Web panel surveys – can they be designed and used in a scientifically sound way?

Jörgen Svensson
Statistics Sweden, Örebro, SWEDEN
e-mail: jorgen.svensson@scb.se

Abstract

Web access panels have received a widespread use in the survey industry for cost-effective surveys. It is not clear, however, if web panel surveys are plausible for high quality products. Can these surveys be adapted in their design to meet the quality requirements of an agency producing official statistics? The self-selection problems of web panel surveys, arising from the recruitment method and the nonresponse, obviously need to be counteracted. This paper presents the ordinary methodology used for web panel surveys as well as a few ideas for combinations of web panel surveys and traditional surveys in order to achieve a more controlled inference situation.

Keywords: inference, official statistics, poststratification, selection bias

1. The web panel challenge

The rapid development of the internet in recent years has entailed a new type of survey: the web panel survey. A **web panel** – or online/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 cooperate for future data collection if selected (ISO 2009). A web panel survey is a survey utilizing samples from web panels. (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**. 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.

Web panel surveys do not seem to be frequently used by national statistical institutes. 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 approach and move to access panels. Moreover, surveys through web panels are often much cheaper than traditional surveys. National statistical institutes 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.

Apparently, a question for national statistical institutes and other producers of statistics for governmental use is: Can web panel surveys be designed and used in a **scientifically sound way?**

2. How web panel surveys work

Below, the methodology (most often) used for web panel surveys is described briefly. Of course, the methods applied vary a lot between different countries and different companies.

2.1 Recruitment of web panels

Web panels are needed as 'sampling frames' for web panel surveys. All persons in the panels must have (up to date) e-mail addresses. **Recruitment** for web panels can be made in different ways. Respondents can be sourced from offline channels: telephone, TV ads, radio ads, ads in newspapers and magazines, addressed letters, outdoor posters, customer registers, etc. Respondents can also be sourced from online channels: e-mails, websites, banners, community sites, member programs, etc. Often, many channels are used in order to achieve the necessary diversity. After the recruitment, a profile survey is conducted in order to collect information on the new panelists.

The recruitment procedure above can be **classified** into the following alternatives:

- Panels recruited using probability sampling.
    - Sampling aiming only on panel recruitment.
    - Sampling with other aims, but where panel recruitment is included.
- Self-recruited panels.

One **example** from sampling with other aims is the approach to ask a question about participation in the web panel during a telephone survey. The response rate is often less than 50 percent in a telephone survey conducted by a market research company. It is said to be common that 20–30 percent of the respondents accept an offer. Thus, so far there is a selection of about one tenth of the individuals. The next step is to conduct a profile survey, getting data on attitudes and behavior and on possession of products. Let us say that 70 percent of the respondents remain, then labeled as active panel members. The final step is to draw a sample from the panel. Often, about 35–50 percent of the sampled individuals respond to the questionnaire. Then, the respondents correspond to less than 5 percent of the individuals. The market research industry avoids talking of nonresponse in this situation, but actually it would still in some way be fair to say that the nonresponse rate is higher than 95 percent.

In practice, at least from our Swedish experience, the distinction between probability-based and self-recruited panels will not be very important, as the probability-based panels will end up in **nonprobability** web panel surveys anyway, like in the example above. This is probably the case for most surveys, with the exception of some very ambitious surveys based on probability samples. The quality among nonprobability surveys may of course vary widely.

2.2 Sampling methods

Simple random sampling, stratified sampling or quota sampling is usually offered from web panels. **Stratification** is an effective method which often increases the precision of the estimators. Here however, stratification is utilized outside the concept of probability sampling, and it is unclear how much is gained with the technique.

Quota sampling is in itself a dubious method, even if the market research industry often advocates it. One begins by identifying important subgroups of the population, and estimates their proportions (quotas) in the population. The estimates can be based on other studies or register data. Then in the data collection phase, respondents are recruited so that the proportions in various subgroups agree reasonably with the estimated proportions in the population. For instance, specified numbers of individuals are recruited in different age groups. Quota sampling is to some extent suggestive of stratified sampling, but the similarity is illusory, since quota sampling lacks the

scientific basis of a stratified probability sample.

The first panelists responding, within a sample, will be used as the response set for the specific web panel survey. The **response rate** used may vary, and we have no figure on a standard rate. In order to avoid nonresponse bias at this stage, the invitation e-mails are normally neutral and do not reveal the topic of the survey.

2.3 Estimation methods

Estimation methodology can vary for web panel surveys. If the parameters sought are proportions, unweighted estimation may work under certain conditions. However, reliability of the estimators will most often benefit from using **weighting techniques**. The main effect here is the correction of recruitment bias and nonresponse bias. Some examples of weighting methods are poststratification, generalized regression estimation, calibration estimation and raking ratio estimation. For weighting, auxiliary variables are required. They must be measured in the web panel survey, and moreover information on their population distribution (or on their web panel distribution according to the profile survey, for the specific survey part) needs to be available. There needs to be a correlation between auxiliary variables and target variables or response behavior. Such auxiliary variables are not always available. However, it should be noticed that estimations of changes are less prone to bias than estimations of levels.

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 objects in the web panel and for the objects 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.

2.4 Panel management

When using panels, there are many quality issues to address. Firstly, is it an **actively managed panel** – where the panelists have taken an active and conscious decision to participate in web surveys on a regular basis – or just a database? Furthermore, professional survey-takers have to be under control. Panelists need to be taken away: the panelists resign or the panel owner removes the panelists, for different reasons. The size of the panel is an important factor, especially when trying to reach a given target group. Is breadth of coverage ensured in relevant target groups? The recruitment channels should give necessary diversity. Panelists should be admitted to the panel through a double opt-in validation procedure, which means that panelists must confirm their joining in the panel by e-mail. The respondent integrity is ensured by using the panel solely for surveys and not for sales or marketing purposes.

3. Assessing web panel surveys

In this section, an attempt will be made to point out some of the most important benefits and drawbacks of using web panel surveys instead of traditional surveys utilizing probability sampling. This will complement what has been said above.

3.1 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 probability-based panels.

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

3.2 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 **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. Undercoverage 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 estimates of sampling error or confidence intervals can be calculated for ordinary web panel surveys. This follows from the nonprobability character of the response set.

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.

One idea here is to use a **poststratification approach**. Firstly, a sample is drawn from a web panel, and a survey is conducted as described in Section 2 above. Then, a probability sample, 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 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. This idea was mentioned at a meeting in 2012 with Statistics Sweden and its Scientific Advisory Board. However, we have not elaborated on or tested the idea in practice to show whether it is feasible for Statistics Sweden. 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.

A second 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. 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. This idea was also mentioned by the Scientific Advisory Board, but any further investigations have not been performed.

5. Concluding remarks

When using **probability samples** of sufficient size, there is no need to worry about the ‘representativeness’ of the sample (however, see the nonresponse problem below). The randomization in the sampling procedure ensures ‘representativeness’ with respect to all relevant variables of the objects. Bias control is then not primarily based on background variables or comparisons with other data sources, but on the probability sampling itself. The procedure is transparent and built on a scientific basis. For a national statistical institute like Statistics Sweden, it is generally considered essential to use probability sampling in sample surveys as much as possible.

However, we are facing a growing **nonresponse problem**, at least for surveys on individuals and households, and **have to counteract** that problem. *One approach* for a national statistical institute would be to say: It is true that the nonresponse (if it is not

completely at random) leads to a set of respondents that is not a probability sample from the population, but it is not considered reasonable to build in a risk of bias as early as in the sampling methodology. Thus, the risk of systematic nonsampling errors is not a strong argument for giving up the whole idea of probability sampling. Rather, it is essential to build on a solid foundation of probability sampling and minimize nonresponse etc. and adjust as effectively as possible for the remaining nonresponse. *Another approach* would be to say: The theory of probability sampling and design-based estimation is invalidated if there is (a high rate of) nonresponse. Therefore, one alternative might be to gather data from a set of objects that may not satisfy the rigorous criteria of a probability sample, but instead some other criteria for controlled selection. National statistical institutes like Statistics Sweden have an advantage here of having a wealth of auxiliary information (mainly from administrative registers) on which to build. However, it takes considerable effort, over a long period, to build alternative systems that are scientifically acceptable. A *combination approach*, like the poststratification method described in Section 4, where web panel surveys are complemented by traditional studies, could also be a viable option, but needs to be investigated thoroughly.

Quality requirements vary for different types of surveys, from surveys for official statistics, to ad hoc surveys for governmental investigations, to research surveys, to opinion polls and market research. For the official Labor Force Survey, e.g., quality is of utmost importance to the users at ministries etc., as far-reaching political decisions are based on the results. It may be considered that quality requirements are less critical for market research, even if the industry may partly disagree. A question is whether web panel surveys may be considered suitable only for surveys with relatively less critical lower quality requirements. In that case, the survey objectives and the quality requirements have to be communicated in a clearer way than they often are today. It is a far from evident judgment to make, when web panel surveys are motivated and when traditional probability sampling surveys must be conducted. Therefore, methodological research should continue.

So far, web panel surveys are not commonly used for official or other statistics produced by **national statistical institutes**. Probably, institutes will be much more challenged by competition from web surveys 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. National statistical institutes therefore need to examine the use of web panels and maybe conduct some quality studies in order to learn more about the methods and if and how they can be adapted to meet the quality requirements of official statistics. Combinations of web panel surveys and traditional surveys in order to achieve a more controlled inference situation should also be studied in detail. National statistical institutes should be careful with their reputation and not legitimate the use of bad web panel surveys. To compete in the segment ‘low price, low quality’ is not advisable.

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.