Survey Research Methods (2009)
Vol.3, No.2, pp. 59-72
ISSN 1864-3361
http://www.surveymethods.org

Nonresponse in the Recruitment of an Internet Panel Based on Probability Sampling

Adriaan W. Hoogendoorn
GGZ inGeest

Jacco Daalmans
Statistics Netherlands

In this paper we study the selectivity in the recruitment of respondents for a Dutch Internet panel (the CentERpanel). This recruitment is based on a probability sample. It involves three stages: participation to a first telephone interview, willing to be re-contacted and final agreement to participate in the Internet panel. By matching data of the recruitment process with registries of Statistics Netherlands we are able to distinguish selectivity with regard to age and income in all stages and with pc-ownership in the latter two stages only. Interestingly, we hardly find any selectivity with key variables on living conditions. Finally we will make some explicit recommendations for the recruitment process.
Keywords: Internet Panel Surveys, Nonresponse Error, Pre-recruited panels, recruitment process

Introduction

Web surveys provide a low cost option for data collection and can save time to publish results (Dillman 2000). In describing the various categories of Web-based surveys Couper (2000) distinguishes eight types, grouping these either as non-probability or as probability-based and discusses the uses and disadvantages of each type. The overview shows that the goals of Web surveys range from entertainment to serious research. In this paper we put ourselves at the 'top end' and want to use Web survey data to make inferences about the 'full population'. In his overview, Couper argues that two approaches are most appropriate for this aim that we will discuss here: 'volunteer opt-in panels' and 'pre-recruited panels of full populations'.

If one uses a volunteer opt-in panel to make inferences about the 'full population', one has to deal with problems of non-coverage (not all members in the population are connected to the Internet) and selection bias (volunteers usually differ to a great extent from non-volunteers). Harris Interactive, an American market research company, that uses Web based interviewing on a large scale, claims that these problems can be resolved by use of a method called *propensity scoring* (see Terhanian et al. (2001)). The propensity scoring technique uses a *reference survey* to calibrate the Web survey (see e.g. Schonlau et al. (2004), Lee (2006a), Schonlau et al. (2007)). Although the reference survey and the Web survey are independent, they share certain demographic and attitudinal variables (so-called *webographics*) that are used in a propensity model. In practice Harris Interactive uses a Random Digit Dialing telephone survey as the reference survey. Propensity scores for the respondents in the Web survey are obtained as follows. First the data of the Internet survey and the data of the reference survey are combined into a single data set. Then for the combined data containing both types of respondents a logistic model is fitted that uses the *webographics* as independent variables to predict the probability that a respondent is a Web survey respondent rather than a reference survey respondent. These predicted probabilities are called propensity scores. For respondents with the same propensity score, the observed differences are due to chance rather than systematic bias (Schonlau et al. 2004). This explains the attractiveness of the procedure. Criticisms to the propensity scoring technique point at several weaknesses: the reference survey itself is likely to have problems of both non-coverage and selection bias, the *webographics* turn out to be poor predictors for being a Web survey respondent and a small size of the reference survey results into high variances of the outcomes of the main survey (see Bethlehem 2007).

The second approach in Couper's (2000) overview that may be appropriate for making inferences for the full population is called 'pre-recruited panels of full populations'. In this approach the recruitment of respondents for the Internet panel is based on a probability sample and uses a conventional data collection method, such as a doorstep survey or a telephone survey, for respondent recruitment. Examples are panels of Knowledge Networks (see Huggins and Eyerman (2001) and Smith (2003)) or CentERdata (see Saris (1998) and Hoogendoorn et al. (2000)). In order to solve the problem of non-coverage of households without Internet access, the survey organization provides the infrastructure if necessary. However, the approach still suffers from nonresponse. In fact, the structure of nonresponse is very complex since nonresponse occurs at various stages in the survey process. We distinguish *initial nonresponse* that occurs in the recruitment stage of the Internet panel, *wave nonresponse* that occurs when panel members do not participate in a particular wave of the survey and *panel attrition* as a result of respondents dropping out of the panel. Couper (2000) mentions that the first and the second type of nonresponse have been

Contact information: Adriaan W. Hoogendoorn, GGZ inGeest, AMSTAD/NESDA, A.J. Ernststraat 887, 1081 HL Amsterdam, the Netherlands, e-mail: a.hoogendoorn@ggzingeest.nl

seen as the key problem with this approach. The problem of nonresponse is especially cumbersome if not all groups are equally well represented. Ignoring this problem may lead to biased estimates of population characteristics. Therefore it is important to study the selectivity of the nonresponse. For probability-based Web surveys it is possible to measure the amount and selectivity of the nonresponse, if the sample framework is based on a register and data on background variables are available.

Several studies have already been carried out. Couper et al. (2007) found selectivity with respect to demographic, financial and health related variables in an Internet survey among persons of at least 50 years old in the US. Lee (2006) came to the same conclusion when she studied the selectivity of nonresponse and non-coverage of a Web Sample of Knowledge Networks.

The purpose of this paper is to measure the selectivity of the initial nonresponse of an Internet panel. We analyzed data of the recruitment process of the CentERpanel, one of CentERdata's Internet panels. CentERdata is a research institute at Tilburg University that is specialized in online data collection. The institute manages and maintains several panels that regularly complete questionnaires through the Internet. The process of recruiting respondents for the CentERpanel can be split into several stages. For each stage we will study the selectivity with respect to a large selection of demographic and socio-economic characteristics. The analysis is possible by matching the information of the recruitment of CentERdata with Statistics Netherlands data that come from registrations. The findings will show the selectivity of nonresponse both at the level of 'total initial nonresponse' and at the different stages and lead to suggestions to take bias reduction measures for each stage. The paper will proceed as follows. First we will describe the process of recruiting respondents in the panel of CentERdata. Next we will describe the data that we will use to study the selectivity in response. Then we will provide results and finally we will draw some conclusions.

Respondent recruitment in the CentERpanel

CentERdata maintains an Internet panel consisting of two thousand households in the Netherlands. The CentERpanel has its founding in the Dutch Telepanel that started in 1988 as a method for interviewing without interviewers (Saris 1998). Every weekend, members of all households in the panel are asked to complete a questionnaire of approximately thirty minutes on the Internet. Households are recruited using a CATI survey on a sample of households drawn from KPN's telephone directory (KPN is a Dutch telecommunications company, owner of the fixed telephone network). Although we will focus our analysis on the process of nonresponse of *sampled households*, we emphasize that a sample of households with registered phone numbers, will result into under-coverage of households that have no registered phone number. This is an important source of selectivity, see for instance Cobben and Bethlehem (2005). The process of nonresponse of the sampled households involves four decision steps:

1. CATI-response;
2. Intention;
3. Selection;
4. Participation

Ad 1: The recruitment procedure starts with a sample of households drawn from the Dutch telephone registries. Contacted persons within sampled households are requested to cooperate to a three minute interview: *the recruitment survey*. Since at this stage the contacted person within the household may either cooperate or refuse, this decision can be seen as the first source of nonresponse.

Ad 2: If the contacted person decides to grant the interview request, the interview continues with questions on demographics and living conditions and ends with the (main) question: "Are you willing to participate in the Internet Panel?" Hard refusals to this question, i.e. the answer 'absolutely not', are the second source of nonresponse.

Ad 3: CentERdata copies the information on households without hard refusals into a pool of potential panel members. When there is a need to replace the households that dropped out of the Internet panel, CentERdata contacts respondents from this pool. A selection method is used that intends to keep the distribution of certain demographic variables of the panel close to the distribution of the population. These demographic variables are: region, urban character, household composition, age, monthly gross income and voting behavior at latest parliament elections (see Sikkel 2000, for details on this method). Note that this selection is carried out by CentERdata, so that this step does not yield nonresponse, but it does yield selectivity, since the selection process is not random.

Ad 4: Selected households receive a letter with information about CentERdata and are re-contacted for a second telephone interview. In this second interview contacted persons are asked whether they are prepared to make an appointment for installation of the software or hardware. Since respondents may reconsider their earlier intention to participate in the CentERpanel, a third source of nonresponse arises. The selection process is summarized in Figure 1. The time between the first and second contact varies from a few weeks to more than a half year, but is usually within three months.

We want to make two remarks here. The first is that CentERdata recently initiated another Internet panel for measurement and experimentation in the social sciences (the LISS panel, as part of the MESS project). The household sample will be drawn from the Municipal Basic Register and a doorstep method will be applied in the case the sampled household does not have a registered telephone (see Scherpenzeel (2007), Vis (2007)), thus overcoming the problem of non-coverage discussed earlier. The second remark concerns alternative ways to run an Internet panel based on probability sampling. Lee (2006) describes the situation of Knowledge Networks where panel members are not supposed to fill out a questionnaire every week, but may or may not be

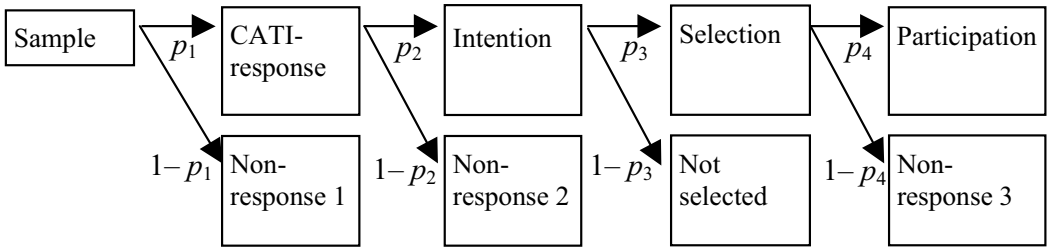

Figure 1. Selection process in recruiting respondents for the CentERpanel

(sub)sampled for a particular survey. She uses a similar scheme as Figure 1, but more detailed (more steps) and due to their approach in somewhat different order. With respect to nonresponse she compared the distributions of variables of the respondents of the Internet survey with those of the sample taken from the pool of Knowledge Networks (what she calls the 'total sample'). We make a similar comparison but compare the respondents that pass CentERdata's recruitment process with the initial sample of the recruitment survey (what we call 'the total sample').

Data

We will analyze data of the recruitment survey in the period 2001 until 2003. In this period in total 16,531 telephone numbers were sampled. We matched the sampled telephone numbers with the Municipal Basic Administration (in Dutch: Gemeentelijke Basis Administratie or GBA). The GBA contains data like date of birth, gender and address of all inhabitants registered. The link between the data of CentERdata and the GBA was carried out by matching on the address variables.

The match failed for 6.6% of all records, partly due to errors in the address variables and partly due to the fact that more than one household can be registered on the same address, for instance students sharing the same apartment. Records that could not be matched were deleted. In case of ambiguous matches one of the matching households was randomly selected, thus making the assumption that households living on the same address do not differ on the relevant characteristics. Information of the 'Social Statistics Database' (SSB)[1] was matched to the GBA and the data from CentERdata, by using identification numbers of persons. The SSB provides detailed information on social-economic variables of all persons registered in the Netherlands. It involves a number of separate data sets. We used data sets on:

- Municipality (such as province and degree of urbanization);
- Characteristics of the direct vicinity (in Dutch: postcodegebied), such as average house value (in Dutch: WOZ-waarde) and the percentage of non-native persons living in that area;
- Household income from the 'Income Panel Survey', a database derived from tax authority data.[2]

Again, for some records the matches failed , and another 227 records were deleted from the analysis. All in all we conclude that the matching procedure is a non-trivial task, making us loose almost 8% of the records. The loss is unfortunate since it may bias our result. Fortunately we found that the response patterns for the records that could not be matched was similar to the records that could be retained, so we expect that the bias in our results from deleting the unmatched records may be within acceptable limits.

The procedure described above resulted into a database of 15,213 records, being the basis of our analysis. Table 1 shows that for most of these telephone numbers (14,670) the survey request could be made, dividing these telephone numbers into respondents (9,179) or non-respondents due to refusal (5,493). A relatively small number of telephone numbers (542) was either never reached, gave problems of language or mental abilities or was for other reasons not usable. Since this group is relatively small we combined it with the refusals into a single group of non-respondents, thus ignoring different reasons for nonresponse (Nonresponse 1 in Table 1). In case of CATI-response we have the data from the recruitment survey, including answers to questions on: demographics, living conditions and, most important the intention to participate in the CentERpanel. In case of an intention (3,504) it is known whether that household was re-contacted by CentERdata.[3] In case of a recontact (2,844) we know the result of the final decision to participate in the CentERpanel.

In analyzing the response variable that indicates how far the sampled household ended up in the selection process, we will distinguish six blocks of variables (see Table 2). The first block refers to the 'paradata' of the CentERdata recruitment survey, which concern the time period of the interview only. We derived two variables from the time period of the interview: the time period of the day and the time period within the survey period measured. These variables were used by Vigderhous (1981) who showed that hour of the day

[1] See http://www.cbs.nl/nl-NL/menu/informatie/onderzoekers/ssb/ssb-info-medio-07.htm

[2] The household income is the sum of the personal income of the household members. The personal income is made up of: income from work, income from enterprise, benefits from income insurance and social benefits (except child allowances). This concept of income is not standardised (i.e. there are not any corrections for differences in household size and composition) and is not the same as the disposable income, which includes deductions for health insurance premiums and income tax.

[3] Since our data set is made at the end of 2004, we only consider re-contacts that are made before 2005.

Table 1: Results of the process of recruiting respondents for the CentERpanel

Step	Result of the selection step	*n*	Source
	Total number of telephone numbers	15,213	
Response	Response to the recruitment survey	9,179	
	Nonresponse due to refusal	5,493	} Nonresponse 1
	Nonresponse due to other reasons than refusal (never reached, problems of language or mental abilities, etc.)	581	
Intention	Intention to participate	3,504	
	Hard refusal on participation request	5,675	Nonresponse 2
Selection	Selected by CentERdata	2,844	
	Not selected by CentERdata	660	Not selected
Participation	Making an appointment to install hardware and software	1,420	
	Late refusal to participate	1,424	Nonresponse 3

Table 2: Six blocks of independent variables in modeling the recruitment process

Block	Source	cat/df*	Description of variables in the block; number of categories within parentheses
1	Contact time variables	16/14	Time period (hour) during the day (6), Time period (quarter) during survey period (10)
2	SSB variables	20/17	Household composition: age by gender (9), Household composition: ethnicity (3), Region (3), Urbanization (2), Income category (3)
3	Recruitment Survey	2/1	PC-ownership (2)
4	Recruitment Survey	28/14	housing: being a house owner (2), having a house with four or more rooms (2), visit cinemas regularly (2), visit theatres regularly (2), being a member of a sports club (2), having a paid job (2), having a relatively long travel time (2), being full time employed (2), were recently ill (2), having chronic complaints (2), high health rating (2), recently a victim of burglary (2), being afraid at home (2), being afraid in the street (2)
5	SSB variables	58/43	Family composition (8), Household size (5), Ethnicity (8), Region (12 provinces, 3 major cities), Urbanization (5), Income category (5), Average house value (7), Percentage non-natives in neighborhood (5)
6	Re-contact variables	7/5	Re-contact with same person (2), Number of weeks between recruitment survey and re-contact (5)

* cat = total number of categories in the block / df = degrees of freedom

and month of the year do matter in terms of response rate to a telephone survey whereas day of the week does not. Purdon et al. (1999) analyzed contact and response data in a face-to-face survey. They modeled the probability of contact by use of event history analysis techniques and found that weekday evenings are most successful times for making a first contact. Because in the recruitment survey callings attempts were all made on weekdays in late afternoons and evenings, we created six categories from time of the day, starting from 'between 16:00 and 17:00' up to 'between 21:00 and 22:00'. The months within year we aggregated to quarters but we did not aggregate over years. This allows us to detect a trend in response rates over the survey period. Since the survey period ended after the first half of 2003, this resulted into ten categories: 'first quarter of 2001' up to 'second quarter of 2003'.

The second block in Table 2 consists variables that were matched from the SSB. From the SSB we derived a set of 60 indicator variables (dummy variables). We select a subset of 17 of these indicators as 'the interesting indicators' into block 2 and the remaining 43 indicators into block 5. Later in this article we will motivate the (sub-) selection, where we show that these 17 indicators can explain the response process almost to the same level as the full set of 60 indicators.

The variables in blocks 3 and 4 concern information that was gathered in the recruitment survey. This survey contains a question on the presence of a personal computer. Since this variable is of special interest we stored it in block 3 separately. Block 4 contains variables on living conditions. The questions were formulated exactly in the same way as in the Integrated Survey on Living Conditions (POLS, in Dutch). We used all questions that were asked in the recruitment survey, but some POLS questions, e.g. questions on life satisfaction, were not asked in the recruitment survey. In order to reduce this information we dichotomized these variables in 'a sensible way', choosing groups of equal size.

Obviously, the information in blocks 3 and 4 is only available for the respondents of the recruitment survey. Missing values in block 4 do appear due to nonresponse and partly due to the design of the recruitment survey. According to the design not all respondents get all the questions on living conditions. Note that the missing data problem only refers to 'step 1: CATI-response' and only to blocks 3 and 4. To overcome the missing data problem in this part of the data set

we imputed these variables using a sequential hot deck imputation method, where we used the quarter within the survey period (ten categories), income (five categories) and family size (five categories) as sorting variables. The correctness of the imputation method relies on the MAR assumption of the model used in the hot deck imputation method, i.e. in a given sorting cell non-respondents and respondents are similar with respect to pc-ownership and living conditions.

Finally, block 6 contains two variables that are of interest to the final decision of participation. The first indicates whether the same household member is contacted in the first interview (the recruitment survey) and the second interview (after being selected by CentERdata).[4] One may expect that (final) agreement is more likely if the same person is interviewed, since that person may remember his spoken intention to cooperate from the first interview. The second variable is the time interval between the first and second interview. We expect that a positive decision on participation is more likely if the time interval is short, since it is more likely that the same household is still living on the address and that the contacted household member remembers his spoken intention to participate.

The six blocks of variables described above will be used in the order indicated in block-wise hierarchical regression models for each of the four steps of the selection process. Since we are interested in the selectivity with respect to respondent characteristics – not with respect to survey process characteristics – we will control for the latter group of variables by using these in the first block. The second block contains socio-demographic variables only, the third and fourth block contain key variables on living conditions. Block five contains a set of socio-demographic variables that will be shown to be irrelevant if we control for the variables previous blocks. Finally in block six we add variables related to the re-contacting process of respondents that earlier agreed to participate. Taking these variables into the analysis will give results that will help to improve the re-contacting process.

Results

Bivariate analysis

Above we divided the recruitment process of the CentERpanel up into four steps: CATI-response, Intention, Selection and Participation. Table 3 shows the conditional (response) rates for each of the four steps, which are the probabilities of the selection step given that the results of all preceding steps are positive. The response rate of the first step, i.e. the CATI recruitment survey, is 60% and a common response rate for telephone interviews in the Netherlands (see e.g. Bethlehem et. al (2006)). For the second step (Intention) we find that only 38% of the respondents to the recruitment survey were no hard refusals to the request stated to participate in the CentERpanel. Combining these first two steps results into 23% of the sampled households that end up in the file of potential panel households. Of all these potential households 81% is selected and re-contacted. About 50% of the re-contacted households agrees to participate, i.e. agrees to make an appointment to install software and hardware devices (if necessary) and to do the first profile interview. The 'total initial response rate' of entering the CentERpanel can be computed by multiplying the rates of CATI-response, Intention and Participation. Here, the selection step is left out, since it does not contribute to nonresponse.

Table 4 shows that there are only minor differences in the total response rates for different hours of the day, deviating only 1 or 2 percent from the total sample rate of 12%. However, the differences are more profound for the first step, the CATI-response (p_1), than for the remaining steps. The best time appears to be between 16:00 and 17:00. This result is slightly different from the finding of Vigderhous (1981) who stated that the best time is between 18:00 and 19:00. Note that in our data the hour of the day is not necessarily when the first contact was made, but is the time point of the interview. With respect to the quarter in which the interview took place, we find substantial differences: the 'total response rate $p_1 \times p_2 \times p_4$' ranges from 5 percent point below the total sample rate of 12% to 10 percent points above that level, although the latter refers to only 1% of the total sample. The decomposition into the selection steps shows that these differences are only partly due to the CATI-response rate. The largest differences are found in the intention step, where the rates range from 25% in the first quarter of 2001 to 56% in the fourth quarter of 2001. Intention rates and CATI-response rates roughly show the same pattern over time, having a peak at the fourth quarter of 2001. The picture of the participation rates over time is different: after the first quarter of 2001 they are remarkably higher than before. It is not clear what caused these large variations. The series of intention rates do not reveal a seasonal pattern. The large variations in intention rates may be a result of differences in monitoring and training of the interviewers or of over-sampling difficult groups (elderly), but this is mere speculation.

If we look at 'total response rate $p_1 \times p_2 \times p_4$' across the presence of household members of specific age and gender (see Table 5) we find very low total response rates for elderly (65+), being 5% or less. This result is not very different for elderly men and women: the total response rate ($p_1 \times p_2 \times p_4$) is 12 percent with a deviation of -7 percent for elderly men and -8 percent for elderly women. The decomposition into CATI-response, intention and participation shows that these groups have low rates in all three steps, especially in the intention and participation steps. CentERdata corrects this selectivity in the selection step, but only partial. With respect to ethnicity Table 5 hardly shows any total effect: the lower CATI-response rate, selection rate and participation rates are compensated by a higher intention rate. Some regional effects are shown in Table 5: the northern part of the Netherlands shows lower total response rates than the other parts, caused by relatively low response, intention and participation rates, which are not offset by a higher selection rate. Finally, it can be concluded from Table 5 that there are large differences with respect to household income groups:

[4] We assume that the same person has been contacted if gender and age are the same in the first and second interview

Table 3: Success rates at the successive steps of the selection process

	CATI Response	Intention	Selection	Participation	Total
	p_1	p_2	p_3	p_4	$p_1 x p_2 x p_4$
total sample	.60	.38	.81	.50	.12

Table 4: Subgroup deviations of success rates for contact time variables[a,b]

	CATI Response	Intention	Selection	Participation	Total	
	p_1	p_2	p_3	p_4	$p_1 x p_2 x p_4$	Proportion
total sample	.60	.38	.81	.50	.12	1.00
time of day of interview						
between 16:00 and 17:00	.14**	.03	.09**	.06*	.02	.06
between 17:00 and 18:00	.03**	.00	.01	.03	.00	.12
between 18:00 and 19:00	.00	.01	.00	.02	.01	.27
between 19:00 and 20:00	.02*	.01	.01	.01	.00	.23
between 20:00 and 21:00	.02**	.01	.00	.01	.01	.23
between 21:00 and 22:00	.03**	.04**	.04	.00	.02*	.09
time of survey period of interview						
first quarter of 2001	.04**	.13**	.10**	.02	.04**	.10
second quarter of 2001	.00	.08**	.02	.14**	.05**	.15
third quarter of 2001	.07**	.06**	.11**	.11**	.00	.13
fourth quarter of 2001	.13**	.18**	.06**	.04*	.07**	.10
first quarter of 2002	.03	.14**	.05	.14*	.10**	.01
second quarter of 2002	.04**	.00	.13**	.16**	.05**	.06
third quarter of 2002	.06**	.02	.08*	.13**	.01	.06
fourth quarter of 2002	.01	.02	.13**	.13**	.04**	.10
first quarter of 2003	.05**	.04**	.02	.16**	.01	.11
second quarter of 2003	.08**	.02	.07**	.04	.01*	.17

* $p < .05$

** $p < .01$

[a] Significance levels refer to difference between the rate of the subgroup and the rate of the total sample and were obtained by an application of the bootstrap method.

[b] The time in the table refers to the first interview (this also holds for Participation p_4).

the lowest income group has a total response rate of only 6%, while the highest income group shows a rate of 16%. Notice that all three steps of recruitment process contribute to the relatively low or high overall rate and that the selection step compensates for some of this selectivity.

Table 6 shows that we find considerable differences in 'total initial response rate' for pc-ownership: 16% for pc-owners versus 5% for non-pc-owners. This result shows that the strategy of providing non-pc-owners with the infrastructure only works partially: it is not as bad as ignoring this group – which would be equivalent to a total response rate of 0% – but the total rate of non-pc-owners is far from the 16% of pc-owners. The decomposition into the successive steps shows that the low total response rate for non-pc-owners is not so much a result of a lower response rate to the recruitment survey, but of lower intention and participation rates. This means that non-pc-owners are at least contacted almost at a similar rate as the pc-owners. This allows the survey organization to put special efforts to convince the non-pc-owners to participate in the Internet panel. However, this result heavily relies on the correctness of the imputation model for pc-ownership. The decomposition also shows that the lower probability for non-pc-owners of ending up in the file of potential panel members (.13 versus .23 for the total sample, obtained by multiplication of the first two steps) is only slightly compensated in the selection stage (.87 versus .81 for the total sample), which means that there is a slight relation between the demographic variables CentERdata uses for selection and pc-ownership. With respect to living conditions the result is impressive in a different way: the general picture is that the total response rates between groups differ significantly from a statistical point of view, but at the same

Table 5: Subgroup deviations in success rates for SSB variables[a,b]

	CATI Response	Intention	Selection	Participation	Total	
	p_1	p_2	p_3	p_4	$p_1 x p_2 x p_4$	Proportion
total sample	.60	.38	.81	.50	.12	1.00
household composition: presence of						
child (age ≤ 18)	.06**	.08**	-.03**	.03*	.04**	.29
man between 19 and 29	.07**	.06**	.05**	-.02	.03**	.22
woman between 19 and 29	.08**	.07**	.04**	-.01	.04**	.24
man between 30 and 44	.04**	.06**	-.07**	.06**	.04**	.23
woman between 30 and 44	.06**	.06**	-.07**	.06**	.05**	.21
man between 45 and 64	.00	.01	-.01	.02	.01	.24
woman between 45 and 64	-.01	-.01	.00	.00	-.01	.26
man older than 64	-.08**	-.14**	.08**	-.12**	-.07**	.19
woman older than 64	-.12**	-.18**	.08**	-.18**	-.08**	.24
household composition: presence of						
native person	.00**	.00	.00	.00*	.00**	.95
non-native western person	-.06*	.09*	-.06	.07	-.01	.13
non-native non-western person	.01	.04**	-.01	-.05*	.00	.02
region						
northern part	-.03**	-.02	.01	-.06*	-.02**	.14
southern part	.04**	.01	.01	.00	.01*	.22
remaining	-.01**	.00	-.01	.01*	.00	.64
urbanization						
urban (≥1000 adresses/km^2)	.00	.02**	-.01*	.02**	.01**	.46
rural (<1000 adresses/km^2)	.00	-.02**	.02*	-.03**	-.01**	.54
household income						
low (first quintile)	-.09**	-.10**	.04*	-.10**	-.06**	.20
medium	.02**	.00	.01	.00	.00	.60
high (last quintile)	.04**	.06**	-.04**	.05**	.04**	.20

* $p < .05$,

** $p < .01$.

[a] Significance levels refer to difference between the rate of the subgroup and the rate of the total sample and were obtained by an application of the bootstrap method.

[b] The categories with respect to household composition may overlap: for instance one household may involve one or more men between 30 and 44 and a woman between 45 and 64.

time only moderately in a practical sense. The decomposition into the successive steps shows the same picture. The lowest total response rate is still 9% for not regularly visiting cinemas (this figure is not in Table 6 since it only shows the 'yes-part').

Table 7 shows that the time between first and second contact significantly affects the participation rate: it is 60% for a time interval up to 10 weeks, but only 40% for a span of 15 weeks or more. This outcome suggests that the survey organization should not wait long to re-contact a household. We find an even larger effect from the fact of whether in the second interview the same person was contacted as in the first interview. In case of contact with the same person – which happened in about 69% of the cases – the participation rate was 63%, while the participation rate for other persons is as low as 24%! Obviously this outcome shows that the survey organization should put more efforts in re-contacting the same person.

Multivariate analysis

The effects of the four sets (blocks) of variables listed in Table 4, 5, 6 and 7 on each of the four cooperation steps of the recruitment process may overlap. For instance, it is known that in the Netherlands the elderly are more likely to belong to lower income groups and are less likely to own a PC. If we want to study the *relative importance* for every single variable we need a multivariate approach. Therefore, for each of the four steps in the selection process we estimated hierarchical logistic regression models with six blocks of predictors.

The coefficients of model fit together with the omnibus

Table 6: Subgroup deviations of success rates pc-ownership and living conditions[a]

	CATI Response	Intention	Selection	Participation	Total	
	p_1	p_2	p_3	p_4	$p_1 x p_2 x p_4$	Proportion
total sample	.60	.38	.81	.50	.12	1.00
owning a pc[b]						
yes	.02**	.07**	-.02**	.05**	.04**	.66
no	-.03**	-.15**	.06**	-.19**	-.07**	.34
living conditions[b]						
housing: ownership	.01**	.01	-.01*	.01	.01**	.60
housing: four or more rooms	.01**	.01**	.00	-.01	.00**	.70
activities: visit cinema	.02**	.04**	.00	.03**	.03**	.50
activities: visit theatre	.01	.03**	-.01	.01	.01**	.42
activities: sports club member	.01*	.03**	-.02*	.02*	.02**	.48
work: paid job	.02**	.04**	-.02**	.03**	.02**	.48
work: long travel time	.04**	.04**	-.02	.07**	.04**	.21
work: hours employed	.03**	.04**	-.02*	.00	.02**	.36
health: recently ill	.02*	.03**	.01	.00	.01	.19
health: chronic complaints	-.02**	-.03**	.00	-.02	-.02**	.25
health: rating	.01**	.00	.00	.00	.00*	.86
safety: burglary	.00	.01	-.03*	.02	.01	.17
safety: afraid at home	-.02	.03	.01	-.06	-.01	.08
safety: afraid in the street	.00	.01	.00	-.06*	-.01	.14

* $p < .05$,

** $p < .01$.

[a] Significance levels refer to the difference between the rate of the subgroup and the rate of the total sample. These were obtained by an application of the bootstrap method.

[b] For PC ownership the model for response was estimated on imputed data; for living conditions all models were estimated on imputed data (see text). For living conditions only the 'yes-part' is shown to save space.

tests of the logistic regression models in Table 8 show that the logistic regression models do not fit that well: R^2_{cs}, the Cox and Snell's pseudo measure of R^2 and a crude measure of model fit, ranges from 0.05 to 0.11, which is rather low. This should make us humble in making statements about the extent to which we can explain participation to the CentER-panel from the available data. Only by adding the re-contact variables R^2 obtained a value of 0.20. Another conclusion that can be drawn from Table 8 is that the effect of the SSB variables is captured quite well by the 17 selected variables since the addition of the remaining SSB variables in block 5 only shows a small and insignificant increase in Cox and Snell's R^2 (with only an exception in '*step 1: CATI-response*' where the p-value = .043). Furthermore, Table 8 confirms our earlier finding of a very small contribution of living condition variables in explaining the several steps in the response process. For example: only in '*step 2: Intention*' the model fit increases significantly although with a tiny increase in R^2_{cs} of 0.002. Finally, Table 8 shows that the re-contact variables are very important in explaining '*step 4: Participation*': adding these variables to the model almost doubles the R^2_{cs}!

Estimated parameters of the logistic regression models are shown in Table 9. These parameters are estimated using weighted effects coding of the categorical variables (see Cohen et al. 2003). The estimates show controlled effects of 'belonging to a certain category of a categorical variable' on the success of passing a step in the selection process compared to a case with average sample characteristics on the same categorical variable. For example, in the model for '*step 1: CATI-response*' the parameter of the indicator variable '*household income: low (first quintile)*' has a value of –.234**. This value represents the log-odds-ratio of this quintile versus the average income, while keeping the other categorical variables in the model (contact time, household composition, region, income, living conditions, etc.) constant. The average corresponds to: 0.2 for the indicator variable '*low (first quintile)*', 0.6 for '*medium*' and 0.2 for '*high (last quintile)*'.

Table 9 shows controlled effects (and significance levels) of the independent variables on the success of passing each of the four steps in the selection process. However, we are particularly interested in studying the effects on the 'total initial response rate' which is a combination of the first, second and fourth step. Therefore we introduce the notion of '*isolated subgroup differences of response rates*' that are similar to the 'subgroup differences of response rates' in Tables 4, 5 and 6, except that the 'subgroup differences' show uncontrolled effects while the 'isolated subgroup differences' show

Table 8: Coefficients of model fit and omnibus tests of logistic regression models[a]

				model			*block*		
		R^2_{CS}	R^2_N	χ^2	df	sig	χ^2	df	sig
block	*step 1: CATI -response*								
1	Contact time variables	.021	.029	327.9	14	<.001	327.9	14	<.001
2	SSB variables	.045	.061	704.9	31	<.001	377.0	17	<.001
3	PC ownership	.046	.062	708.8	32	<.001	3.9	1	.049
4	Living conditions	.046	.063	723.3	46	<.001	14.4	14	.417
5	Remaining SSB	.050	.068	783.3	89	<.001	60.7	43	.043
block	*step 2: Intention*								
1	Contact time variables	.033	.045	306.6	14	<.001	306.6	14	<.001
2	SSB variables	.065	.088	612.6	31	<.001	306.0	17	<.001
3	PC ownership	.082	.111	781.7	32	<.001	169.1	1	<.001
4	Living conditions	.084	.114	808.6	46	<.001	26.9	14	.020
5	Remaining SSB	.090	.122	866.4	89	<.001	57.8	43	.065
block	*step 3: Selection*								
1	Contact time variables	.052	.083	185.5	14	<.001	185.5	14	<.001
2	SSB variables	.067	.108	242.3	31	<.001	56.8	17	<.001
3	PC ownership	.068	.109	245.6	32	<.001	3.2	1	.072
4	Living conditions	.072	.115	260.1	46	<.001	14.5	14	.413
5	Remaining SSB	.083	.135	305.4	89	<.001	45.3	43	.376
block	*step 4: Participation*								
1	Contact time variables	.047	.063	137.1	14	<.001	137.1	14	<.001
2	SSB variables	.072	.096	212.9	31	<.001	75.8	17	<.001
3	PC ownership	.090	.120	267.3	32	<.001	54.4	1	<.001
4	Living conditions	.095	.127	284.1	46	<.001	16.8	14	.266
5	Remaining SSB	.110	.146	330.6	89	<.001	46.5	43	.329
6	Re-contact variables	.204	.272	648.2	95	<.001	317.6	6	<.001

[a] The coefficients of model fit and omnibus tests concern the logistic regression models that predict the likelihood of making a successful step (CATI-response, intention, selection or participation) in the recruitment process. The results of block i show the outcomes of the models that have block 1 up to i as explaining variables.

Table 7: Subgroup deviations of participation (p_4) for re-contact variables

	Participation p_4	proportion
total sample	.50	1.00
time between first and second contact		
5 weeks or less	.07**	.26
between 6 and 10 weeks	.10**	.21
between 11 and 15 weeks	.01	.14
between 16 and 25 weeks	-.09**	.19
more than 25 weeks	-.13**	.20
contact with same household member		
yes	.12**	.69
no	-.26**	.31

* $p < .05$,
** $p < .01$, the significance levels refer to the difference between the rate of the subgroup and the rate of total sample. These were obtained by an application of a bootstrap method.

controlled effects. We define the 'isolated subgroup differences of response rates' for each category of a categorical variable (category or subgroup represented by an indicator variable χ_k) and for each step s in the selection process (s= 1, 2, 3, 4) as the difference between the probabilities $p_k^{(s)}$ and $p_0^{(s)}$ where $p_k^{(s)}$ stands for the predicted probability of passing selection step s if the household belongs to the category k, while all other categorical variables in the model are fixed to the average value of the CATI-sample. The probability $p_0^{(s)}$ is evaluated at the mean values on all characteristics. For example the probability $p_0^{(2)}$ of passing '*step 2: Intention*' is .37, while the 'isolated subgroup difference' $p_k^{(2)} - p_0^{(2)}$ of the group 'pc-owners' is .06. This number indicates – that keeping all other variables fixed – pc-owners have a higher probability (higher by .06) than persons with average scores on all characteristics.

For the 'total initial response rate' the 'isolated subgroup difference' can then be computed as the difference between $p_k^1 \times p_k^2 \times p_k^4$ and $p_0^1 \times p_0^2 \times p_0^4$. The results are shown in the last column of Table 10. Significance levels are obtained by an application of the bootstrap method. Note that the significance levels in Table 9 and Table 10 are (almost) the same.

Comparing the uncontrolled effects of the independent variables on the selection process (Table 4, 5, 6 and 7) with the controlled effects (Table 10) shows the common phenomenon of partial redundancy: controlled effects are generally smaller than uncontrolled effects due to an overlap of effects. For example, for '*step 1: CATI-response*' Table 5 shows the uncontrolled difference in response probability between the lowest income group (–.09) and the highest income

Table 9: Estimated coefficients of logistic regression model[a]

	CATI-Response	Intention	Selection	Participation
time point of interview				
between 16:00 and 17:00	.420**	-.167*	-.001	.080
between 17:00 and 18:00	.130**	-.007	-.057	-.118
between 18:00 and 19:00	-.006	.059	.041	.047
between 19:00 and 20:00	-.021	.110**	.007	-.012
between 20:00 and 21:00	-.085**	-.019	.053	.002
between 21:00 and 22:00	-.154**	-.274**	-.226	-.046
time point of interview				
first quarter of 2001	.001	-.294**	.601*	.294*
second quarter of 2001	.125**	-.124*	-.265	-.407**
third quarter of 2001	.195**	.283**	.867**	-.472**
fourth quarter of 2001	.380**	.603**	.425**	-.209*
first quarter of 2002	-.075	.324	-.361	.448
second quarter of 2002	-.026	-.220*	-.749**	.517**
third quarter of 2002	-.172*	-.120	-.426*	.474*
fourth quarter of 2002	.050	.032	-.737**	.461**
first quarter of 2003	-.206**	-.276**	-.143	.599**
second quarter of 2003	-.307**	-.179**	-.439**	.097
household composition: presence of				
child (age ≤ 18)	.027	.103*	-.035	.062
man between 19 and 29	-.058	-.152**	.171	-.085
woman between 19 and 29	.046	-.054	.159	-.194*
man between 30 and 44	-.082	-.103	-.133	-.053
woman between 30 and 44	.053	-.155*	.004	-.292*
man between 45 and 64	-.111*	-.148*	-.048	-.044
woman between 45 and 64	-.149**	-.179**	.149	-.359**
man older than 64	-.107*	-.153*	.433*	-.054
woman older than 64	-.471**	-.543**	.318	-.671**
household composition: presence of				
native person	.016**	.005	-.007	.016
non-native: western person	-.239*	.224	-.395	-.142
non-native : non-western person	.072	.158*	-.073	-.232*
region				
northern part	-.137**	-.044	-.104	-.196
southern part	.185**	.090*	.077	.016
remaining parts	-.033*	-.024	-.008	.034
urbanization				
urban (≥ 1000 adresses/km^2)	.011	.077**	-.091*	.064
rural (<1000 adresses/km^2)	-.012	-.089**	.117*	-.079
household income				
low (first quintile)	-.234**	-.135*	.149	-.248*
medium	.030*	-.011	.030	.011
high (last quintile)	.136**	.132**	-.147	-.079
owning a pc[b]				
yes	-.029*	.236**	-.044	.164**
no	.057*	-.501**	.182	.106

* $p < .05$,

** $p < .01$.

[a] The logistic regression models predict the likelihood of making a successful step (CATIresponse, intention, selection or participation) in the recruitment process. Significance levels are obtained from Wald test statistics on a single predictor. The probability p_0 is the predicted probability evaluated at mean characteristics: $p_0 = 1/(1 + exp(-b_0))$, where b_0 is the constant in the model.

[b] For PC ownership the model for response was estimated on imputed data; for living conditions all models were estimated on imputed data (see text).

Table 9: Continued[a]

	CATI-Response	Intention	Selection	Participation
living conditions[b]				
housing: ownership	-.003	-.022	.005	.003
housing: four or more rooms	.015	.015	-.012	-.016
activities: visit cinema	.009	.042	.065	.060
activities: visit theatre	.013	.055	-.056	-.020
activities: sports club member	-.020	-.004	-.015	-.025
work: paid job	-.021	.030	-.019	.016
work: long travel time	.033	-.043	-.047	.207**
work: hours employed	.005	-.023	-.111	-.058
health: recently ill	.078*	.105*	.048	-.048
health: chronic complaints	-.001	-.024	-.134	.082
health: rating	.012	-.003	-.015	-.006
safety: burglary	.059	.099*	-.088	-.013
safety: afraid at home	-.059	.170*	-.226	-.001
safety: afraid in the street	-.014	.010	-.065	-.191
constant	.441**	-.538**	1.630**	-.012
p_0	.608	.369	.836	.497
R^2 Cox and Snell	.046	.084	.072	.095

* $p < .05$,

** $p < .01$.

[a] The logistic regression models predict the likelihood of making a successful step (CATIresponse, intention, selection or participation) in the recruitment process. Significance levels are obtained from Wald test statistics on a single predictor. The probability p0 is the predicted probability evaluated at mean characteristics: $p_0 = 1/(1 + exp(-b_0))$, where b_0 is the constant in the model.

[b] For PC ownership the model for response was estimated on imputed data; for living conditions all models were estimated on imputed data (see text).

group (+.04) is .13, while according to Table 10 the controlled difference between the lowest income group (–.06) and the highest income group (+.03) is only .09, meaning that a part (but not all) of the differences between income groups can be contributed to other variables in the model!

The results for the 'total initial response rate' are shown in the sixth column (i.e. the last column) of Table 10. For example: the isolated subgroup differences of 'the quarter of the interview' in Table 10 range from .08 (= .11 – .03) to .16 (= .11 + .05), while the uncontrolled effects vary more: from .07 to .22 (see Table 4). This means that differences in 'total initial response rates' among different quarters can partly be explained by differences in other variables (age, gender and ethnicity of the household composition, region, income, etc.). Still, after controlling for these variables, some of the differences between quarters remain. We can summarize the selectivity in 'total initial response rates' with respect to the other variables (subgroups) in the model by saying that the (controlled) effects of household composition (presence of elderly women), income and pc-ownership remain, while the effects of the living condition variables almost completely disappear.

Conclusion and Discussion

This paper provides valuable insight into the amount and the selectivity of nonresponse in an Internet panel (in this case the CentERpanel) that is founded on a probability-based recruitment survey. The recruitment process of this panel starts with a sample of telephone numbers and then involves four separate steps: 1. response to the recruitment survey, 2. giving an intention to participate in the Internet panel, 3. being selected by the survey organization to be in the panel and 4. making an appointment to install the hardware and/or software to fill out a first survey. Steps 1, 2 and 4 are of special interest since these steps concern decisions of (individuals within) the sampled households to end up in the Internet panel. Multiplication of the response rates corresponding to these three steps defines the 'total initial response rate' which comes to 11.5% for the CentERpanel.

Matching the sampled telephone numbers to the Municipal Basic Administration (GBA) and to the 'Social Statistics Database' (SSB) together with the questions on living conditions in the recruitment survey gives us a rich data set and a unique opportunity to study the selectivity with respect to a large set of interesting variables. One general conclusion is that we find selectivity with respect to age and income. Households with older persons and low incomes have much lower probabilities to end up in the panel than households without older persons and in higher income groups. It is remarkable that the selectivity takes place in all three decision steps in the recruitment process. We also found selectivity with respect to pc ownership due to intention and final par-

Table 10: Isolated subgroup deviations of response rates derived from logistic regression models[a]

	CATI-Response	Intention	Selection	Participation	Total
total sample	.61	.37	.84	.50	.11
time point of interview					
between 16:00 and 17:00	.09**	-.04*	.00	.02	.00
between 18:00 and 19:00	.00	.01	.01	.01	.01
between 19:00 and 20:00	-.01	.03**	.00	.00	.01
between 20:00 and 21:00	-.02**	.00	.01	.00	-.01
between 21:00 and 22:00	-.04**	-.06**	-.03	-.01	-.03**
time point of interview					
first quarter of 2001	.00	-.07**	.07**	.07	-.01
second quarter of 2001	.03**	-.03*	-.04	-.10**	-.03**
third quarter of 2001	.05**	.07**	.09**	-.12**	.00
fourth quarter of 2001	.09**	.15**	.05**	-.05**	.05**
first quarter of 2002	-.02	.08	-.06	.11	.05
second quarter of 2002	-.01	-.05**	-.13**	.13**	.01
third quarter of 2002	-.04**	-.03	-.07*	.12**	.01
fourth quarter of 2002	.01	.01	-.13**	.11**	.03**
first quarter of 2003	-.05**	-.06**	-.02	.15**	.00
second quarter of 2003	-.08**	-.04**	-.07**	.02	-.02**
household composition: presence of					
child (age ≤ 18)	.01	.02*	-.01	.02	.01*
man between 19 and 29	-.01	-.04**	.02	-.02	-.02**
woman between 19 and 29	.01	-.01	.02	-.05*	-.01*
man between 30 and 44	-.02	-.02	-.02	-.01	-.01
woman between 30 and 44	.01	-.04*	.00	-.07*	-.02**
man between 45 and 64	-.03*	-.03*	-.01	-.01	-.02**
woman between 45 and 64	-.04**	-.04**	.02	-.09**	-.04**
man older than 64	-.03*	-.04	.05*	-.01	-.02*
woman older than 64	-.12**	-.12**	.04	-.16**	-.07**
household composition: presence of					
native person	.00**	.00	.00	.00	.00*
non-native; western person	-.06*	.05	-.06	-.04	.00
non-native; non-western person	.02	.04	-.01	-.06	.00
region					
northern part	-.03**	-.01	-.02	-.05	-.02**
southern part	.04**	.02*	.01	.00	.02**
remaining parts	-.01**	-.01	.00	.01	.00
urbanization					
urban (≥1000 adresses/km^2)	.00	.02*	-.01*	.02	.01**
rural (<1000 adresses/km^2)	.00	-.02*	.02*	-.02	-.01**
household income					
low (first quintile)	-.06**	-.03*	.02	-.06*	-.03**
medium	.01*	.00	.00	.00	.00
high (last quintile)	.03**	.03**	-.02	.03	.02**
owning a pc[b]					
yes	-.01*	.06**	-.01	.04**	.03**
no	.01*	-.11**	.02	-.15**	-.06**

* $p < .05$,

** $p < .01$.

[a] Significance levels refer to the difference between the rate of the subgroup and the rate of the total sample. These were obtained by an application of the bootstrap method.

[b] For PC ownership the model for response was estimated on imputed data; for living conditions all models were estimated on imputed data (see text).

Table 10: Continued[a]

	CATI-Response	Intention	Selection	Participation	Total
total sample	.61	.37	.84	.50	.11
living conditions[b]					
housing: ownership	.00	-.01	.00	.00	.00
housing: four or more rooms	.00	.00	.00	.00	.00
activities: visit cinema	.00	.01	.01	.02	.01*
activities: visit theatre	.00	.01	-.01	-.01	.00
activities: sports club member	-.01	.00	.00	-.01	.00
work: paid job	-.01	.01	.00	.00	.00
work: long travel time	.01	-.01	-.01	.05*	.01
work: hours employed	.00	-.01	-.02	-.02	-.01
health: recently ill	.02*	.03*	.01	-.01	.01
health: chronic complaints	.00	-.01	-.02	.02	.00
health: rating	.00	.00	.00	.00	.00
safety: burglary	.01	.02	-.01	.00	.01
safety: afraid at home	-.01	.04*	-.03	.00	.01
safety: afraid in the street	.00	.00	-.01	-.05	-.01

* $p < .05$,

** $p < .01$.

[a] The logistic regression models predict the likelihood of making a successful step (CATIresponse, intention, selection or participation) in the recruitment process. Significance levels are obtained from Wald test statistics on a single predictor. The probability p0 is the predicted probability evaluated at mean characteristics: $p_0 = 1/(1 + exp(-b_0))$, where b_0 is the constant in the model.

[b] For PC ownership the model for response was estimated on imputed data; for living conditions all models were estimated on imputed data (see text).

ticipation. On the other hand we found little selectivity with respect to the presence of non-native persons. Although this finding may come as a surprise, we should keep in mind that this non-selectivity refers to sampled households that have a registered phone number so that we started with a selective subgroup here. Another interesting finding is that we hardly found any selectivity with respect to the variables on living conditions for this sample.

We found large effects from the re-contact variables: they turn out to be good predictors in explaining the final participation:

1. The more time elapses between the first and second interview, the less households will grant the participation request. Thus, our advice to the survey organization is not to wait that long to re-contact a household
2. A household is less inclined to participate to the panel, if different members of the household are re-contacted. Therefore our advice is to re-contact the same member, although this will be not so easy.

There are two positive messages from this study: 1. Although we found severe selectivity with respect to age and income, we also found that this selectivity is not all due to the first step of the recruitment process. Most households with older persons and within the low income group at least responded to the recruitment survey. It is a challenge to develop strategies to motivate the respondents of the recruitment survey – especially those that are in the low income group or in the higher age group – not to drop out of the recruitment process, thus increasing the success rate in the successive decision steps. 2. We found very little selectivity with respect to variables on living conditions.

However there are also some limitations to this study. The initial nonresponse described in this paper is only a part of the total nonresponse of the CentERpanel. We did not study the selectivity of panel attrition and wave nonresponse. This can be a topic of further research. It would also be interesting to study the effects of the interviewer selection to the final response rate. Does it matter much which interviewer contacts a household? Does it matter whether the same interviewer makes a re-contact? Unfortunately we had no information on the interviewer.

It is important to have knowledge of the role of undercoverage in this analysis. The results in this paper refer to a sample of households with a registered telephone number. The sample design has an explicit problem of undercoverage. However, we can ask ourselves if the supply of a PC or Internet connection solves the problem of undercoverage of households that do not own a PC or Internet connection within this restricted population of households with a registered telephone number. We find that the 'total initial response rate' of non-pc-owners is about 5% and still far from the 16% of pc-owners. We also find that if we control for age and income and the other variables in the multivariate models the difference in 'total initial response rate' between non-pc-owners and pc-owners remains substantial: 14% for pc-owners and 5% for non-pc owners. Still, this is not as bad as ignoring the non-pc owners – which corresponds to a rate of 0% – but we should be aware of the limits of this remedy.

Acknowledgements

The authors would like to thank Jelke Bethlehem and Hans Schmeets for their suggestions and Rainer Schnell and two anonymous referees for their constructive comments.

References

Bethlehem, J. (2007). *Reducing the Bias of Web Survey Based Estimates*. Voorburg: Statistics Netherlands.

Bethlehem, J., Cobben, F., & Schouten, B. (2006). *Nonresponse in Household Surveys*. Voorburg: Statistics Netherlands.

Cobben, F. (2005). *Adjusting undercoverage and nonresponse bias in telephone surveys.* Statistics Netherlands. Discussion paper 05006. http://www.cbs.nl/NR/rdonlyres/7FD00F42-15A3-4151-9DAA-2D54566CF59A/0/200506x10pub.pdf.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/ correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly*, *64*, 464-494.

Couper, M. P., Kapteyn, A., Schonlau, M., & Winter, J. (2007). Noncoverage and nonresponse in an Internet survey. *Social Science Research*, *36*, 131-148.

Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method* (2nd ed.). New York, NY: John Wiley & Sons.

Hoogendoorn, A. W., Sikkel, D., & Weerman, B. (2000). *The Internet, Blaise and a Representative Panel.* Proceedings of the 6th International Blaise Users Conference, Central Statistical Office, Cork, Ireland.

Huggins, V., & Eyerman, J. (2001). *Probability Based Internet Surveys: A Synopsis of Early Methods and Survey Research Results.* Paper presented at the November 14-16, 2001 Research Conference of the Federal Committee on Statistical Methodology.

Lee, S. (2006a). An Evaluation of Nonresponse and Coverage Errors in a Prerecruited Probability Web Panel Survey. *Social Science Computer Review*, *24*(4), 460-475.

Lee, S. (2006b). Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys. *Journal of Official Statistics*, *22*(2), 329-349.

Purdon, S., Campanelli, P., & Sturgis, P. (1999). Interviewers' Calling Strategies on Face-to-Face Interview Surveys. *Journal of Official Statistics*, *15*(2), 199-216.

Saris, W. E. (1998). Ten Years of Interviewing Without Interviewers: The Telepanel. In M. P. Couper et al. (Eds.), *Computer Assisted Survey Information Collection* (p. 409-429). New York: Wiley.

Scherpenzeel, A. (2007). *Recruiting a Sample for an Online Panel: Effects of Contact Mode, Incentives and Information Letter.* Paper presented at ESRA conference 2007, June 25-29, Prague, Czech Republic.

Schonlau, M., van Soest, A., & Kapteyn, A. (2007). Are "Webographic" or attitudinal questions useful for adjusting estimates from Web surveys using propensity scoring? *Survey Research Methods*, *1*(3), 155-163.

Schonlau, M., Zapert, K., Simon, L. P., Sanstad, K., Marcus, S., Adams, J., et al. (2004). A Comparison Between Responses From a Propensity-Weighted Web Survey and an Identical RDD Survey. *Social Science Computer Review*, *21*(1), 128-138.

Sikkel, D. (2000). *CentERpanel, representativiteit en weging*. Tilburg University, Tilburg: Internal CentERdata Report.

Smith, T. W. (2003). An Experimental Comparison of Knowledge Networks and the GSS. *International Journal of Public Opinion Research*, *15*(2), 167-179.

Terhanian, G., Smith, R., Bremer, J., & Thomas, R. K. (2001). Exploiting analytical advances: Minimizing the Biases Associated with Internet-based Surveys of Non-random Samples. *ARF/ESOMAR: Worldwide Online Measurement*, *248*, 247-272.

Vigderhous, G. (1981). Scheduling Telephone Interviews: A study of Seasonals Patterns. *The public opinion quarterly*, *45*, 250-259.

Vis, C. (2007). *Who Wants to Become an Online Panel Member? Effects of Respondent Characteristics on the Recruitment of an Online Panel.* Paper presented at ESRA conference 2007, June 25-29, Prague, Czech Republic.