Proceedings of Statistics Canada Symposium 2014
Beyond traditional survey taking: adapting to a changing world

Web panel surveys – a challenge for official statistics

Jörgen Svensson[1]

Abstract

During the last decade, web panel surveys have been established as a fast and cost-efficient method in market surveys. The rationale for this is new developments in information technology, in particular the continued rapid growth of internet and computer use among the public. Also growing nonresponse rates and prices forced down in the survey industry lie behind this change. However, there are some serious inherent risks connected with web panel surveys, not least selection bias due to the self-selection of respondents. There are also risks of coverage and measurement errors. The absence of an inferential framework and of data quality indicators is an obstacle against using the web panel approach for high-quality statistics about general populations. Still, there seems to be increasing challenges for some national statistical institutes by a new form of competition for ad hoc statistics and even official statistics from web panel surveys.

This paper explores the question of design and use of web panels in a scientifically sound way. An outline is given of a standard from the Swedish Survey Society for performance metrics to assess some quality aspects of results from web panel surveys. Decomposition of bias and mitigation of bias risks are discussed in some detail. Some ideas are presented for combining web panel surveys and traditional surveys to achieve controlled cost-efficient inference.

Key Words: selection bias, quality indicators, inference, access panels.

1. Introduction

The fast development of the internet in recent years has entailed a new type of survey: the web panel survey. A **web panel** (or online or internet panel) could be defined as an access panel of people willing to respond to web questionnaires. By access panel we mean a sample database of potential respondents who declare that they will co-operate for future data collection if selected (ISO 2009). A web panel survey is a survey utilizing samples from web panels, i.e. a survey via the web of persons who are willing to participate in such surveys. (This is something different than fixed or rotating panels in longitudinal surveys based on probability sampling.)

The concepts of web panels and web panel surveys should not be confused here with web surveys in the meaning surveys using **web questionnaires**, i.e. utilizing the web mode. The latter concept only means that data are collected via the internet, regardless of how the sampling is made. Statistics Sweden carries out much of its data collection via web questionnaires, particularly for business surveys and public sector surveys. Although crucial, the measurement issues of web questionnaires will not be dealt with in this paper, which will focus on web surveys using access panels and their inherent selection problems.

There is still no consensus about how to carry out web panel surveys so as to provide correct inference and adequate data quality. An important issue is the **self-selection** approach, seemingly more or less giving up the usual control of sampling errors.

Web panel surveys have so far not been frequently used by national statistical institutes (NSIs), and Statistics Sweden has up to this point not worked with web panels. However, the market research industry appears to increasingly abandon traditional postal enquiries and telephone interviews in favor of self-selection web panels. One reason for this may be the growing nonresponse problem connected to rigorously designed surveys: if nonresponse increasingly threatens the validity of the probability sampling approach, it may seem no worse to abandon this

[1]Jörgen Svensson, Statistics Sweden, SE-701 89 Örebro, SWEDEN.

approach and move to access panels. Moreover, surveys through web panels are often much cheaper than traditional surveys. NSIs that conduct surveys on commission will probably be **increasingly challenged** by a new form of competition for ad hoc surveys and by invitations on cooperation on data collection through web panels.

A **description** of how web panel surveys work more in detail can be found in Svensson (2013). There, recruitment of web panels, sampling methods, estimation methods and panel management are described.

2. Assessing web panel surveys

In this section, some of the most important benefits and drawbacks of using web panel surveys instead of traditional surveys utilizing probability sampling are pointed out.

2.1 Advantages of web panel surveys

An advantage for web panel surveys is that they are considered to be **uncomplicated**. The approach gives easy access to large groups of potential respondents, once the web panels have been set up.

A second advantage is that web panel surveys are **inexpensive**, as they need no interviewers, no printing and mailing, and notably less demanding tracing and persuasion of nonrespondents. However, to manage high-quality web panels could be rather costly. Costs are associated with recruitment, panel administration and support, incentive programs, web portals for the panelists, and sampling and administration of specific surveys. Self-recruited panels are normally much cheaper than panels that are probability-based from the start.

A third advantage is that web panel surveys are **fast**. A survey can be launched and finished very quickly.

A fourth advantage of web panel surveys is that the **respondent burden is decreased** so far that the respondents *volunteer* to participate in surveys.

2.2 Disadvantages of web panel surveys

The most important disadvantage of web panel surveys is probably the **self-selection problem**. The sample selected from the web panel survey is not a probability sample, even if the recruitment was done through probability sampling or if the sample from the web panel frame is a probability sample. This is why a task force of AAPOR (American Association for Public Opinion Research) recommends that nonprobability online panels should be avoided when estimating population values (AAPOR 2010). There is also a risk of manipulation, with examples of campaigns resulting in biased statistics. The existence of professional survey-takers is also a problem. There is nothing to prevent an individual from joining many different web panels. Web panel surveys will then reflect the loud minority instead of the silent majority.

The *bias* resulting from self-selection is proportional to the correlation coefficient between the target variable and the participation probability, to the standard deviation for the participation probability, to the standard deviation for the target variable, and to the reciprocal value of the average participation probability. See Bethlehem (2009). Thus, the bias will be diminished if the relationship between participation behavior and the target variable is reduced, if the variation in the participation probabilities is reduced, or if the average participation probability is increased. The same applies to the situation in which a probability sample has been drawn and subsequently nonresponse occurs during data collection. By comparison however, it can be seen that the risk of bias is much higher in web panel surveys, since the participation probabilities (or response propensities) in most cases are much lower. An argument for not embarking on web panel surveys for official statistics could be that one mostly is interested in means and totals, for which there is an obvious risk of biased estimators, while market research institutes often focus on parameters that include purchasing potential, which might be related to the participation probability, and then more reliable results are probable from web panel surveys.

A second disadvantage of web panel surveys is the **noncoverage** (undercoverage) **problem**. The target population of a survey is often wider than those having access to the internet. People without internet will never be selected for

a web panel survey and may differ from those with internet. Noncoverage bias results from the fact that online panels cannot represent people who are not online. The bias risk may be most severe for surveys on elderly, low-educated and ethnic minority groups, since they have lower internet coverage. However, there is also a risk of bias for general population surveys. The bias is, according to Bethlehem (2009), proportional to the relative size of the group of people without internet and to the average difference between people with and without internet. A consequence is that the bias would diminish if internet coverage increases, with everything else alike. But of course everything else may not necessarily be alike over time, as the contrast between the means of the internet population and the non-internet population may change.

A third disadvantage, connected mostly to the first one described above, is the **difficulty of assessing the quality** of the results. No ordinary estimates of sampling error or confidence intervals can be calculated for the usual type of web panel survey. This follows from the nonprobability character of the response set. However, a credibility interval, reflecting the statistical uncertainty generated by a statistical model that relies on Bayesian statistical theory, could be calculated. A drawback then is that the statistical insights from nonprobability samples will depend on explicit models linking the sample to the target population.

3. A Swedish standard with performance metrics for web panel surveys

The Swedish Survey Society has had a **web panel committee**, with the purpose of proposing measures for assessing quality aspects of results from web panel surveys. A main goal was to increase the **transparency** in procedures and methods in the industry. The committee proposed several quality indicators or rather performance metrics. Below some measures are presented briefly, see Nyfjäll (2013) for more details. It is too early to see if the industry will use the proposed metrics.

First, some **measures connected to traditional surveys and nonresponse** are proposed: cumulative participation rate = recruitment rate × profile rate × participation rate. The recruitment rate can only be calculated for a probability-based sample. The profile rate is taken from the profile survey. The participation rate (or completion rate) concerns the specific survey with a sample from the panel. The cumulative (or more correctly *final*) participation rate takes all steps in a web panel survey into account, from the sample in the recruitment to the response set in a specific survey. This measure is similar to the response rate in a traditional survey and can be used for coarse comparisons. The cumulative participation rates are often much lower in Swedish web panel surveys than in Swedish traditional surveys.

However, the selection bias resulting from noncoverage and nonresponse can be decomposed according to the following (not taken from the web panel committee or Nyfjäll, 2013). Here $\hat{\bar{Y}}$ is the straight expansion estimator of a mean value, e.g. an average income. It is based only on respondents in a specific survey based on a sample from a web panel.

- The first term, *noncoverage rate* × $(\bar{Y}_{cov} - \bar{Y}_{noncov})$, shows the rate of those not having access to internet multiplied by the contrast for that group compared with those having internet, resulting in the noncoverage bias.
- The second term, *nonrecruitment rate* × $(\bar{Y}_{recr} - \bar{Y}_{nonrecr})$, shows the nonresponse bias from those with internet access but not recruited from the probability sample.
- The third term, *nonprofile rate* × $(\bar{Y}_{prof} - \bar{Y}_{nonprof})$, shows the nonresponse bias from those recruited but not profiled in the profile survey.
- The fourth term, *nonparticipation rate* × $(\bar{Y}_{part} - \bar{Y}_{nonpart})$, shows the nonresponse bias from those profiled but not participating in the specific survey.

Thus, the measures on recruitment, profile and participation rates described above are **not enough to estimate the bias**. There is also a need to know how much bias is associated with each drop-out stage. Some kinds of estimates of the contrast between those recruited and those not, etc., are needed. This is however very difficult and costly to achieve.

There are also metric **measures related to the panel**. One measure is the response burden, described by the total number of invitations sent out in relation to the panel size or by the total number of responses in relation to the panel size. Dominance from certain panelists can be described by the proportion of responses that the 20 percent most active (most responding) panelists stand for during a year. Other measures are panel size and attrition rate.

One example of a **measure related to a specific survey** deals with panel conditioning: the proportion of partial nonresponse or most positive scale answers in a group of respondents who have participated in many surveys; high rates may indicate inattentiveness or professionalism. The break-off rate is the proportion of opened but not completed web questionnaires; a high rate may indicate technical problems or a too long or boring questionnaire.

Then, there is also a need for **verbal descriptions** of the recruitment method(s) for the web panel, the sampling design for a specific survey, and the weighting methods for estimation from a specific survey.

4. Some ideas for combinations of web panel surveys and traditional surveys

An interesting question is if it would be possible to enjoy the advantages in convenience of web panel surveys, but essentially without having to give up **controlled inference**. Combinations of web panel surveys and traditional probability sampling surveys could then be a solution. This will move us away from the traditional concept of a sample survey as a free-standing entity for data collection and analysis. Instead, emphasis will be put on how to integrate data from different sources. The importance of sampling errors will probably be diminished, while the combination of different sampling and nonsampling errors will have to be focused.

One solution for the noncoverage problem is to conduct an **extra survey** or have an extra stratum, i.e. for the group of **people without internet**. This group could be approached by a mail or CATI survey. A drawback is the mixed-mode set-up. Another possibility, however costly, is to provide this group with hardware, software and training and conduct the extra survey by the internet.

An approach for handling the selection problems is to **calibrate a web panel survey to a traditional probability survey.** The two surveys should be run in parallel and be as similar as possible. If the two surveys give almost equal results, there is an empirical base for conducting the web panel survey for some rounds and expecting good enough results. Otherwise, coefficients for a calibration model are estimated using the two surveys. The coming survey round, a model-based estimation using the web panel results is applied in order to achieve a traditional probability survey result. A re-calibration needs to be done periodically, depending on the nature of the subject matter.

Another idea is to conduct **follow-up studies** in a systematic way. Firstly, a web panel survey is conducted, and results are presented. Then, a probability sample, much smaller than the web panel sample, is drawn from a comprehensive frame, like the Population Register of Statistics Sweden. The sampled individuals are surveyed by a telephone interview with the same questions. The results from this follow-up (traditional) study are then reported, maybe one or two months later. More important, an interesting, empirical material to describe the quality of different web panel surveys will fairly soon be available to both the producer and the users. When combined over several surveys, the sample sizes do not have to be very large for each follow-up study. Still the nonresponse in these studies might cause problems. Another drawback is the mode effects. This idea was mentioned at a meeting in 2012 with Statistics Sweden and its Scientific Advisory Board. However, the idea has not been elaborated on or tested in practice to show whether it is feasible for Statistics Sweden.

There are also examples of weighting under **propensity score adjustment** in the market research industry. This method, see e.g. Lee (2006), requires good access to auxiliary information for the persons in the web panel and for the persons in a so-called reference sample that has been selected with probability sampling and with ignorable nonresponse. For the reference sample, variables are reported that are relatively simple and inexpensive to measure. However, background variables like gender and age are often not enough; psychographic data (attitude and lifestyle data) are also required. The propensity score adjustment method falls within the framework of calibration estimation and leads to unbiased estimators if all assumptions on relations are true, but this is probably too often not the case. A disadvantage is that the method leads to substantial increase in the variance of the estimators.

Another alternative is to use a **poststratification approach**. Firstly, a sample is drawn from a web panel, and a web panel survey is conducted. Then, a probability sample, smaller than the web panel sample, is drawn from a comprehensive frame. The sampled individuals are surveyed by a telephone interview or by a postal (or web) questionnaire with the same questions as for the web panel sample. In the end of the interview or questionnaire, a question is asked whether the respondent would like to be a member in a web panel. The persons responding positively will then be considered to correspond to the same poststratum as the web panel survey sample has been drawn from. So the poststratification divides the individuals into two categories: panel positives and panel negatives. Then, the ordinary poststratified estimator would be utilized to produce results, which will be (technically) unbiased. One feature is that a traditional measure of quality in surveys, the sampling variance, is introduced. This will raise the credibility of the combined survey, although the precision might be too low if the probability sample is too small. Mode effects occur when adding responses from the web mode and the interview (or postal) mode and can result in measurement errors. Another issue might be the quality of the division into poststrata. This approach has not been investigated in practice.

5. Some challenges to Statistics Sweden

During the last five years, some **challenges** from web panel surveys have been met by Statistics Sweden. Other public agencies in Sweden, even statistical authorities, have started to purchase web panel data. Some agencies have started to send call for tenders, also for framework agreements with statistical producers, where the capacity to conduct telephone surveys, postal enquires *and* web panel surveys is required.

The Swedish Agency for Economic and Regional Growth is responsible for data about tourism in Sweden. The agency purchase statistics from Statistics Sweden and other producers. Last year, there was a call for tender for **consumer tourism data**. These survey data are EU regulated and concern Swedes travelling in Sweden: their number of overnight stays in different counties, their expenses and the like. The data are used a lot for analyses, although not so far being official statistics in Sweden.

The former contractor had conducted a traditional telephone survey based on probability sampling for some 25 years and had a large tourism database for analyses. The agency chose a new contractor with a **web panel survey** approach. One of the motivations was that the new contractor had a better description of the calculations of confidence intervals. The description included a kind of implicit simple random sampling assumption. Thus, the formula was correct in a formal way, but irrelevant. The former contractor appealed against this in court, but lost the case. No formal mistake was done in the purchase, according to the court.

Statistics Sweden is an intermediary and sends these statistics together with other tourism statistics to the EU statistical office, Eurostat. Furthermore, Statistics Sweden utilizes the data for tourism satellite statistics (a satellite account to the national accounts). This product has been on commissioned services from the Swedish Agency for Economic and Regional Growth to Statistics Sweden for several years. It is not decided yet whether the new web panel data will be utilized in the work by Statistics Sweden during 2015.

6. An approach from Statistics Denmark

The Interview Department at Statistics Denmark has embarked on **testing web panel surveys** for producing population estimates. They started collecting email addresses from samples in traditional, probability-based surveys. Up to the summer of 2014, they had some 10,000 addresses or individuals in their panel. The recruitment rate was about 32 percent. They hope that the final cumulative participation rate will end up at above 20 percent. Two surveys for non-official statistics – OECD Better Life Index and Gross National Happiness Index – were launched in November 2014. The results are not known yet.

Statistics Denmark uses the term **pseudo-representative design** for these web panel surveys. The surveys are said to be *representative* since a lot of register information, like gender, age, region, income, education, ethnic background, socioeconomic category and type of family, can be used in sampling design and in estimation aiming at reaching sufficient quality. The term *pseudo* is used since there are still issues with the noncoverage and non-

response bias. The rationale for the Danish approach is that NSIs should be present in the expanding area of the survey industry and study and learn from experiences. It might be possible to find a new customer segment for rapid and cheap surveys, while also keeping the segment of traditional probability-based surveys.

7. Possible next steps for national statistical institutes

There are several **possible alternatives** for Statistics Sweden and other NSIs in the coming years:

- A baseline alternative is of course **not to embark** at all into the approach of web panel surveys. When customers ask for web panel surveys, they will then be told that we do not use them since they are not scientifically sound or that no guarantee of quality can be given. At the same time Statistics Sweden and other institutes will increasingly work with the web mode, also for surveys on individuals.
- A second alternative is to start using web panels for **qualitative surveys**, generating ideas or exploring the existence of a phenomenon. The web panels may also be used for experiments, randomizing parts of the panel for studies of different questionnaires.
- A third alternative is to start testing the web panel approach for **population estimates** in a small scale to start with. This would be similar to the Danish alternative, see Section 6. Preferably, Statistics Sweden should build its own web panels from probability-based samples. The estimators used should utilize the wealth of auxiliary register information at Statistics Sweden. Quality studies need to be conducted in order to gain more information on benefits and drawbacks with the new approach. Studies would be required on if and how web panel surveys can be adapted to meet the quality requirements of official statistics.
- A fourth, more radical, alternative, perhaps relevant in a later phase, is to **implement web panel surveys for mostly non-official statistics**. Here, the nonresponse problems are often profound. The theory of probability sampling and design-based estimation may often seem of questionable correct applicability due to nonresponse. A risk of this approach is that Statistics Sweden encourages the use of nonprobability surveys.
- A fifth alternative might very hypothetically be interesting in the long run. This alternative includes **extensive use of web panel surveys** and model-based estimation and would imply a new paradigm for a NSI. It would take considerable effort over a long period to build an alternative system like this that is scientifically acceptable. In the choices for specific statistics, important considerations have to be made on subject-matter aspects and quality needs.

So far, web panel surveys are not commonly used for official or other statistics produced by NSIs. Probably, institutes will be more challenged by competition from web panels for at least ad hoc surveys in the coming years. Invitations on cooperation on data collection through web panels may also be expected to increase. **NSIs should be careful with their reputation** and not encourage the use of bad web panel surveys.

References

AAPOR (2010). AAPOR Report on Online Panels. Deerfield, IL: American Association for Public Opinion Research.

Bethlehem, J.G. (2009). *Applied Survey Methods, A Statistical Perspective.* John Wiley & Sons, Hoboken, NJ.

ISO (2009). ISO 26362:2009, Access panels in market, opinion and social research – Vocabulary and service requirements.

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

Nyfjäll, M. (2013). A standard with quality indicators for web panel surveys: a Swedish example. *Paper presented at the 59th ISI World Statistical Congress, 2013, Hong Kong, China.*

Svensson, J. (2013). Web panel surveys – can they be designed and used in a scientifically sound way? *Paper presented at the 59th ISI World Statistical Congress, 2013, Hong Kong, China.*