and j' ($j \neq j'$) in class i of the random factor (here: two respondents with the same interviewer i).

In the case of the beta-binomial model, ϱ_I is a function of the parameter α. To see this we consider first for every interviewer the covariance between the n_i binary observations, X_{ij} $(i = 1, \ldots, I,\ j = 1, \ldots, n_i)$

$$\begin{aligned}\operatorname{cov}(X_{ij}X_{ij'}) &= \mathrm{E}_{P_i} \operatorname{cov}(X_{ij}X_{ij'} \mid P_i) \\ &\quad + \operatorname{cov}_{P_i}(\mathrm{E}X_{ij} \mid P_i,\ \mathrm{E}X_{ij'} \mid P_i) \\ &= 0 + \operatorname{cov}_{P_i}(P_i, P_i) \\ &= \operatorname{var}(P_i) = \frac{\pi(1-\pi)}{a+b+1},\end{aligned} \tag{10}$$

and the variance of X_{ij}

$$\begin{aligned}\operatorname{var}(X_{ij}) &= \mathrm{E}_{P_i} \operatorname{var}(X_{ij} \mid P_i) + \operatorname{var}_{P_i}(\mathrm{E}X_{ij} \mid P_i) \\ &= \mathrm{E}\,\{P_i(1-P_i)\} + \operatorname{var}(P_i) \\ &= \mathrm{E}P_i - \mathrm{E}P_i^2 + \mathrm{E}P_i^2 - (\mathrm{E}P_i)^2 \\ &= \pi(1-\pi).\end{aligned} \tag{11}$$

The intra-class correlation coefficient for the binary variables X_{ij} can now be written as

$$\varrho_I = \frac{1}{a+b+1} = \frac{\alpha}{1+\alpha}. \tag{12}$$

Thus ϱ_I is a simple function of α.

The variance of the cell proportion $P_+ = F_+ / n_+$ with $F_+ = \sum_{i=1}^{I} F_i$ and n_+ the total number of observations in the sample can now (using (8) and (12)) be written as

$$\operatorname{var}(P_+) = \frac{1}{n_+^2} \sum_{i=1}^{I} n_i \pi(1-\pi)(1+(n_i-1)\varrho_I), \tag{13}$$

and for n interviews for each interviewer this simplifies to

$$\begin{aligned}\operatorname{var}(P_+) &= \frac{1}{In}\pi(1-\pi)\ (1+(n-1)\varrho_I) \\ &= \frac{\pi(1-\pi)}{In} + \frac{(n-1)\pi(1-\pi)}{In}\varrho_I.\end{aligned} \tag{14}$$

In (14) the variance of the sample proportion is expressed as the sum of two variance components: one component due to the variance between respondents and a second component due to the variance between interviewers. As n approaches infinity the first variance component approaches zero, but the second variance component approaches $\pi(1-\pi)\varrho_I/I$. The second variance component decreases as the number of interviewers (I) increases and as ϱ_I decreases and vanishes when the number of interviewers reaches its maximum n_+ or if $\varrho_I = 0$. The variance of a proportion can thus be decreased by increasing the sample size *and* by increasing the number of interviewers. Even small values of ϱ_I can have substantial effects on the variance of a proportion. For instance, if $\varrho = 0.02$ and $n = 50$, then the variance of a proportion will be about two times larger than when there are no interviewer effects. If the costs of hiring and training an interviewer and of completing one interview are known, and if ϱ_I is known or can be estimated, it is possible to calculate, for a given budget, the workload per interviewer that minimizes the variance. In this respect, Formula (14) is helpful when making decisions about the planning of the field work.

If we in (13) or (14) substitute the population variance σ^2 of a continuous variable for $\pi(1-\pi)$ we obtain analogous expressions for the variance of the sample mean of a continuous variable. The formulas thus obtained are the same as the variance formulas for cluster samples (see, e.g., Cochran (1963)). Kish (1962) uses the continuous variable analogue of (14) as an approximation in the case of unequal numbers of interviews per interviewer, where the average number of interviews per interviewer is used instead of n.

4. Maximum Likelihood Estimators

Assuming independent beta-binomial distributions for the I observed frequencies F_i, the likelihood function (L) follows from (6). The maximum likelihood estimators of α and π are the solutions of the likelihood equations given by

$$\partial L/\partial\alpha = \sum_{i=1}^{I} \left\{ \sum_{j=0}^{f_i-1} j(\pi + j\alpha)^{-1} + \sum_{j=0}^{n_i-f_i-1} j(1-\pi+j\alpha)^{-1} - \sum_{j=0}^{n_i-1} j(1+j\alpha)^{-1} \right\} = 0 \tag{15}$$

and

$$\partial L/\partial\pi = \sum_{i=1}^{I} \left\{ \sum_{j=0}^{f_i-1} (\pi + j\alpha)^{-1} - \sum_{j=0}^{n_i-f_i-1} (1-\pi+j\alpha)^{-1} \right\} = 0. \tag{16}$$

The solution to these equations must be determined iteratively. For the application in this article we used the Newton-Raphson algorithm. The covariance matrix of the ML estimators for α and π, $\hat{\alpha}$ and $\hat{\pi}$, can be estimated by the observed information matrix, I, with elements given by

$$-\left.\frac{\partial^2 L}{\partial\alpha^2}\right|_{\substack{\pi=\hat{\pi}\\ \alpha=\hat{\alpha}}} = \sum_{i=1}^{I} \left\{ \sum_{j=0}^{f_i-1} j^2(\hat{\pi} + j\hat{\alpha})^{-2} - \sum_{j=0}^{n_i-f_i-1} j^2(1-\hat{\pi}+j\hat{\alpha})^{-2} + \sum_{j=0}^{n_i-1} j^2(1+j\hat{\alpha})^{-2} \right\}, \tag{17}$$

$$-\left.\frac{\partial^2 L}{\partial\pi^2}\right|_{\substack{\pi=\hat{\pi}\\ \alpha=\hat{\alpha}}} = \sum_{i=1}^{I} \left\{ \sum_{j=0}^{f_i-1} (\hat{\pi} + j\hat{\alpha})^{-2} - \sum_{j=0}^{n_i-f_i-1} (1-\hat{\pi}+j\hat{\alpha})^{-2} \right\}, \tag{18}$$

and

$$-\frac{\partial^2 L}{\partial\alpha\partial\pi}\bigg|_{\substack{\pi=\hat{\pi}\\ \alpha=\hat{\alpha}}} = \sum_{i=1}^{I}\left\{-\sum_{j=0}^{f_i-1} j(\hat{\pi}+j\hat{\alpha})^{-2} + \sum_{j=0}^{n_i-f_i-1} j(1-\hat{\pi}+j\hat{\alpha})^{-2}\right\}. \quad (19)$$

The ML estimator $\hat{\varrho}_I$ of ϱ_I is obtained by substituting $\hat{\alpha}$ for α in (12). And a linear approximation for the variance of $\hat{\varrho}_I = f(\hat{\alpha})$ is given by

$$\begin{aligned}\text{var}(\hat{\varrho}_I) &= \left(\frac{df(\alpha)}{d\alpha}\bigg|_{\alpha=\hat{\alpha}}\right)^2 \text{var}(\hat{\alpha}) \\ &= (1+\hat{\alpha})^{-4}\,(I)^{-1}_{11}.\end{aligned} \quad (20)$$

5. Test Statistics

Before we draw conclusions on the basis of the estimated parameters of the beta-binomial model, we want to know how well our model fits the data. That is, we want to test the hypothesis

$$H_{0(1)} : F_i \sim B - B(n_i, \pi, \alpha) \quad \forall\, i, \quad (21)$$

against the general alternative, $H_{A(1)}$, that $H_{0(1)}$ is not true. As a test statistic for this "goodness-of-fit" problem we can use a generalization of Pearson's χ^2-statistic, the Wald-statistic, W^2. The statistic W^2 is given by

$$W^2 = (\mathbf{F} - \hat{\mathbf{F}})'\hat{\mathbf{V}}^{-1}(\mathbf{F} - \hat{\mathbf{F}}), \quad (22)$$

where $\mathbf{F}$ is the vector with observed frequencies, $\hat{\mathbf{F}}$ the vector with the estimated frequencies under $H_{0(1)}$ and $\hat{\mathbf{V}}$ the estimated covariance matrix of $\hat{\mathbf{F}}$. Under $H_{0(1)}$ the frequencies F_i are mutually independent and beta-binomially distributed with the same α and π but with different n_i. Therefore the matrix $\hat{\mathbf{V}}$ is given by

$$\hat{\mathbf{V}} = \text{diag}\,\{n_i\hat{\pi}(1-\hat{\pi})\,(1+(n_i-1)\,\hat{\varrho}_I)\}. \quad (23)$$

If $H_{0(1)}$ cannot be rejected and we accept the beta-binomial model, we can test if the interviewer variance component is significantly greater than zero. That is to say we want to test the hypothesis

$$H_{0(2)} : \alpha = 0,$$

against the alternative

$$H_{A(2)} : \alpha > 0.$$

The distribution of the estimated frequencies under the hypothesis $H_{0(2)}$ is binomial and under $H_{A(2)}$ this distribution is beta-binomial. Unfortunately, the likelihood-ratio statistic for testing $H_{0(2)}$ against $H_{A(2)}$ (i.e., twice the difference between the maximized beta-binomial and binomial log-likelihoods) is not suitable in this case. The problem is that since $\alpha > 0$, the value of α specified under $H_{0(2)}$ lies on the boundary of the parameter set. In such cases the usual asymptotic χ^2 distribution of the likelihood-ratio statistic does not apply. However, Tarone (1979) derived a one-sided test of the binomial distribution against beta-binomial alternatives based on the asymptotically standard normal distributed statistic given by

$$Z = \frac{(pq)^{-1}\sum_{i=1}^{I}(f_i - n_i p)^2 - \sum_{i=1}^{I} n_i}{\left\{2\sum_{i=1}^{I} n_i(n_i-1)\right\}^{1/2}}, \quad (24)$$

where p is the observed cell proportion and $q = 1 - p$ (see Prentice (1986)).

6. Application

In this section the beta-binomial model will be used to estimate the interviewer variance component in the December 1986 Netherlands Consumer Survey. The data were collected by 36 interviewers using CATI. Together they completed 1 400 interviews which amounts to an average interviewer workload of 39 interviews. The model was fitted for each category of the following three variables: net family income, with categories 1 (the lowest income category) to 7 (the highest income category), 8 (refusal) and 9 (don't know); the question "what is your opinion on the general economical situation? Has it gotten better or worse or remained the same in the last 12 months in the Netherlands" with categories 1 (clearly better), 2 (somewhat better), 3 (remained the same), 4 (somewhat worse), 5 (clearly worse), and 6 (don't know); and the question on the respondents' opinions on what will happen to the general economical situation in the next 12 months, with the same six categories as in the question concerning the past 12 months. The results of the goodness-of-fit test for the beta-binomial model are displayed in Table 1.

The results in Table 1 show an excellent fit of the model for these data and therefore we can use this model to make inferences about the interviewer variance component. First, the hypothesis $H_{0(2)}$:$\alpha = 0$, or equivalently $\varrho_I = 0$, was tested. The results of the likelihood-ratio test are displayed in Table 2.

It appears that for 9 out of the 21 categories examined, $\hat{\varrho}_I$ is significantly greater than zero (at the 5 % significance level). Significant values of ϱ_I seem to occur more often with special

Table 1. Goodness of fit test for the beta-binomial distribution ($H_{0(1)}$)

Income			Opinion on economical situation				
				Opinion last 12 months		Opinion next 12 months	
Category	W^2	p	Category	W^2	p	W^2	p
1. low	38.0	0.29	1. clearly better	41.5	0.18	37.2	0.32
2.	35.6	0.39	2. somewhat better	35.7	0.39	33.1	0.51
3.	37.0	0.33	3. the same	38.0	0.29	36.7	0.35
4.	22.4	0.94	4. somewhat worse	33.2	0.51	34.7	0.43
5.	36.9	0.34	5. clearly worse	35.2	0.41	36.9	0.34
6.	35.4	0.40	6. don't know	36.7	0.35	36.2	0.37
7. high	32.5	0.54					
8. refusal	31.1	0.61					
9. don't know	36.0	0.37					

Table 2. Test of the hypothesis $H_{0(2)}$: $\varrho_I = 0$

Income			Opinion on economical situation				
				Opinion last 12 months		Opinion next 12 months	
Category	Z	p	Category	Z	p	Z	p
1. low	2.8	<0.01	1. clearly better	−0.2	0.58	1.7	0.05
2.	1.5	0.07	2. somewhat better	1.6	0.06	−0.5	0.70
3.	2.7	<0.01	3. the same	−0.6	0.72	1.6	0.06
4.	−1.2	0.89	4. somewhat worse	3.2	<0.01	1.3	0.10
5.	1.6	0.06	5. clearly worse	2.5	0.01	0.8	0.20
6.	2.0	0.02	6. don't know	−0.4	0.67	3.7	<0.01
7. high	−0.8	0.79					
8. refusal	3.0	<0.01					
9. don't know	2.6	<0.01					

Table 3. Estimated ϱ_I-values and standard errors

Income			Opinion on economical situation				
				Opinion last 12 months		Opinion next 12 months	
Category	$\hat{\varrho}_I$		Category	$\hat{\varrho}_I$		$\hat{\varrho}_I$	
1. low	0.015	(0.010)	1. clearly better	<0.001	(<0.001)	0.010	(0.009)
2.	0.010	(0.010)	2. somewhat better	0.008	(0.008)	<0.001	(<0.001)
3.	0.023	(0.013)	3. the same	<0.001	(<0.001)	0.011	(0.009)
4.	<0.001	(<0.001)	4. somewhat worse	0.014	(0.009)	0.007	(0.007)
5.	0.010	(0.009)	5. clearly worse	0.014	(0.010)	0.005	(0.007)
6.	0.011	(0.009)	6. don't know	<0.001	(<0.001)	0.023	(0.012)
7. high	<0.001	(<0.001)					
8. refusal	0.022	(0.013)					
9. don't know	0.017	(0.010)					

kinds of categories: three of the four "don't know" or "refusal" categories have significant ϱ_I values.

In Table 3 the estimated ϱ_I-values and standard errors are displayed.

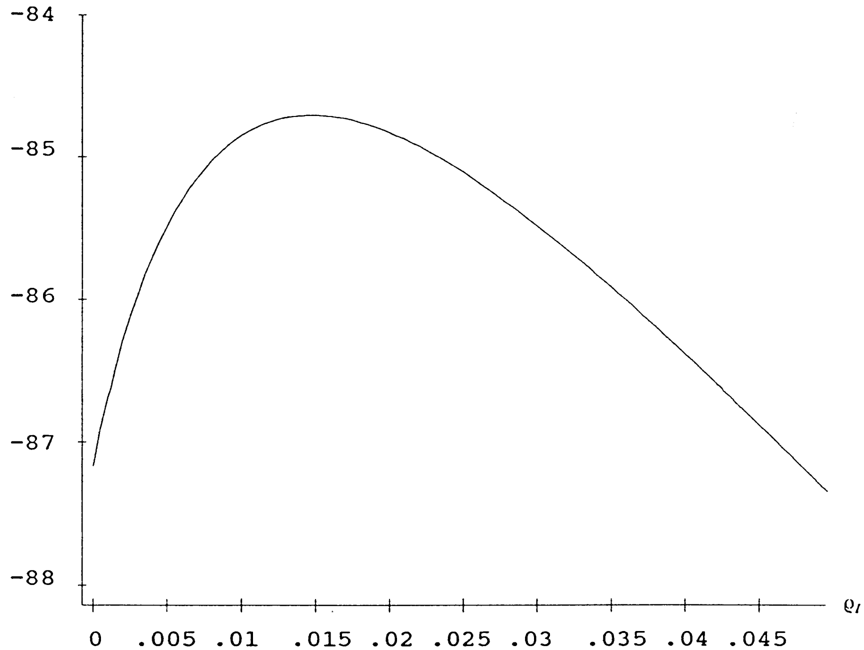
The estimated ϱ_I values exhibit large variability over the different categories examined. The mean value over all categories is 0.009 with a range of zero to 0.023. Two of the four "don't know" and "refusal" categories have values greater than 0.02 but only for one of the 17 substantive categories a value greater than 0.02 is found.

There are problems with the interpretation of the large values of the estimated standard errors in Table 3. If these standard errors are used to construct 95 % confidence intervals based on the limiting normal distribution of $\hat{\varrho}_I$, then the conclusion is that none of the $\hat{\varrho}_I$ values is significantly different from zero. This contradicts the results in Table 2. A plausible explanation is that the distribution of $\hat{\varrho}_I$ is not well approximated by a normal distribution. Evidence that this is indeed the case can be obtained from Figure 1, where, for income category 1, the log-likelihood is plotted as a function of ϱ_I (for each value of ϱ_I the likelihood is maximized over π). This profile log-likelihood function is negatively skewed, whereas it would be symmetric if $\hat{\varrho}_I$ was approximately normally distributed. Plots of the profile log-likelihood function for the other significant values of $\hat{\varrho}_I$ (based on the results in Table 2) were very similar in shape. In such cases, the normally based symmetric confidence intervals can be seriously misleading (see, e.g., Cox and Hinkley (1974, p. 343)).

7. Conclusions

In this study on interviewer variance in a telephone survey some evidence was found for the occurrence of interviewer effects. The results from this study are in concordance with results found in the literature. For instance, Groves and Magilavy (1986) report a mean ϱ_I value of 0.009 for nine telephone surveys, exactly the same mean value obtained for the three questions examined in this article. Other results from the literature indicate that for face to face interviews the mean ϱ_I values per survey range from 0.00 to 0.04 with an overall mean value of about 0.02 (see Hagenaars and Heinen (1982) and Groves and Magilavy (1986)). These results may be taken as an indication that the use of CATI reduces the interviewer variance compared to face to face interviewing. However, we have to reckon with a great variability of ϱ_I values between variables and between categories and only carefully designed experiments can answer the question whether or not CATI reduces interviewer variance. In our study, the ϱ_I

Fig. 1. Profile log-likelihood function for income category 1



values obtained for the "don't know" and "refusal" categories were more often significant and also larger than the values obtained for substantive categories. Although for these categories the effect of interviewer variance on the variance of a proportion is not of much interest. It is, however, worth noting that interviewers differ significantly in their ability to obtain a response and that our estimator is effective in picking up this expected variability.

Although for most of the categories examined in this study no significant value of ϱ_I was found, for some of the categories significant values in the range from about 0.01 to 0.02 were found. The effect of the interviewer variance on the variance of an estimated proportion can be calculated using (14). For $n = 39$, $\varrho_I = 0.01$ and $I = 36$ it follows that the variance of a proportion is 1.38 times larger than the expected variance in the case where no interviewer variance occurred. The usual variance estimator does not take interviewer variance into account and will therefore have a downward bias of a factor of 1.38. Another way of expressing this result is by using the effective sample size, that is, the sample size that results in the same variance if no interviewer effects had occurred. For our sample with 1 400 cases the effective sample size is 1 015 for ϱ_I=0.01 and only 796 for $\varrho_I = 0.02$.

Formula (14) can also be of use in planning the field work. Suppose, for instance, that if no interviewer variance occurs, we need precision corresponding to a sample of 1 000 cases. That is to say that if we do have to reckon with interviewer variance, we need an *effective* sample size of 1 000. If ϱ_I=0.02 and we have 36 interviewers, we need an interviewer workload of 61, which results in a total sample size as large as 2 196. But this is not the only way of obtaining the effective sample size of 1 000. We can also use 45 interviewers with a workload of 39, which amounts to a total sample size of 1 755. If the costs per interviewer and per completed interview are known, this kind of calculation can be used to obtain the optimal combination of number of interviewers and number of interviews.

8. References

Anderson, D.A. and Aitkin, M (1985): Variance Component Models With Binary Response: Interviewer Variability. Journal of the Royal Statistical Society, B, 47, pp. 203–210.

Bailar, B. and Dalenius, T. (1969): Estimating the Response Variance Components of the U.S. Bureau of the Census Survey Model. Shankhya, B, 33, pp. 341–360.

Bailar, B., Bailey, L., and Stevens, J. (1977): Measures of Interviewer Bias and Variance. Journal of Marketing Research, 14, pp. 337–343.

Cannell, C.F., Miller, P.V., and Oksenberg, L. (1981): Research on Interviewing Techniques. In: Leinhardt, S. (ed.), Sociological Methodology 1981. Jossey-Bass, San Francisco.

Cochran, W.G. (1963): Sampling Techniques. Second edition, Wiley, New York.

Cox, D.R. and Hinkley, D.V. (1974): Theoretical Statistics. Chapman and Hall, London.

Fellegi, I.P. (1964): Response Variance and Its Estimation. Journal of the American Statistical Association, 59, pp. 1016–1041.

Feller, W. (1968): An Introduction to Probability Theory and Its Applications. Volume 1, third edition. Wiley, New York.

Groves, R.M. and Magilavy, L.J. (1986): Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys. Public Opinion Quarterly, 55, pp. 251–266.

Hagenaars, J.A. and Heinen, T.G. (1982): Role-independent Interviewer Characteristics. In: Dijkstra W. and van der Zouwen, J. (eds.), Response Behaviour in the Survey-interview, pp. 91–127. Academic Press, London.

Hanson, R.H. and Marks, E.S. (1958): Influence of the Interviewer on the Accuracy of Survey Results. Journal of the American Statistical Association, 53, pp. 635–655.

Johnson, N.L. and Kotz, S. (1969): Discrete Distributions. Wiley, New York.

Kish, L. (1962): Studies of Interviewer Variance for Attitudinal Variables. Journal of the American Statistical Association, 57, pp. 92–115.

Miller, V.P. and Cannell, C.F. (1982): A Study of Experimental Techniques for Telephone Interviewing. Public Opinion Quarterly, 46, pp. 250–269.

Paul, S.R. and Plackett, R.L. (1978): Inference Sensitivity for Poisson Mixtures. Biometrika, 65, pp. 591–602.

Prentice, R.L. (1986): Binary Regression Using an Extended Beta-binomial Distribution, With Discussion of Correlation Induced by Covariate Measurement Errors. Journal of the American Statistical Association, 81, pp. 321–327.

Searle, S.R. (1971): Linear models. Wiley and Sons, New York.

Stokes, S.L. and Hill, J.R. (1985): Modeling Interviewer Variability for Dichotomous Variables. American Statistical Association, Proceedings of the Section on Survey Research Methods, pp. 344–348.

Stokes, S.L. and Mulry, M.H. (1987): On the Design of Interpenetration Experiments for Categorical Data Items. Journal of Official Statistics, 3, pp. 389–401.

Tarone, R.E. (1979): Testing the Goodness of Fit of the Binomial Distribution. Biometrika, 66, pp. 585–590.

Tucker, C. (1983): Interviewer Effects in Telephone Surveys. Public Opinion Quarterly, 47, pp. 84–95.

Winer, B.J. (1971): Statistical Principles in Experimental Design. McGraw-Hill, London.

Received July 1987
Revised October 1988