Developments in Survey Methodology
Anuška Ferligoj and Andrej Mrvar (Editors)
Metodološki zvezki, 15, Ljubljana: FDV, 2000

Balancing MSE and Number of Contacts

Vasja Vehovar and Katja Lozar Manfreda[1]

Abstract

This paper addresses the balance between costs and errors in telephone surveys. Two alternatives are compared: a larger sample with less follow-up calls and a smaller sample with more follow-up calls. The comparison involves a detailed elaboration of mean squared errors and cost functions. Based on the model, the key variables are discussed. The empirical example refers to the percentage of companies with access to the Internet. It is shown that the above elaboration can be helpful for practical decisions.

1 Introduction

In survey research, we often discuss various procedures for improving the quality of data but rather rarely do we discuss survey costs. However, when quality improvement efforts are discussed in such an isolated form a heavy mismatch between theory and practice may occur (Groves, 1989: vi, vii).

In this paper, we pose the following practical question: What is the optimum balance between the errors and the costs of a telephone survey? A similar approach was used by the authors previously in the case of a mail survey (Vehovar and Lozar, 1998).

We use the general understanding of survey errors and survey costs (Kish, 1965; Groves, 1989). Specifically, we narrow our analysis down to the issues of sample size and number of contacts. Both parameters, large initial sample size and large number of follow-ups, lead to smaller error, but at the same time, they also produce higher costs. We search for the precise balance between survey costs and errors.

Empirically, we present a case of a telephone survey with several follow-up calls. As an example, we discuss whether or not to use the second follow-up, however the same reflection could be made for any of the follow-up contacts. In our example, the dilemma can be expressed with the following question: Is it

[1] Faculty of Social Sciences. University of Ljubljana. Kardeljeva ploščad 5. 1000 Ljubljana, Slovenia.

better to have smaller initial sample with two follow-ups, or larger initial sample with only one follow-up? We explore the following two factors: the expected nonresponse conversion rate for the second follow-up and the relative bias after the first follow-up.

2 Errors and costs

2.1 Errors

The components of survey errors have already been well-elaborated (Groves 1989). In this paper we limit ourselves to the sampling error and to the component that belongs to the nonresponse bias. The sampling error can be regulated with an increase/decrease of the sample size and the nonresponse error can be, at least in this context, reduced with additional contacts.

We use the following expression of the estimate of the root mean square error (rmse), a measure of the total survey error (Groves, 1989: 8) which takes into account the different number of follow-up contacts and final sample size as the function of the completion rate at the *i*-th follow-up and initial sample size (Vehovar and Lozar, 1998: 140-141):

$$rmse(p)=\sqrt{\frac{1}{n^{*}\sum_{i=0}^{v}CR_i}\left[\frac{\sum_{i=0}^{v}p_iCR_i}{\sum_{i=0}^{v}CR_i}\left(1-\frac{\sum_{i=0}^{v}p_iCR_i}{\sum_{i=0}^{v}CR_i}\right)\right]+\left[P-\frac{\sum_{i=0}^{v}p_iCR_i}{\sum_{i=0}^{v}CR_i}\right]^2} \qquad (1)$$

This is an expression in the case of v follow-ups. $v = 0$ for the first contact since there are no follow-ups. Here we have n^* as initial sample size, P as the population percentage, p_i as the estimate at the i-th follow-up and CR_i as the completion rate at the i-th follow-up, where $i=0, \ldots, v$.

RMSE is thus a function of the population percentage P, nonresponse bias, completion rates, initial sample size n^* and number of follow-ups v. We have - at least in this context – no influence on the value of the population percentage P nor can we regulate the sample estimate p (i.e. the bias). We also have no influence on the completion rate at the i-th follow-up (CR_i). On the other hand, we can regulate the initial sample size n^* and the number of follow-ups v. In the expressions above we have assumed that all the previous nonrespondents have been included in each additional follow-up.

2.2 Costs

As we have already mentioned, the changes in the sample size and the number of follow-ups influence the survey costs. The cost function in our example is thus not a continuous one as in the case of one contact. Simplified it can be written:

$$T = K + C_v = K + (K_v + C_{vn}) = K + (v+1)A + n^* \sum_{i=0}^{v+1} (1 - CC_i) * c_i \ , \qquad (2)$$

where the total costs (T) of the survey consist of constant costs (K) and variable costs (C_v). The constant costs include: design of the survey, construction of a questionnaire, data analysis, printing as well as circulation of survey reports, overhead costs etc. The variable costs have two components:

- the costs K_v that vary with the number of follow-ups (but not with the initial sample size). These are costs for data management and preparation before the telephone calls are made. We assumed that these costs are the same for each follow-up, therefore we can express K_v as the product of number of phases in the survey research process and a certain level of these costs $((v+1)A)$. The data management costs partially depend also on sample size what is included in the second component of the variable costs.
- the variable costs C_{vn} that are proportional to the initial sample size n^* and to the variable costs at the i-th follow-up. These costs are the function of the cumulative contact rate (CC_i) and c_i - the costs per unit within each follow-up[2]. The contact rate (CC_i) in this case means the proportion of the sample for which additional contacts are not needed anymore (because they have already responded or are noneligible). In the case of telephone surveys this costs include costs for calls where interviewer starts to speak with the household or organisation. In a simplified case these costs are the same for all calls[3]. Of course, $CC_0=0$ since all units are called at least once. The last component ($i=v+1$) does not refer to any follow-up but includes only the costs of telephone interviews (data entry). Therefore, CC_i takes value $CC_{v+1} =(1-n/n^*)$, so that $n^*(1-CC_{v+1})$ gives the number of all responding units denoted as n.

2.3 Optimisation

The aim of the optimal design may be stated in two alternative ways: achieving minimum MSE for fixed costs, or achieving minimum costs for fixed MSE. Both principles would generally lead to the same solution (Kish, 1965: 263-264). However, unlike with the standard sampling theory, it is difficult to find the

[2] At this we have in mind the initial sample since survey costs comprise costs for respondents and nonrespondents.

[3] In practice they depend on the time spent for a call but in a simplified model we can use an average time and therefore average costs for a call.

analytical solution when the variable to optimise is a discrete one – the number of follow-ups *v*.

In the case of the above two equations (costs, RMSE) we can calculate costs and RMSE for each value of *v* and then compare the values. We will concentrate on the optimisation of the RMSE for fixed costs. It is possible to increase the sampling error (with a decrease of n^*, initial sample size) and simultaneously reduce the nonresponse error (with an increase of *v*, number of follow-ups), or the other way around, but the total costs must remain the same. Typically, we can have a large initial sample and a small number of follow-ups, or the opposite, a small initial sample and a large number of follow-ups.

3 Example

We present an example of a telephone survey on the use of Internet among Slovenian companies which was conducted in 1998 as a part of a larger research project *Research on Internet in Slovenia* at the Faculty of Social Sciences, University of Ljubljana (http://www.ris.org). The key population parameter was a percentage of companies with the access to the Internet. Since the true population value was unknown we have assumed that the true value was the value achieved after the last follow-up. The bias after the last follow-up was therefore automatically set to zero.

One of the problems associated with such assumption regarding the true value is that with a small sample for the last follow-up, the last follow-up almost has no weight and no extra value. Other methods of establishing the true value would be extrapolating the changes after each follow-up or using an external information source. Nevertheless, we have decided for the value achieved after the last follow-up since this has a very practical meaning: by comparing the value before the last follow-up to the value after the last follow-up we actually assess how much worse the variable estimate would be if any of the follow-ups was omitted.

The RIS98 telephone survey was done using CATI in a centralised telephone facility. Interviewers tried to contact the person responsible for Internet or information technology within the company. Only the calls where the interviewer actually spoke to someone in the company are included in the following calculations. Therefore one call is defined as a contact where someone in the company answered the phone. Several calls (up to 12) were made in order to contact the right persons and conduct the interviews.

1,120 companies were included in the initial sample. The final response rate was 70.0% and completion rate was 61.4%.

We can observe that after all eleven follow-up calls 59.7% of companies had access to Internet. This estimated percentage is slightly changing after different number of follow-up calls. The highest estimate is achieved after the fifth follow-up call: with five follow-up calls the estimated percentage would be 61.2%. The estimated percentage actually increases with additional follow-up calls until the fifth one, but then starts to decrease slightly. However, none of the two consecutive pairs of estimates differ significantly (at α=0.005).

All designs in Table 1 assume the same fixed budged. For the budget needed for eleven follow-up calls and an initial sample of 1,120 units, we could for example omit the last three follow-up calls and enlarge the initial sample size to 1,460. If we did only three follow-up calls, we could have an initial sample of 2,451 companies. In case of no follow-up we could have a sample size of 7,737. In all these cases the sample variance vary considerable, however the bias stays more or less the same. Therefore the optimum for this fixed budget is achieved in the case of only one follow-up (the smallest RMSE).

Table 1: Optimum design in a telephone survey (costs are expressed in Slovenian tolars).

No. of fol-low-up calls	Estimate p of P	Costs for a unit of initial sample (phone call)	Costs for completed interview (data entry)	% of total co-sts	Cumula-tive overall costs	Cumula-tive compli-tion rate	Initial sample size	Final sample size (res-pon-dents)	Rmse(p) – estimate of RMSE
0	59.0%	30	1000	17.5	17.5	16.8	7737	1299	0.015316
1	59.5%	30	1000	15.2	32.7	30.9	4053	1252	0.014016
2	60.4%	30	1000	11.5	44.2	40.2	2972	1194	0.015814
3	60.8%	30	1000	9.1	53.3	46.3	2451	1134	0.018032
4	60.7%	30	1000	7.6	60.9	50.4	2128	1074	0.017779
5	61.2%	30	1000	6.5	67.4	53.2	1905	1014	0.021449
6	60.7%	30	1000	6.1	73.5	55.4	1724	956	0.018537
7	60.2%	30	1000	5.1	78.6	56.3	1592	897	0.017042
8	59.8%	30	1000	5.4	84.0	57.7	1460	842	0.016901
9	59.8%	30	1000	5.1	89.1	58.6	1345	788	0.017470
10	59.5%	30	1000	4.7	93.8	59.0	1244	734	0.018342
11	59.7%	30	1000	6.1	100.0	61.4	1120	688	0.018697

4 When should additional follow-up be used?

The example above was, no doubt, a very specific one. However, we would like to find the basic principles for deciding on the use of additional follow-ups by varying the parameters in such example. For this purpose we are going to discuss the decision whether to use the second follow-up or not. We could choose any follow-up, the second is just an example chosen since in our case it is already not worthwhile to use.

We are interested whether the situation changes if some parameters are different. We will thus compare the situation after the first and after the second follow-up changing the following two key parameters:

1. relative bias after the first follow-up,
2. nonresponse conversion rate for the second follow-up.

4.1 The bias after one follow-up

How large should the relative bias be in order to justify the use of an additional contact? In our situation the RMSE is constant for two follow-ups and variable for one follow-up, since the variable factor is the relative bias after one follow-up. It changes from -2% to +2%. We can observe that RMSE is smaller for two then for one follow-up if the relative bias after one follow-up is larger than 1.2%[4].

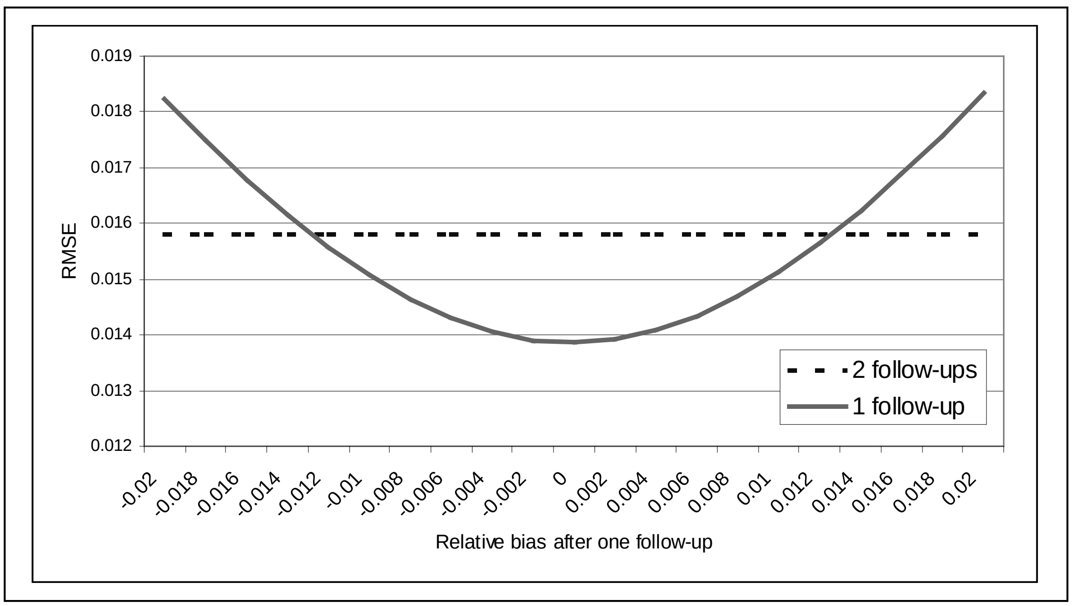

Figure 1: Influence of the relative bias after one follow-up[5].

The above figure shows only a specific case. For example, if the nonresponse conversion rate after the second follow-up was higher, the line representing its RMSE would be lower. In such a case two follow-ups would be reasonable also in the case of even smaller relative bias after the first follow-up.

[4] In our case the relative bias after one follow-up was 0.3%, so the second follow-up was not needed.

[5] In this case initial sample sizes are 4.053 for one and 2.972 for two follow-ups (this guarantees same costs for both situations), completion rates 30.9% for one and 40.2% for two follow-ups, 'true' value of 0.597 and relative bias of –1.2% for two follow-ups.

4.2 The nonresponse conversion rate after the second follow-up

Another factor that can influence the decision as regards the additional follow-up is the nonresponse conversion rate after this additional follow-up. In our case, how large should this conversion be in order to use the second follow-up?

The variable factor is thus the nonresponse conversion rate after the second follow-up. This rate influences the total completion rate and therefore also the final sample size. The larger the conversion rate, the larger the final sample and smaller sampling variance and RMSE. Of course, this variable factor has no impact on the completion rate in case of one follow-up, so this line is a constant.

We can see that RMSE is smaller for two then for one follow-up when nonresponse conversion rate after the second follow-up is larger then 30%[6]. The second follow-up is therefore justified if we expect a nonresponse conversion rate of more than 30%.

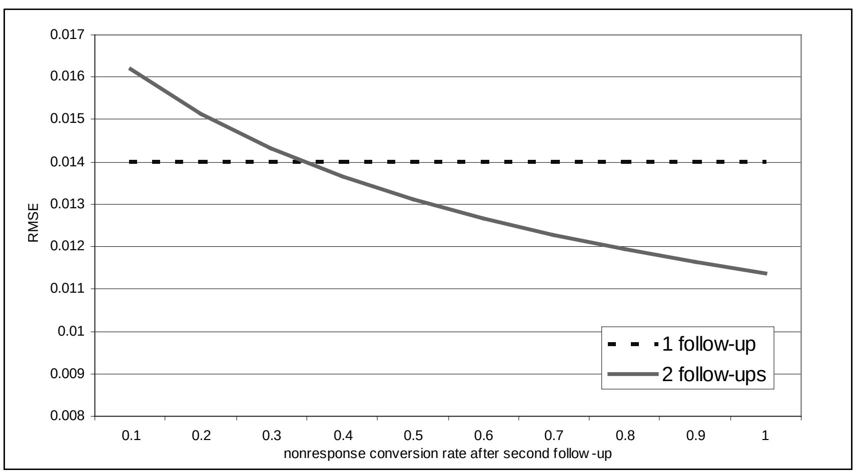

Figure 2: The impact of the expected nonresponse conversion rate[7].

Again, the above figure shows one specific case. If bias after the first follow-up was larger, the horizontal line would be higher and again two follow-ups would be reasonable even with smaller nonresponse conversion rates.

---

[6] In our case the nonresponse conversion rate for the second follow-up was 13.4%, therefore the second follow-up was not worthwhile to use.

[7] The constant factors in this case are relative bias of 0.3% for one and –0.1% for two follow-ups, 'true' value of 0.597, initial sample size of 4.053 for one and 2.972 for two follow-ups, and completion rate 30.9% for one follow-up. In this case total costs for one or two follow-ups are not completely the same, because costs for a unit of initial sample vary also with the nonresponse conversion rate for the second follow-up.

5 Conclusion

We have demonstrated the impact of various factors effecting the decision of whether or not to use additional follow-up contact in a telephone survey. Of course, the relationship between the parameters involved is complex and depends on many specific circumstances. For the simultaneous understanding of several factors together (bias, costs, response) a multivariate presentation in a multi-dimensional space may be helpful. In such a space a sort of multi-dimensional pyramid can be drawn. Only within the body of such a pyramid can the parameters take on values that may justify the use of an additional follow-up.

It is somewhat difficult to perform the above calculation in practice. One obvious complication is the case of different sub-populations which behave differently. Another obstacle may be that we have no information about the bias and the nonresponse rates. There may even be difficulties with the accurate anticipation of the costs. Of course, in such situations a good decision cannot be reached. However, it is reasonable to make certain estimates from previous surveys or, at least an educated guess. It is also possible to make estimates from earlier stages of the same survey. Based on these assumptions we can - with the aid of the above-described model - obtain a better understanding of the interaction between costs and errors in telephone surveys.

Beside costs, bias and nonresponse conversion rate, other factors also may play an important role, such as time constraints or low quality of late responses[8]. However, when faced with a clear dilemma between sampling error and nonresponse bias the above results can be useful.

References

[1] Dillman, D.A. (1978): *Mail and Telephone Survey. The Total Design Method.* New York: John Wiley and Sons.

[2] Green, K.E. (1991): Reluctant Respondents: Differences Between Early, Late, and Nonresponders to a Mail Survey. *Journal of Experimental Education*, **59**, 268-276.

[3] Groves, R.M. (1989): *Survey Errors and Survey Costs*. New York: John Wiley and Sons.

[4] Kish, L. (1965): *Survey Sampling*. New York: John Wiley and Sons.

[5] Kojetin, B.A., Borgida, E., and Snyder, M. (1993): Survey Topic Involvement and Nonresponse Bias. *Proceedings of the Section on Survey Research Method, Volume II.,* American Statistical Association, 838-843.

[8] Several authors recognize that late respondents are less interested in the survey topic and less willing to cooperate in he survey (Green 1991, 268-276; Kojetin et al. 1993, 838-843).

[6] Vehovar, V. and Lozar, K. (1998): How Many Mailings are Enough? In: A. Koch and R. Porst (Eds.), *Nonresponse in Survey Research. Proceedings of the Eighth International Workshop on Household Survey Nonresponse, 24-26 September 1997*. Mannheim: ZUMA, 139-149.