Survey Research Methods (2007)
http://w4.ub.uni-konstanz.de/srm
Vol. 1, No. 1, pp. 3-10

Minimizing survey refusal and noncontact rates: do our efforts pay off?

Dirk Heerwegh, Koen Abts and Geert Loosveldt
Center for Survey Methodology, Katholieke Universiteit Leuven

This study investigates the link between the effort undertaken to collect survey data and the nonresponse error on a key survey estimate. For this purpose a threefold analysis was conducted. First, the level of nonresponse error and its composition is charted. Second, it is investigated whether these levels change throughout the fieldwork period. This helps answering the question whether collecting more data implies higher data quality. This type of analysis also provides a possible framework for a dynamic process control during the fieldwork period. A third and final analysis links interviewer efforts (in terms of number of contact attempts) to nonresponse error and its composition. The results show that error due to noncontact is 2.6 times higher than error due to refusal, even though the refusal rate is almost two times higher than the noncontact rate. Also, the results suggest that collecting more data does not necessarily imply higher data quality and that a higher number of contact attempts does not markedly reduce the nonresponse error in absolute terms. The analysis uncovers the underlying process responsible for this latter finding.

Keywords: Survey research, nonresponse rate, nonresponse error, data quality.

Introduction

The objective of this paper is to conduct exploratory analyses concerning the general question whether reducing the survey nonresponse rate leads to data of appreciably higher quality.[1] The nonresponse rate refers to the percentage or proportion of sample cases not included in the eventually realized sample, for whatever reason (refusal, noncontact, other reasons). Data quality is measured by the difference between the parameter estimate based on the successfully surveyed sample cases and the parameter value in the total sample, which is thought of as the true value. A sample estimate closer to the true value indicates higher data quality. This definition of data quality is relatively strict, and usually not attainable in common survey research since the 'true' sample value is unknown and if it were known, no survey research would be necessary. In certain cases however, the possibility exists to obtain the true value in the sample about a limited number of variables. It is for instance possible to match the survey results on voting behavior to actual election results (Voogt and Saris 2005). Alternatively, administrative data may be available for all sample cases on an individual basis.

The current survey, commissioned by the Minister of Housing, intended to determine the living conditions of the residents in Flanders (the Northern part of Belgium). A large array of characteristics of all sampled households' residences was collected independently from sample cases' cooperation with the face-to-face survey interview.[2] One of the characteristics recorded was whether the household occupied a house or an apartment, which was a key variable of interest in the study. Nonresponse bias will be assessed using this target variable. This paper will compare the proportion of households occupying a house vs. an apartment in the sample with the proportion that would have been obtained if residence information had been collected only for the respondents (based on the face-to-face interview alone).[3]

Although the proportion of households living in apartments vs. houses is a parameter of interest in this specific study, it remains a parameter of broader interest since type of residence can correlate with the noncontact and refusal probabilities (Groves and Couper 1998:139-140). It is therefore interesting to investigate the error on this estimate, which should give some overall impression of the data quality and which could give some insight into the question whether the amount of effort spent on collecting the data is related to the degree of obtained data quality.

In the following section, the research questions are formulated. The subsequent section details the specifics of the data used in this paper. The analyses are presented in a next

Contact information: Center for Survey Methodology, Katholieke Universiteit Leuven, Parkstraat 45 - bus 3601, 3000 Leuven, Belgium (Dirk.Heerwegh@soc.kuleuven.be).

[1] The data used in this report were collected as part of the 'Flemish Housing Survey' conducted by the 'Research Network on Sustainable Housing Policy' commissioned by the Ministry of the Flemish Community, Housing Policy Department.

[2] In the Data section, this independent data collection process is elaborated on.

[3] This analysis evaluates the net effect of nonresponse, eliminating other potential sources of survey error (Groves 1987). Although the sample remains susceptible to sampling error, the analysis does not make an attempt at comparing the sample estimate with the true population proportion. For the same reason, noncoverage error is not a factor in the current analysis. And finally, although measurement error may have occurred (a housing unit may have been misclassified), the analysis does not compare a measurement value with a true value, but rather compares the same measurement across two groups in the sample (the respondents and the nonrepondents).

section, followed by a discussion and conclusions section. Since these analyses are rather explorative of nature, no specific hypotheses are presented.

Research questions and expectations

This paper deals with three related research questions on the general topic of nonresponse error. In this paper, nonresponse error refers to the discrepancy observed between the estimated proportion of a specific residence type based on the respondent data and the estimate based on the full dataset including both respondents and nonrespondents.

A first research question seeks to determine the magnitude of the nonresponse error obtained on the key estimate (proportion of households living in a house as opposed to living in an apartment), as well as the nonresponse error composition (refusal, noncontact, other).

A second research question is whether nonresponse error composition varies by the amount of data collected. A possible strategy to answer this question is to evaluate data quality chronologically by dividing the entire fieldwork period into smaller chronologically ordered periods (e.g. weeks of data collection).This allows a detection of 'trends' in data quality throughout the entire fieldwork period. In this paper, the fieldwork period is divided into 10 segments or 'slices' of 10% fielded addresses. Nonresponse error can be assessed at ten intervals, enabling an evaluation of the effect of fielding an additional 10% of the available addresses on nonresponse error.

The third research question is whether the nonresponse error and/or its composition would have been different if fewer contact attempts had been made. This question is related to the research about the impact of extended interviewer efforts on nonresponse bias (Lynn et al. 2002). With extended interviewer efforts one tries to contact difficult-to-contact respondents and to convert reluctant respondents.[4] It is a 'hunt for the last respondent' (Stoop 2005). It is common practice in survey research to increase the number of contact attempts to obtain a higher response rate (and, hopefully, data of higher quality). As more contact attempts are made, a lower noncontact rate should be observed. If the 'continuum of resistance model' holds, a lower noncontact rate will decrease the noncontact error. If all other error components remain equal, this should reduce nonresponse error. The continuum of resistance model assumes that nonrespondents are similar to the most reluctant respondents (Stoop 2005:222). Recent research however suggests that this model cannot be assumed to be routinely justified (Lin and Schaeffer 1995; Stoop 2005). Keeter et al. (2000) and Curtin et al. (2000) found that greater efforts to increase or maintain response rates did not lead to substantially different point estimates. It is conceivable that the difference between noncontacts and respondents increases as the noncontact rate is decreased by conducting more contact attempts, because the hardest to contact people remain non contacted while the easier to contact people who resemble the respondents more are added to the pool of respondents. Nonetheless, it remains unknown whether this potentially increasing difference between respondents and noncontacts is large enough to increase the resulting noncontact error. This remains an empirical matter to resolve.

Data

To answer the aforementioned research questions, we draw upon data from a large scale study on housing conditions and housing needs in Flanders. The study consists of a random sample of 7,770 addresses, drawn from the National Register. Since the interview needed to be conducted with the reference person of the household, these names were also extracted from the National Register. The sample was stratified disproportionately with respect to geographical area in Flanders (district). For reasons of comparability with a previous study and to avoid very small numbers of sample cases per district, the minimum number of sample cases per district was fixed at 210. The larger districts were consequently underrepresented in the sample.

The Flemish Housing study consists of two parts. The first part is a face-to-face survey. Interviewers were used to conduct a face-to-face interview (CAPI) with the reference person. Interviewers were instructed to fill in a contact sheet for each contact attempt. These contact sheet data will also be used in the current analysis. The realized sample includes 5,216 respondents who were effectively interviewed in their homes.

Independently from the face-to-face survey, inspectors were used to register auxiliary data on the exterior characteristics of the private residences of each of the sample units. This is the second part of the Flemish Housing study. The inspectors recorded the type of residence (house vs. apartment), and a large number of other aspects of the residence (some 560 variables were registered).[5] Since data are available for all 7,770 sample cases for these variables, this dataset can be used to evaluate nonresponse bias. The 'type of residence' variable contained within this dataset will constitute the focus of our analyses.

The survey was fielded from April 2005 to February 2006, although data were collected during only a few days in February 2006. The period of data collection therefore is about 10 months. Since the survey overrepresented small districts and underrepresented larger districts, and because the fieldwork agency had more interviewers for the larger districts and fewer for the smaller districts, there is a certain

[4] The fieldwork procedures required that any refusal obtained by telephone would be followed up by a personal contact in an attempt to convert the refusal. At the discretion of the interviewer, sample cases that refused tacitly (without overtly saying "I don't want to do this") could also be followed up. In all, only 9.2 percent of all refusals were followed up by an attempted refusal conversion. Because of this small percentage, we focus exclusively on the number of contact attempts in the analysis.

[5] These data were collected without entering the respondent's residence and without contacting the resident. E.g., data on the condition of the roof were collected.

Table 1: Outcome for the fielded addresses

Outcome	Full sample		Usable sample	
	n	%	n	%
Interview	5,216	67.13	5,112	69.74
Noncontacts	661	8.51	646	8.80
Refusals	1,242	15.89	1,223	16.68
Other	418	5.38	349	4.67
Ineligible	233	3.00	–	–
Total	7,770	100.00	7,330	100.00

amount of chronological displacement noticeable in the fieldwork. The larger districts were fielded sooner in the fieldwork, while the data collection for smaller districts was at its peak later in the fieldwork period. In the analyses, we correct for this factor.

Results

Table 1 lists the outcome of the 7,770 fielded addresses. Using the AAPOR response rate definition (AAPOR 2004), this survey obtained a 69.2% response rate (5,216/(7,770-233)).

For each of these 7,770 cases, an external evaluation of the private residence was made. The 233 ineligible cases (vacant dwellings) are excluded from further analyses since by definition survey interviewers cannot contact such cases. In addition, inspectors were unable to record the type of residence for 149 cases, and 58 interviewed cases were later deleted from the dataset after a quality check.[6] The net result is that the sample decreases from 7,770 to 7,330 cases. This is the sample on which the following analyses are based. The last two columns of Table 1 detail the outcome of these 7,330 fielded addresses.

Nonresponse error and error composition

Based on the administrative data (n=7,330), 81.15% of the sample cases live in a house and the remainder (18.85%) live in an apartment. If data on type of residence had only been available for the sample cases that cooperated with the survey (n=5,112), we would have estimated the percentage of households to live in a house at 84.86%, and the percentage living in an apartment at 15.14%. This (realized) sample distribution is significantly different from the true sample distribution available from the administrative data (χ^2=45.98; df=1; $p < 0.0001$; n=5,112). These results indicate that a significant overestimate of the percentage of households occupying a house would have resulted as a consequence of nonresponse.

If the percentage of households living in a house is our target estimate, the nonresponse error in the survey estimate is 0.0371 or 3.71% (84.86% – 81.15%).[7] Using the following expression (taken from Groves and Couper 1998:3),

$$\bar{y}_r = \bar{y}_n + \left(\frac{m}{n}\right)[\bar{y}_r - \bar{y}_m] \tag{1}$$

where r refers to the respondents, n to the total sample, and m to the missing observations, it can be seen that $\bar{y}_r$ (0.8486) equals $\bar{y}_n$ (0.8115) plus the nonresponse error (0.0371). Therefore, $\bar{y}_m$ equals 0.7260. This means that 72.6% of the eligible sample cases not interviewed live in a house. This is quite different from the respondent group where 84.86% lives in a house.

Further differences between the response and the nonresponse groups can be observed when the latter is segmented into subclasses that are different in origin (Mayer and Pratt Jr. 1966). The most important subclasses are noncontacts and refusals. It is possible to rewrite Equation 1 to a more elaborated equation (cf. Groves and Couper 1998:12),

$$\bar{y}_r = \bar{y}_n + \left(\frac{m_{ref}}{n}\right)\left[\bar{y}_r - \bar{y}_{m_{ref}}\right] + \left(\frac{m_{nc}}{n}\right)\left[\bar{y}_r - \bar{y}_{m_{nc}}\right] + \left(\frac{m_{oth}}{n}\right)\left[\bar{y}_r - \bar{y}_{m_{oth}}\right] \tag{2}$$

where m_{ref} refers to the missing observations due to refusal, m_{nc} to the missing observations due to noncontact and m_{oth} to the missing observations due to other reasons (not including ineligible cases). The data allow filling in Equation 2 as follows:

$$.8486 = .8115 + \left(\tfrac{1,223}{7,330}\right)[.8486 - .8021] + \left(\tfrac{646}{7,330}\right)[.8486 - .6161] + \left(\tfrac{349}{7,330}\right)[.8486 - .6619]$$

These figures show that refusal error (0.0078) and error due to other reasons for nonresponse (0.0089) make the smallest contributions to the nonresponse error (21.0 and 24.0%, respectively), while noncontact error (0.0205) makes a larger contribution to the nonresponse error (55.26%). These figures show that the error introduced due to noncontact is higher than the error introduced due to refusal even though the refusal rate is appreciably higher (16.7%) than the noncontact rate (8.8%). This is because the sample cases that refused cooperation resemble the respondents more in terms of type of residence than the noncontacts. This nicely illustrates that error is a combination of the missing rate and the difference between respondents and missing observations.[8] It is therefore difficult to assess data quality solely on missing rates, as already noted by some authors (Curtin, Presser and Singer 2000; Groves, Presser and Dipko 2004). If only the refusal and the noncontact rates were known, one could easily come to the wrong conclusion that the error due to refusal should be higher than the error due to noncontact.

[6] The quality check consisted of, among other things, an inspection of the duration of the interview.

[7] All subsequent analyses take the percentage of households living in a house as the target estimate. Since this is a dichotomous variable, overrepresentation of one category by a certain percentage implies underrepresentation of the other category by the same percentage.

[8] The term 'missing' is meant as a general term which needs to be specified regarding the reason why data are missing, e.g. "refusal error is a combination of the missing due to refusal rate and the difference between respondents and refusers."

Table 2: Type of residence and type of nonresponse

Nonresponse	House		Apartment		Overall	
	n	%	n	%	n	%
Interview	4,338	72.93	774	56.01	5,112	69.74
Noncontact	398	6.69	248	17.95	646	8.81
Refusal	981	16.49	242	17.51	1,223	16.68
Other	231	3.88	118	8.54	349	4.79

χ^2=258.95; df=3; p < 0.0001

In line with these results, Table 2 shows that although the refusal percentage is about equal across both types of residences, the noncontact rate is 2.7 times higher for apartments than for regular houses. Also, the 'other' reasons are twice as high for apartments as for houses. Because of these large differences in interview, noncontact, and 'other reasons for nonresponse' rates across both types of residences, the survey estimate of the proportion of households living in either a house or an apartment are biased, as shown earlier.

Changes in nonresponse error composition as more data are collected

The previous analyses showed that at the end of the fieldwork period (i.e. after fielding 100% of the available sample addresses), the survey overestimates the percentage of households living in a house and consequently underestimates the percentage of households occupying an apartment. It is of interest to determine whether these components remain at about the same level throughout the entire fieldwork period.

Contact sheet data are used to answer this question. More specifically, the date of the first contact attempt of each address is taken into account. Since this was a large survey, not all addresses were fielded simultaneously (the total fieldwork period spanned 10 months). It is therefore meaningful to identify the 10% first fielded addresses, the next 10% fielded addresses, and so on.[9] It is of importance to note that this process of 'slicing up' the total fieldwork period into 10 segments or 'slices' of 10% fielded addresses was conducted by geographical area. Some geographical areas were covered early in the fieldwork period, while others were covered later on in the fieldwork period. Identification of the fielding date of the addresses independently from geographical area was hence necessary. Thus, the first 10% fielded addresses as displayed in Figure 1 refers to the first 10% fielded addresses in each region. It should also be noted that this type of figure is cumulative of nature. In order to observe when errors stabilize, it is necessary to add each next slice of 10% addresses to the already fielded addresses (for a similar, but distinct, operation see Groves and Heeringa (2006) when they graphically demonstrate the notion of phase capacity). Also, while this type of figure takes into account the date at which the first contact attempt was made, the outcome of the fielded address (interview, noncontact, refusal, other) is of course the final outcome, which is not necessarily the outcome of that first contact attempt.

Figure 1 displays the nonresponse error, as well as the

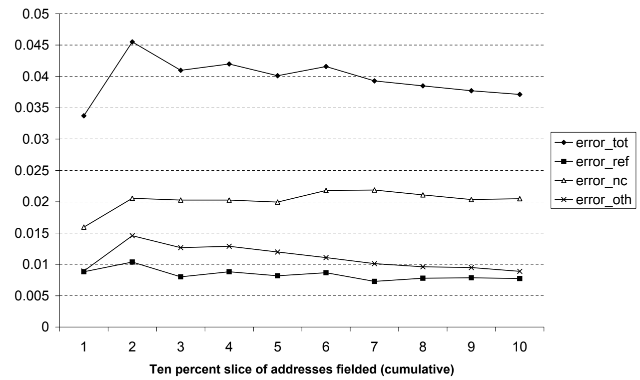

Figure 1. Decomposition of the nonresponse error, by ten percent slices of fielded addresses (error_tot = nonresponse error; error_ref = error due to refusal; error_nc = error due to noncontact; error_oth = error due to other reasons for nonresponse)

three distinguished components of this error. This graph shows that the error due to refusal and other reasons is smaller than the error due to noncontact. In addition, this graph shows no striking fluctuations throughout the entire fieldwork period. At each stage of the fieldwork, the error due to noncontact is larger than the error due to refusals and other reasons. When investigating the noncontacts (Figure 2) and the refusals (Figure 3) more closely, the same conclusions can be drawn. Throughout the entire fieldwork period, the noncontact rate fluctuates from a low of 7.5% to a high of 8.8%, and the difference between respondents and noncontacts also fluctuates slightly (a low of 0.021 and a high of 0.025), resulting in a relatively stable error level (range: 0.016 to 0.022). Similarly, the refusal rate fluctuates, as does the differences between respondents and refusals, but the resulting error is relatively stable.

These figures show that although the noncontact and refusal rates may fluctuate throughout the fieldwork period, the associated error terms are relatively stable, leading to a relatively stable nonresponse error throughout the entire fieldwork period. This provides an answer to the second research question. It can consequently be concluded that collecting more data does not necessarily imply improvements in the nonresponse error or in the various nonresponse error components.

Simulating the effect of fewer contact attempts

In this particular survey, interviewers were instructed to make at least four contact attempts, of which at least one in the evening and another one during the week-end. The data show that the interviewers made 4.78 contact attempts on average for the 646 households that eventually remained noncontacts. This indicates that the interviewers exceeded the minimum required effort level. It was also shown earlier that

[9] Alternatively, one could look at the data by week, but this would result in rather small numbers of observations per data collection period.

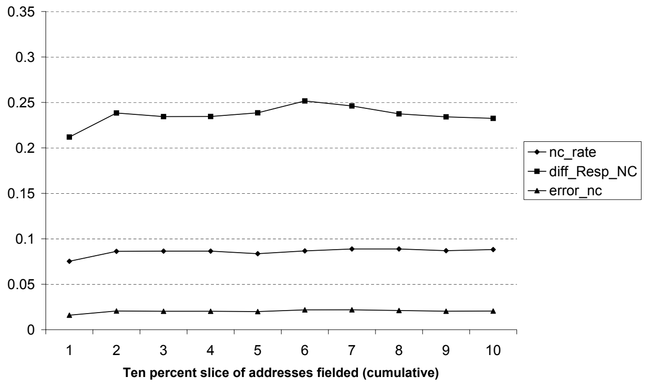

Figure 2. Error due to noncontact, by ten percent slices of fielded addresses (nc_rate = noncontact rate; diff_Resp_NC = difference between Respondents and Noncontacts; error_nc = error due to noncontact)

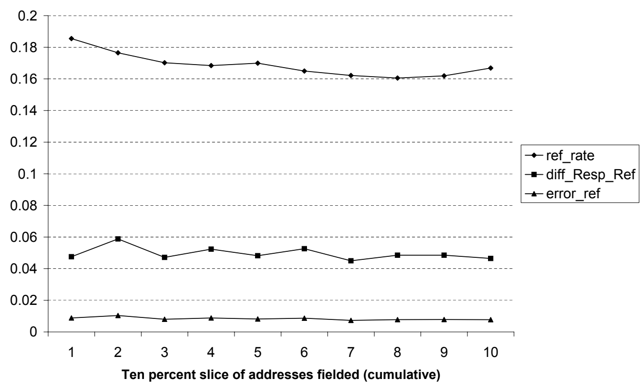

Figure 3. Error due to refusal, by ten percent slices of fielded addresses (ref_rate = refusal rate; diff_Resp_ref = difference between Respondents and Refusals; error_ref = error due to refusal)

the noncontact rate was almost three times higher for apartments than for houses (cf. Table 2). Nonetheless, the interviewers made more contact attempts for noncontacted apartments (average of 5.19, n=248) than for noncontacted houses (average of 4.52, n=398; Kruskal-Wallis χ^2=24.82; df=1; p < 0.0001). Moreover, the interviewers undertook more personal contact attempts for apartments (on average 4.87) than for houses (on average 4.10; Kruskal-Wallis χ^2=27.9; df=1; p < 0.0001); and fewer telephone attempts for apartments (on average 0.32) than for houses (on average 0.43; Kruskal-Wallis χ^2=4.15; df=1; p=0.0416) that remained uncontacted at the end of the fieldwork period. This indicates that despite more efforts of the interviewers (more, and more personal contact attempts), apartments remain more difficult to contact.[10]

This raises the question regarding the effects of more contact attempts on the noncontact rate, but also on the other rates and the associated errors. It is of interest to observe the possible changes in the nonresponse error, and whether the composition of the nonresponse error changes with the number of contact attempts. To answer these questions, data from the contact sheets was used. These data show the result of each of the contact attempts, and it is possible to 'censor' some information from these data (which comes down to truncating the data set, see Curtin, Presser and Singer 2000). More specifically, it is for instance possible to discard all contact attempts after the fourth attempt, and hence to treat the result of the fourth attempt as the final result. This would simulate the results of having made fewer contact attempts. Five such simulations were run. The outcome was simulated if no more than 2, 3, 4, and 5 contact attempts had been made.[11] These situations can be compared with each other and with the situation in which we take into account all contact attempts (this maximum number of contact attempts equals 12).

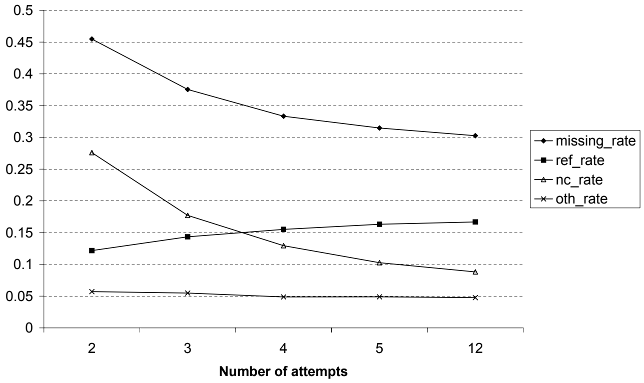

Figure 4. Overall missing (nonresponse) rate, refusal rate (ref_rate), noncontact rate (nc_rate), and the nonresponse rate due to other reasons (oth_rate), by number of contact attempts

Unsurprisingly, Figure 4 shows that a higher number of contact attempts leads to a decrease in the overall nonresponse rate (from 45.5% to 30.3%). This figure also shows that more contact attempts lead to a lower noncontact, but to a higher refusal rate. This can be explained as follows. As more households are contacted, more interviews are obtained (the nonresponse rate drops), but a number of these now contacted households also refuse cooperation, which in the current survey raises the refusal rate.

Figure 5 shows the relationship between the nonresponse rate and the nonresponse error by depicting the differences between the respondents and the nonrespondents. This figure shows that the nonresponse error decreases from 0.052 to 0.037. This means that the survey estimate of the percentage

[10] An additional analysis revealed that roughly 35% of all contact attempts were made after 6 pm. This figure was not significantly different across type of housing unit (35.1% for houses and 35.8% for apartments). The higher noncontact rate for apartments is consequently not attributable to differences in interviewers' strategies regarding timing of contact attempt.

[11] This effectively means that the result of the i^{th} (i=2,3,4,5) contact attempt is taken as the final outcome for a sample case, unless the outcome of contact i was an appointment in which case the result of a next contact attempt was taken.

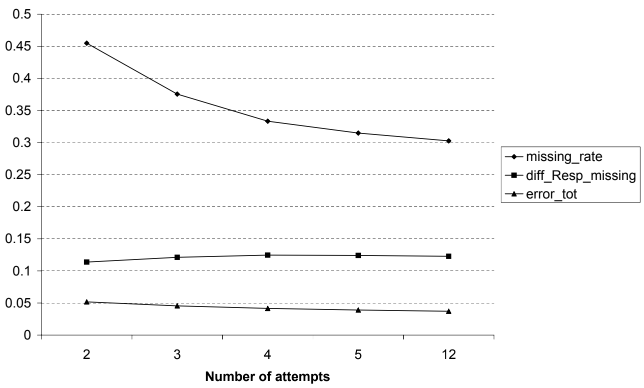

Figure 5. Nonresponse error by number of contact attempts (diff_Resp_missing = difference between respondents and missing observations; error_tot = nonresponse error)

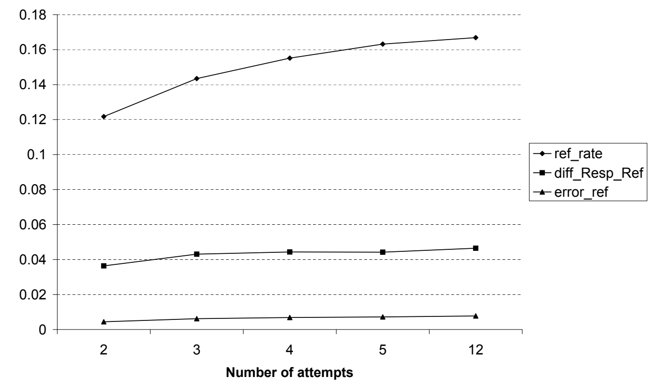

Figure 7. Error due to refusal, by number of contact attempts (ref_rate = refusal rate; diff_Resp_ref = difference between Respondents and Refusals; error_ref = error due to refusal)

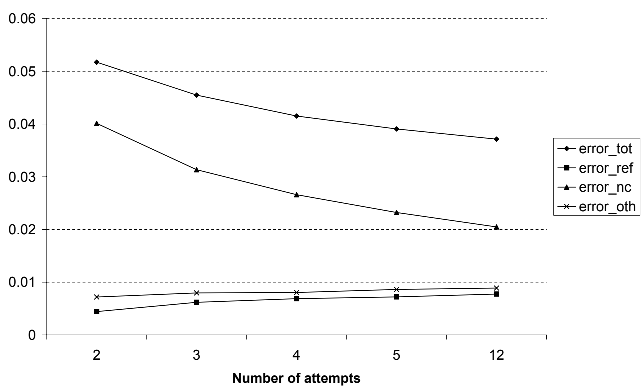

Figure 6. Decomposition of the nonresponse error, by number of contact attempts (error_tot = nonresponse error; error_ref = error due to refusal; error_nc = error due to noncontact; error_oth = error due to other reasons for nonresponse)

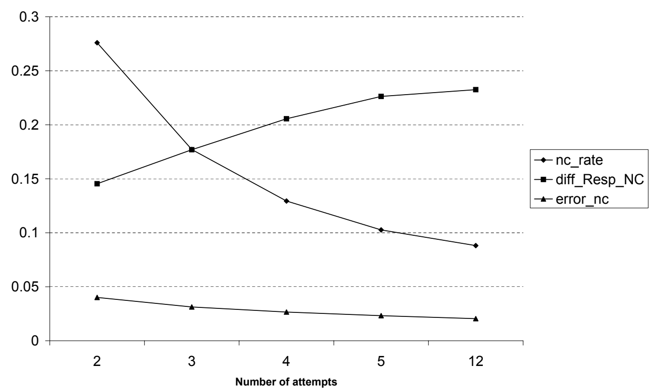

Figure 8. Error due to noncontact, by number of contact attempts (nc_rate = noncontact rate; diff_Resp_NC = difference between Respondents and Noncontacts; error_nc = error due to noncontact)

of households living in a house (as opposed to an apartment) is overestimated by 5.2 percentage points if only two contact attempts are made, and is still overestimated by 3.7 percentage points if up to 12 contact attempts are made. Although the absolute figures seem small, this does represent a reduction of the overestimation of 28.8% by increasing the number of maximum attempts from 2 to 12 (the associated noncontact rate reduction was 33.4%). This analysis was repeated while dividing the fieldwork period into 10% chronological slices of fielded addresses. That analysis showed the same results as above in each stage of the fieldwork period.

Figure 6 shows that the error composition changes with the number of contact attempts. This figure suggests that a larger number of contact attempts primarily influences the noncontact error. The other error components show a subtle increase with an increasing number of contact attempts.

When investigating the refusals and the noncontacts more closely (Figure 7 and Figure 8), an interesting pattern emerges. Regarding the error due to refusal (Figure 7), it can be noted that the refusal rate increases as more contact attempts are made, the difference between the respondents and the refusals increases slightly, and the error due to refusal consequently rises modestly as well. The same pattern is observed throughout the entire fieldwork period (data not shown). Undertaking up to 12 contact attempts instead of only 2 effectively halves the noncontact error (from about 0.04 to 0.02), although the noncontact rate is affected more strongly (from 28 to 9%, a reduction with factor 3.1). The reason for the error to be less influenced than the noncontact rate is the ever increasing difference between respondents and remaining noncontacts as more contact attempts are undertaken (see Figure 8).

These findings suggest that as more contact attempts are made, a lower noncontact rate results. Also, more contact attempts reduce the noncontact error, and late respondents (21.4% apartment dwellers) are more similar to noncontacts (38.4% apartment dwellers) than early respondents (14.5% apartment occupants) are to noncontacts.[12]Still, late respon-

[12] This analysis compares 4,607 cases who cooperated using less than five contact attempts with 495 cases who needed 5 or more contacts before they cooperated and 646 remaining uncontacted cases.

dents remain more similar to early respondents than to noncontacts. The unequal contact probability across residence type persists and the remaining group of noncontacted sample cases becomes more and more distinct from the group of respondents.

The results also suggest that an error component that does not appear to change much may conceal a very dynamic mechanism at the level of a specific missing rate and the difference between the respondents and the missing observations (due to noncontact, refusal, or other reasons).

Relationship between type of residence and other survey variables

Even though most researchers conducting in-house household surveys will be confronted with a variable 'type of residence' (house vs. apartment), and even though this will possibly affect the noncontact probabilities, not all survey researchers will be interested in estimating the proportion of households living in each type of residence. Nonetheless, the current analysis may still be of interest to the larger survey research community if it is shown that type of residence is associated with other survey measures. The current survey did not include many questions on opinions or attitudes, since the main interest was in the housing and living conditions of the households. Nonetheless, some variables referring to attitudes are available. First, there was a question on "how safe do you feel in your neighborhood after dark?" Responses were collected on a six point scale ranging from 1=very safe to 6=very unsafe. The average score of the respondents living in an apartment was significantly higher (2.37, n=768) than the average score of the respondents living in a house (2.00, n=4,331; Kruskal-Wallis χ^2=44.08, df=1, $p < 0.0001$), which reflects a lower feeling of safety in the group of respondents living in an apartment.

Similarly, those living in an apartment reported a lower average level of satisfaction with their neighborhood (average score of 1.90 on a 1 to 5 scale where 5=very unsatisfied) than the respondents living in a house (average score of 1.71; Kruskal-Wallis χ^2=50.67, df=1, $p < 0.0001$). More apartment occupants also reported to be inclined to 'certainly' or 'probably' move out of their neighborhood if they had the opportunity than respondents living in a house (28.48% vs. 12.51%; χ^2=131.25, df=1, $p < 0.0001$). And, along the same lines, more apartment occupants reported an inclination to move out of their residence if they had the opportunity than respondents living in a house (43.21% vs. 17.88%; χ^2=247.38, df=1, $p < 0.0001$).

Since the survey underrepresents people living in apartments, the survey probably overestimates the overall feeling of safety in the neighborhood and the overall level of satisfaction with the neighborhood. For the same reason, the survey probably underestimates the overall inclination to move out of the neighborhood and out of the residence. Since these are crucial variables in this survey, it is important to establish that these survey estimates are probably biased. This analysis also shows that other surveys, while perhaps not interested in the percentage of people living in certain types of residences, will also produce biased survey estimates if the survey estimate is correlated with type of residence.

Discussion and conclusion

The objective of this paper was to evaluate the error on a single (important) survey variable and to relate it to the amount of effort spent to collect the data. The value of the variable in question, type of residence, was determined by means of data collection conducted for all sample cases independently from their participation or non-participation with the survey. This allows an investigation what would have happened if data on type of residence would have been available only for the survey participants. In addition, contact sheet data allows differentiating between different types of nonresponse error. The results show that the survey overrepresents households living in a house, and underrepresents apartment occupants. The results also indicate that 'refusals' and 'other reasons for nonresponse' make the smallest contribution to the nonresponse error, while the 'noncontacts' category makes a larger contribution to the nonresponse error. This indicates that error (on this variable) is due more to noncontacts than to refusal.

Since this was a large survey, we could divide the fieldwork period in smaller periods ('rounds') of data collection, each of which representing a sizable amount of fielded addresses (around n=730). This allows an investigation of the trends in error over time. The results show that the nonresponse error and its components remain relatively stable throughout the entire fieldwork period. This indicates that simply collecting more data (increasing the sample size) does not decrease the nonresponse error. Incidentally, this type of analysis could be useful to monitor the fieldwork as it progresses, and could provide opportunities to adjust and optimize fieldwork procedures during the fieldwork.

The current study also attempted to evaluate the effect of the number of contact attempts on the error. The results indicate that more contact attempts reduce nonresponse error, although in absolute terms the reduction is not very large. The primary reason for this could be found in the observed pattern for the noncontacts. More contact attempts decrease the noncontact rate, but the remaining noncontacts become more different from the contacted cases. The reduced noncontact rate in combination with the increased difference between noncontacts and respondents makes for an only slight reduction in error due to noncontacts. This illustrates the basic tenet of what constitutes error: the nonresponse rate *and* the difference between respondents and nonrespondents. This simulation study truncated the contact sheet data set, which is not the same as giving interviewers the instruction not to make more contact attempts than 2, 3, 4, or 5. If such a limitation were imposed upon the interviewers, it is possible that they would change their contact strategy, which would lead to different results than the current simulation. Nevertheless, we believe that these simulations provide useful approximations.

We have also shown that type of residence is associated with the response to a number of available variables of interest (e.g. feelings of safety). Consequently, the current results would also be of interest to researchers conducting surveys on other topics, since all surveys will to a certain extent be confronted with lower contact rates in apartments. Perhaps it would be useful to instruct interviewers to record the type of residence on the contact sheets, so that this information can be used in post-survey stratification methods.

In conclusion, these results indicate that collecting more data does not necessarily decrease the error. Care must be taken not to confuse 'more data' with 'better data'. The results provide compelling evidence that the remaining noncontacts are at the heart of the remaining error on the survey estimate concerning type of residence in the current survey. Even though the simulation studies suggest that more contact attempts do not markedly reduce the noncontact error in absolute terms, perhaps the contact strategy itself should be improved (e.g. timing). This could be a venue for future research interested in decreasing the error due to noncontact.

References

AAPOR. (2004). *Standard definitions: Final dispositions of case codes and outcome rates for surveys.* (3 ed.; AAPOR, Ed.). Lenexa, Kansas: AAPOR.

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly*, *64*, 413-428.

Groves, R. M. (1987). Research on survey data quality. *Public Opinion Quarterly*, *51*, S156-S172.

Groves, R. M., & Couper, M. P. (1998). *Nonresponse in household interview surveys*. New York: Wiley.

Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly*, *68*, 2-31.

Groves, R. M., Robert, M., & Heeringa, S. G. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *169*, 439-457.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. (2000). Consequences of reducing nonresponse in a telephone survey. *Public Opinion Quarterly*, *64*, 125-148.

Lin I-F., & Schaeffer, N. C. (1995). Using participants to estimate the impact of nonparticipation. *Public Opinion Quarterly*, *59*, 236-258.

Lynn, P., P.Clarke, J.Martin, & P.Sturgis. (2002). The effects of extended interviewer efforts on nonresponse bias. In R. Groves, D. Dillman, J. Eltinge, & R. Little (Eds.), *Survey nonresponse* (p. 135-147). New York: Wiley.

Mayer, C. S., & Pratt Jr., R. W. (1966). A note on nonresponse in a mail survey. *Public Opinion Quarterly*, *30*, 637-646.

Stoop, I. (2005). *Nonresponse in sample surveys. the hunt for the last respondent*. The Hague: Social and Cultural Planning Office.

Voogt, R. J. J., & Saris, W. E. (2005). Mixed mode designs: Finding the balance between nonresponse bias and mode effects. *Journal of Official Statistics*, *21*, 367-387.