

Article

Symposium 2008:
Data Collection: Challenges, Achievements and New Directions

Can We Make Official Statistics with Self-Selection Web Surveys?

by Jelke Bethlehem

2009



Can We Make Official Statistics with Self-Selection Web Surveys?

Jelke Bethlehem¹

Abstract

At first sight, web surveys seem to be an interesting and attractive means of data collection. They provide simple, cheap and fast access to a large group of people. However, web surveys also suffer from methodological problems. Outcomes of web surveys may be severely biased, particularly if self-selection of respondents is applied instead of proper probability sampling. Under-coverage is also a serious problem. This raises the question whether web surveys can be used for data collection in official statistics. This paper addresses the problems under-coverage and self-selection in web surveys, and attempts to describe how Internet data collection can be incorporated in normal data collection practices of official statistics.

Key Words: Web survey, Self-selection, Under-coverage, Probability sampling, Mixed-mode survey.

1. Introduction

The survey research landscape has undergone radical changes over the last decades. First, there was the change from traditional paper and pencil interviewing to computer-assisted interviewing. And now, particularly in commercial market research, face-to-face, mail and telephone surveys are increasingly replaced by web surveys. The popularity of online research is not surprising. A web survey is a simple means to get access to a large group of people. Questionnaires can be distributed at very low costs. No interviewers are needed, and there are no mailing and printing costs. Surveys can be launched very quickly. Little time is lost between the moment the questionnaire is ready and the start of the fieldwork. And web surveys offer new, attractive possibilities, such as the use of multimedia (sound, pictures, animation and movies).

At first sight, online surveys seem to have much in common with other types of surveys. It is just another mode of data collection. Questions are not asked face-to-face or by telephone, but over the Internet. What is different for many online surveys, however, is that the principles of probability sampling have not been applied. By selecting random samples probability theory can be applied, making it possible to compute unbiased estimates and also the accuracy of estimates can be computed. The probability sampling paradigm has been successfully applied in official and academic statistics since the 1940's, and to a much lesser extent also in more commercial market research. Web surveys often rely on self-selection of respondents instead of probability sampling. This has serious impact on the quality of survey results. The theory of probability sampling cannot be applied and estimates are often substantially biased.

National statistical institutes in many countries are faced with budget constraints on the one hand and demands for more and more detailed information on the other. The question arises whether web surveys can play a role in finding a solution of this dilemma. It is argued in this paper that self-selection surveys cannot play such a role. However, when web surveys are carried out within the framework of probability sampling, there are certainly possibilities, as a single-mode survey or as one of the modes in a mixed-mode survey.

¹Jelke Bethlehem, Statistics Netherlands, P.O. Box 24500, 2490 HA The Hague, The Netherlands (jbtm@cbs.nl)

2. Web surveys

2.1 Sampling the web

The basics of probability sampling as applied now in official statistics have been laid down by Horvitz and Thompson (1952) in their seminal paper. They state that unbiased estimators of population characteristics can always be constructed, provided samples are selected by means of probability sampling and every element in the population has a known and strictly positive probability of being selected. Moreover, under these conditions standard errors of estimates, and thus confidence intervals, can be computed. Therefore it is possible to establish the accuracy of estimates. The Horvitz-Thompson approach can be used also in surveys with complex sampling designs, like stratified random samples, cluster samples and two-stage samples.

Unfortunately, many web surveys are based on some form of self-selection. The survey is simply put on the web. Participation requires in the first place that respondents are aware of the existence of a survey. They have to accidentally visit the website, or they have to follow up a banner, e-mail message, or a call in another commercial. In the second place, they have to make the decision to fill in the questionnaire on the Internet. The survey researcher is not in control of the selection process.

All major opinion polls in The Netherlands use web panels that have been set up by means of self-selection. The values of some demographic variables are recorded during the recruitment phase. Therefore the distribution of these variables in a survey can be compared with their distribution in the population. Weighting adjustment techniques can be attempted to correct for over- or under-representation of specific groups. An example of large online cross-sectional survey in The Netherlands was *21minuten.nl*, a survey supposed to supply answers to questions about important problems in Dutch society. Within a period of six weeks in 2006 about 170,000 people completed the online questionnaires. A similar survey was conducted in Germany (*Perspektive Deutschland*). A study across 19 online panels of Dutch market research organizations showed that most of them use self-selection, see Vonk et al. (2006).

Self-selection web surveys results are sometimes claimed to be 'representative' because of the high number of respondents or as a results of advanced adjustment weighting procedures. The term representative is rather confusing, see Kruskal and Mosteller (1979a, 1979b, 1979c). It can have many meanings and is often used in a very loose sense to convey a vague idea of good quality. A high number of respondents are often considered to ensure validity and reliability. There are serious doubts, however, whether a large sample size as a result of self-selection of respondents has the same meaning as a large sample size in probability sampling.

Currently, many web surveys suffer from two fundamental methodological problems. The first one was already mentioned: self-selection. Researchers have no control over the selection mechanism. Selection probabilities are unknown. Therefore, no unbiased estimates can be computed, nor can the accuracy of estimates be established. The second problem is under-coverage. Since data is collected using the Internet, people without Internet access will never be able to participate in a web survey. This means research results can only apply to the Internet population and not to the complete population. These two problems are analysed in more detail in the subsequent sections.

2.2 Under-coverage

Web survey suffers from under-coverage because the target population is usually much wider than the Internet population. According to data from Eurostat, the statistical office of the European Union, 54% of the households in the EU had access to Internet in 2007. There were large variations between countries. The countries with the highest percentages of Internet access were The Netherlands (83%), Sweden (79%) and Denmark (78%). Internet access was lowest in Bulgaria (19%), Romania (22%) and Greece (25%).

Even more problematic is that Internet access is unevenly distributed over the population. A typical pattern found in many countries is that the elderly, the low-educated and ethnic minorities are severely under-represented among those having access to Internet. See Bethlehem (2007) for a description of the situation in The Netherlands.

To obtain insight in the impact of under-coverage on estimates, suppose a proper random sample is selected from the Internet population. Let the target population of the survey consist of N persons. Associated with each person k is a value Y_k of the target variable Y . Aim of the online survey is assumed to be estimation of the population mean $\bar{Y} = (Y_1 + Y_2 + \dots + Y_N) / N$ of the target variable Y .

The population U is divided into a sub-population U_I of size N_I of persons having access to Internet, and a sub-population U_{NI} of size N_{NI} of persons without Internet access. The sub-population U_I will be called the Internet population. Suppose a simple random sample is selected without replacement from the Internet-population. The sample mean \bar{y}_I is an unbiased estimator of the mean \bar{Y}_I of the Internet population, but not necessarily of the mean of the target population. Bethlehem (2007) shows that the bias of this estimator is equal to

$$B(\bar{y}_{HT}) = E(\bar{y}_{HT}) - \bar{Y} = \bar{Y}_I - \bar{Y} = \frac{N_{NI}}{N} (\bar{Y}_I - \bar{Y}_{NI}). \quad (2.2.1)$$

The magnitude of this bias is determined by two factors. The first factor is the relative size N_{NI} / N of the sub-population without Internet. Therefore the bias decreases as Internet coverage increases. The second factor is the contrast $\bar{Y}_I - \bar{Y}_{NI}$ between the means of the Internet-population and the non-Internet-population. The more the mean of the target variable differs for these two sub-populations, the larger the bias will be.

Since Internet coverage is steadily increasing, the factor N_{NI} / N is decreasing. This has a bias reducing effect. However, it is not clear whether the contrast also decreases. To the contrary, it is not unlikely that the (small) group of people without Internet will be more and more different from the rest of the population. As a result, substantial bias may still remain.

2.3 Self-selection

Participation in a self-selection web-survey requires that respondents are aware of the existence of the survey and that they decide to fill in the questionnaire on the Internet. All this means that each element k in the population has unknown probability ρ_k of participating in the survey, for $k = 1, 2, \dots, N$. Bethlehem (2007) shows that the expected value of the sample mean is equal to

$$E(\bar{y}) \approx \bar{Y}^* = \frac{1}{N} \sum_{k=1}^N \frac{\rho_k}{\bar{\rho}} Y_k \quad (2.3.1)$$

where $\bar{\rho}$ is the mean of all response propensities. The bias of this estimator is equal to

$$B(\bar{y}) = E(\bar{y}) - \bar{Y} \approx \bar{Y}^* - \bar{Y} = \frac{C_{\rho Y}}{\bar{\rho}} = \frac{R_{\rho Y} S_{\rho} S_Y}{\bar{\rho}}, \quad (2.3.2)$$

in which $C_{\rho Y}$ is the covariance between the target variable and the response probabilities, $R_{\rho Y}$ is the correlation coefficient, S_{ρ} is the standard deviation of the response probabilities, and S_Y is the standard deviation of the target variable. It can be shown that in the worst case (S_{ρ} assumes its maximum value and the correlation $R_{\rho Y}$ is equal to either +1 or -1) the absolute value of the bias is equal to

$$|B_{\max}(\bar{y})| = S_Y \sqrt{\frac{1}{\bar{\rho}} - 1}. \quad (2.3.3)$$

Bethlehem (1988) shows the formula (2.3.2) also applies in the situation in which a probability sample has been drawn, and subsequently nonresponse occurs during the fieldwork. Consequently, expression (2.3.3) provides a means to compare potential biases in various survey designs. For example, regular surveys of Statistics Netherlands are all based on probability sampling. Their response rates are around 70%. This means the absolute maximum bias

is equal to $0.65 \times S_y$. One of the largest self-selection web surveys in The Netherlands was *21minuten.nl*. Within a period of six weeks in 2006 about 170,000 people completed the online questionnaire. The target population of this survey was not defined, as everyone could participate. If it is assumed the target population consists of all Dutch from the age of 18, the average response propensity is equal to $170,000 / 12,800,000 = 0.0133$. Hence, the absolute maximum bias is equal to $8.61 \times S_y$. It can be concluded that the bias of the large web survey can be a factor 13 larger than the bias of the smaller probability survey.

The effects of self-selection can also be illustrated using an example related to the general elections in The Netherlands in 2006. Various survey organizations used opinion polls to predict the outcome of these elections. The results of these polls are summarized in Table 2.3-1. *Politieke Barometer*, *Peil.nl* and *De Stemming* are opinion polls carried out by market research agencies. They are all based on samples from web panels. To reduce a possible bias, adjustment weighting was been carried out. The polls were conducted one day before the election. The Mean Absolute Difference indicates how big the differences (on average) are between the poll and the election results. Particularly, differences are large for the more volatile parties like PvdA, SP and the PVV. For example, one poll predicted 32 seats in parliament for the SP (socialist party) whereas this party got only 25 seats.

DPES is the Dutch Parliamentary Election Study. The fieldwork was carried out by Statistics Netherlands in a few weeks just before the elections. The probability sampling principle has been followed here. A true (two-stage) probability sample was drawn from the population register. Respondents were interviewed face-to-face (using CAPI). The predictions of this survey are much better than those based on the online opinion polls. The predictions and election results only differ for four parties, and differences are at most one seat.

Table 2.3-1
Dutch Parliamentary elections 2006. Outcomes and the results of various opinion surveys

	Election result	Politieke Barometer	Peil.nl	De Stemming	DPES 2006
Sample size	...	1,000	2,500	2,000	2,600
Seats in parliament:					
CDA (christian democrats)	41	41	42	41	41
PvdA (social democrats)	33	37	38	31	32
VVD (liberals)	22	23	22	21	22
SP (socialists)	25	23	23	32	26
GL (green party)	7	7	8	5	7
D66 (liberal democrats)	3	3	2	1	3
ChristenUnie (christian)	6	6	6	8	6
SGP (christian)	2	2	2	1	2
PvdD (animal party)	2	2	1	2	2
PVV (populists)	9	4	5	6	8
Other parties	0	2	1	2	1
Mean Absolute Difference	...	1.27	1.45	2.00	0.36

... not applicable

Probability sampling has the additional advantage that it provides protection against certain groups in the population attempting to manipulate the outcomes of the survey. This may typically play a role in opinion polls. Self-selection does not have this safeguard. An example of this effect could be observed in the election of the 2005 Book of the Year Award (Dutch: NS Publieksprijs), a high-profile literary prize. The winning book was determined by means of a poll on a website. People could vote for one of the nominated books or mention another book of their own choice. More than 90,000 people participated in the survey. The winner turned out to be the new Bible translation launched by the Netherlands and Flanders Bible Societies. This book was not nominated, but nevertheless an overwhelming majority (72%) voted for it. This was due to a campaign launched by (among others) Bible societies, a Christian broadcaster and Christian newspaper. Although this was all completely within the rules of the contest, the group of voters could clearly not be considered to be representative of the Dutch population.

3. Weighting adjustment

3.1 Traditional weighting adjustment

Weighting adjustment is a family of techniques that attempt to improve the quality of survey estimates by making use of auxiliary information. *Auxiliary information* is defined as a set of variables that have been measured in the survey, and for which information on their population distribution (or complete sample distribution) is available. By comparing the population distribution of an auxiliary variable with its response distribution, it can be assessed whether or not the sample is representative for the population (with respect to this variable). If these distributions differ considerably, one must conclude that the sample is selective. To correct this, adjustment weights can be computed. Weights are assigned to all records of observed elements. Estimates of population characteristics can now be obtained by using the weighted values instead of the unweighted values. Weighting adjustment is often used to correct surveys that are affected by nonresponse, see e.g. Bethlehem (2002).

Post-stratification is a well-known and often used weighting method. To carry out post-stratification, one or more qualitative auxiliary variables are needed. Together they divide the target population into a number of strata (i.e. sub-populations). Identical adjustment weights are assigned to all elements in the same stratum. The bias of the estimate based on weighted data will be small if there is (on average) no difference between participants and non-participants. This is the case if there is a strong relationship between the target variable and the stratification variables. This situation is referred to in the literature as Missing at Random (MAR). The variation in the values of the target variable manifests itself between strata but not within strata. In other words, the strata are homogeneous with respect to the target variable. Unfortunately, such auxiliary variables are not very often available, or there is only a weak correlation.

3.2 Propensity weighting

Propensity weighting is used by several market research organisations to correct for a possible bias in their web surveys, see e.g. Börsch-Supan et al. (2004) and Duffy et al. (2005). *Propensity scores* are obtained by modelling a variable that indicates whether or not someone participates in the survey. Usually a logistic regression model is used where the indicator variable is the dependent variable and attitudinal variables are the explanatory variables. These attitudinal variables are assumed to explain why someone participates or not. Fitting the logistic regression model comes down to estimating the probability (propensity score) of participating, given the values of the explanatory variables.

The estimated propensity scores are used to stratify the population. Each stratum consists of elements with (approximately) the same propensity scores. If indeed all elements within a stratum have the same response propensity, there will be no bias if just the elements in the Internet population are used for estimation purposes. Cochran (1968) claims that five strata are usually sufficient to remove a large part of the bias. The market research agency Harris Interactive was among the first to apply propensity score weighting, see Terhanian et al. (2001).

To be able to apply propensity score weighting, two conditions have to be fulfilled. The first condition is that proper auxiliary variables must be available. These are variables that are capable of explaining whether or not someone is willing to participate in the web survey. Variables often used measure general attitudes and behaviour. The second condition is that the values of these variables must be available for participants and non-participants.

3.3 Weighting adjustment with a reference survey

If proper auxiliary variables are not available, it might be considered to conduct a *reference survey*. Such a reference survey is based on a small probability sample, where data collection takes place with a mode different from the web, e.g. CAPI (Computer Assisted Personal Interviewing, with laptops) or CATI (Computer Assisted Telephone Interviewing). Under the assumption of no nonresponse, or ignorable nonresponse, this reference survey will produce unbiased estimates of the population distribution of auxiliary variables. The reference survey approach has been applied by several market research organizations, see e.g. Börsch-Supan et al. (2004) en Duffy et al. (2005).

Using an estimated population distribution in post-stratification results in the same expected value of the estimator. So, the conditions under which the bias is reduced are the same as those for the normal post-stratification estimator. An interesting aspect of the reference survey approach is that any variable can be used for adjustment weighting as long as it is measured in both surveys. For example, some market research organizations use 'webographics' or 'psychographic' variables to divide the population in 'mentality groups'. People in the same groups are assumed to have more or less the same level of motivation and interest to participate in such surveys. If this is the case, such variables can be effectively used in weighting adjustment. This requires, of course, that adequate information on psychographics is available for the population, based on high response rate in random samples.

Schonlau et al. (2004) describe the reference survey of Harris Interactive. This is a CATI survey, using random digit dialling. This reference survey is used to adjust several web surveys. Schonlau et al. (2003) stress that the success of this approach depends on two assumptions: (1) the webographics variables are capable of explaining the difference between the web survey respondents and the other persons in the target population, and (2) the reference survey does not suffer from non-ignorable nonresponse. In practical situations it will not be easy to satisfy these conditions. Schonlau et al. (2007) show that use of webographics variables for adjustment weighting may work, but not always.

The reference survey approach also has a disadvantage. Bethlehem (2007) shows that if a reference survey is used to estimate the population distribution of the auxiliary variables, the variance of the post-stratification estimator is for a large part determined by the size of the small reference survey. So, the large number of observations in the online survey does not help to produce accurate estimates. The reference survey approach reduces the bias of estimates at the cost of a higher variance. The effective sample size of the web survey is of the same order of magnitude to that of the reference survey.

It should be remarked that attitudinal questions are much less reliable than factual questions. Respondents may never have thought of the topics addressed in attitudinal questions. They have to make up their mind at the very moment the question is asked. Their answers may depend on their current circumstances, and may vary over time. Therefore, attitudinal question may be subject to substantial measurement errors.

4. Web surveys for official statistics?

Can self-selection web surveys be used for data collection in official statistics? The discussion of the methodological problems in section 2 leads to the conclusion that there are severe methodological problems making it very hard, if not impossible, to make valid inference about the population to be surveyed. Self-selection can cause estimates of population characteristics to be biased. This seems to be similar to the effect of nonresponse in traditional probability sampling based surveys. However, it was shown that the bias in self-selection surveys can be substantially larger.

Adjustment weighting may help to reduce the bias, but only if the sample selection mechanism satisfies the Missing at Random (MAR) condition. This implies that weighting variables must have a strong relationship with the target variables of the survey and the response probabilities. Often such variables are not available. A way out of this problem may be to carry out a reference survey. There are some reports that webographics variables seem to be capable of explaining response behaviour. It should be noted that webographics variables are attitudinal variables. They are much harder to measure than factual variables, and therefore may be subject to large measurement errors.

A reference survey only works well if it is a real probability sample without nonresponse, or with ignorable nonresponse (MCAR). This condition may be hard to satisfy in practical situations. If reference survey estimates are biased due to non-ignorable nonresponse, the web survey bias is replaced by a reference survey bias. This does not really help to solve the problem.

Self-selection is a serious problem, but it can be solved by applying probability sampling. A random sample (e.g. of addresses) can be drawn from a sampling frame. A letter can be sent to each selected address with request to complete a questionnaire on the Internet. Unique identification codes guarantee that the proper persons answer the

questions. In fact, the only difference with a mail questionnaire is that the paper questionnaire form is replaced by an electronic one on the Internet.

The problem of under-coverage in web surveys has to be addressed too. It is interesting to note that only between 60% and 70% of the households in The Netherlands still have a listed landline telephone. So one out of three households is missing if a sample is selected from a telephone directory. This shows that also more traditional phone surveys suffer from under-coverage. It is expected that under-coverage will decrease over time. From the point of view of coverage a web survey may be better than a telephone survey, at least in The Netherlands.

If under-coverage in web survey really is a problem, a possible solution could be to simply provide Internet access to those without Internet. An example of this approach is the LISS panel, see Scherpenzeel (2008). This Dutch panel was established in 2007. It consists of approximately 5,000 households. To construct this panel, Statistics Netherlands selected a probability sample from the Dutch population. Households were invited to become a member of the panel by means of a CAPI or CATI survey. Households without Internet access were offered a free Internet connection and a simple Internet-PC, developed for those who never had worked with Internet before.

Can a web survey be an alternative for a CAPI or CATI survey? With respect to data collection, there is a substantial difference between CAPI and CATI on the one hand and web surveys on the other. Interviewers carry out the fieldwork in CAPI and CATI surveys. They are important in convincing people to participate in the survey, and they also can assist in completing the questionnaire. There are no interviewers in a web survey. It is a self-administered survey. Therefore quality of collected data may be lower due to higher nonresponse rates and more errors in answering questions. According to De Leeuw and Collins (1997) response rates tend to be higher if interviewers are involved. However, response to sensitive questions is higher without interviewers.

A web survey can be one of the modes in a mixed-mode data collection approach. Each mode of data collection (face-to-face, telephone, mail, web, etc) has its advantages and disadvantages. Mixing data collection modes provides an opportunity to compensate for the weakness of each individual mode. This can reduce costs and at the same time increase response rates and data quality. Several mixed-mode data collection strategies are possible. One is the concurrent approach. The sample is divided in groups that are approached by different modes, at the same time. A second mixed-mode strategy is the sequential approach. All sample elements are first approached by the cheapest mode (Internet). The non-respondents are then followed up by the next cheapest mode (CATI), and finally remaining non-respondents by the most expensive mode (CAPI). A third strategy could be to let respondents select their preferred data collection mode. This is flexible and respondent-friendly, but some research seems to indicate that response rates drop as soon as respondents are offered a choice of modes, see Dillman (2008).

A major concern for mixed-mode data collection is that data quality may be affected by the occurrence of mode effects. This is the phenomenon that asking a person the same question in different data collection modes would lead to a different answer. An example is asking a closed question with a substantial number of answer options. The respondent in a face-to-face survey would be presented a show card with all possible answers. In case of a telephone survey, the interviewer would read all possibilities to the respondents. Research indicates that this results in a preference for the last options in the list. Respondents offered a drop-down list in a web survey have a preference for answers early in the list.

The conclusion can be that there may be opportunities to use web surveys as a mode of data collection in official statistics. More research is required to establish how this can be realized without affecting the quality of the collected data.

References

Bethlehem, J.G. (1988). Reduction of the nonresponse bias through regression estimation, *Journal of Official Statistics*, 4, 251-260.

- Bethlehem, J.G. (2002). Weighting Nonresponse Adjustments Based on Auxiliary Information. *Survey Nonresponse* (Eds. R.M. Groves et al.). New York: Wiley.
- Bethlehem, J.G. (2007). Reducing the bias of web survey based estimates, Discussion Paper 07001, Voorburg, The Netherlands: Statistics Netherlands.
- Cochran, W.G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies, *Biometrics*, 24, 205-213.
- Couper, M.P., Baker, R.P., Bethlehem, J.G., Clark, C.Z.F., Martin, J., Nicholls II, W.L. and O'Reilly, J.M. (Eds.) (1998). *Computer Assisted Survey Information Collection*, New York: Wiley.
- De Leeuw, E. and Collins M. (1997). Data collection methods and survey quality. *Survey Measurement and Process Control* (Eds. L. Lyberg et al.). New York: Wiley, pp. 199-220.
- Dillman, D. (2008). How can we convince the general public to respond to Internet surveys, *MESS Workshop*. Zeist, The Netherlands.
- Duffy, B., Terhanian, G., Bremer, J. and Smith, K. (2005). Comparing data from online and face-to-face surveys, *International Journal of Market Research*, 47, 615-639.
- Horvitz, D.G. and Thompson D.J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 47, 663-685.
- Kruskal, W. and Mosteller, F. (1979a). Representative Sampling, I: Non-scientific Literature, *International Statistical Review*, 47, 13-24.
- Kruskal, W. and Mosteller, F. (1979b). Representative Sampling, II: Scientific Literature, Excluding Statistics, *International Statistical Review*, 47, 111-127.
- Kruskal, W. and Mosteller, F. (1979c). Representative Sampling, III: the Current Statistical Literature, *International Statistical Review*, 47, 245-265.
- Scherpenzeel, A. (2008). An online panel as a platform for multi-disciplinary research, *Access panels and online research, panacea or pitfall?* (Eds. I. Stoop and M. Wittenberg), Amsterdam: Aksant, 101-106.
- Schonlau, M., Fricker, R.D. and Elliott, M.N. (2003). Conducting Research Surveys via E-mail and the Web, Rand Corporation, Santa Monica, CA.
- Schonlau, M., Van Soest, A. and Kapteyn, A. (2007). Are 'webographic' or attitudinal questions useful for adjusting estimates from web surveys using propensity scoring?, *Survey Research Methods*, 1, 155-163.
- Schonlau, M., Zapert, K., Payne Simon, L., Sanstad, K., Marcus, S., Adams, J., Spranca, M., Kan, H., Turner, R. and Berry, S. (2004). A Comparison between responses from propensity-weighted web survey and an identical RDD survey, *Social Science Computer Review*, 22, 128-138.
- Terhanian, G., R. Smith, J. Bremer, and R. K. Thomas (2001). Exploiting Analytical Advances: Minimizing the Biases Associated with Internet-Based Surveys of Non-Random Samples, *ARF/ESOMAR: Worldwide Online Measurement*, ESOMAR Publication Services, 248, 247-272.
- Vonk, T., Van Ossenbruggen, R. and Willems, P. (2006). The effects of panel recruitment and management on research results, a study among 19 online panels, *Panel Research 2006, ESOMAR World Research*, ESOMAR Publication Services, 317, 79-99.