DHĀRANA
BHAVAN'S INTERNATIONAL JOURNAL OF BUSINESS
Vol:3, 2 (2009) 51-56 ISSN 0974-0082

Web-Based Surveys: An Emerging Tool

T. Srivenkataramana[1] and M. Saisree[2]
1. UGC Emeritus Fellow, Bangalore University 2. Quantitative methods& Operations, ICFAI Business School, Bangalore

Abstract

With increasing web accessibility and popularity, Web-Based-Surveys (WBS) are becoming common and convenient. In spite of the lower cost and greater speed advantages, there are a few concerns with WBS, including sample randomness. This article reviews WBS scenario briefly, mentions advantages and concerns sources of error. The question of noncoverage is addressed in particular and its impact is examined and this study addresses the category of non-respondents. A few seed ideas for addressing the shortcomings of the WBS are given.

Key words and phrases: Bias; Dual-frame sampling; Internet; Noncoverage; Nonresponse; Randomness; River sampling; Web-based surveys

1.Introduction

Traditional modes of survey are observation, personal interview, telephone and mail. The merits and demerits of each of these methods are well known. These centre around cost, quality and feasibility considerations. With the advent of twenty first century an added dimension for surveying has been Web-Based Surveys (WBS). Internet has truly democratized the survey taking process, where that anyone with a website can conduct a survey quite easily. Thus the ability to conduct surveys is no longer restricted to a few organizations at centers of power in the society, such as Government Departments, Institutions and other bodies. As a result of this democratization, the range and quality of such surveys vary considerably. This variation can become a cause for concern in a quality conscious set up.

2. Two Broad Types of WBS

It is useful to group all WBS into two types: non-probability based and probability based. In the former, the units of target population do not have known, non-zero probabilities of selection. Hence, generalizations to the population are based on hope rather than sound statistical reasoning. Surveys on the web run the gamut from entertainment questionnaires to those with a probability based designs.[8]

Acknowledgement: The authors appreciate the constructive comments of the anonymous referees which have substantially improved the contents of the paper. The first author is grateful to the University Grants Commission for the Emeritus fellowship support. The second author is grateful to ICFAI Business School, Bangalore for the encouragement and support.

2.1 Non- Probability Approaches of WBS

Couper(2000) includes the following under this category:

(a) WBS as entertainment

These surveys consist of questionnaires that request a vote on particular questions or other instant polls. Often there is no control over what questions are posted or who responds.

(b) Self selected WBS

These use open invitations on portals, frequently visited websites or dedicated survey sites. This is arguably the most common form of WBS and also the most threatening to legitimate survey efforts in terms of quality claims. Often these surveys do not have access restrictions or control over multiple submissions.

(c) WBS of volunteer panel of internet users

To begin with, a volunteer panel is formed by wide appeals on popular websites and portals. Basic demographic information is obtained from the volunteers, to create a database of possible respondents for surveys later. Unlike in the first two types of WBS, access to surveys is by invitation only, monitored through e-mail identifiers or passwords. Selection of members from the panel for a particular survey may be by quota sampling or by random sampling. This appears to be currently the fastest growing WBS segment.

2.2 Probability Based Approaches of WBS

These methods begin with probability sampling of some form. However, due to web access being not universal and non-existence of a complete list of web users, there are two options to draw probability based samples:

1. Restrict sampling to only those with web access. This limits the target populations, often seriously.
2. Augment sampling reach by using a mixed mode; e.g., by Random Digit Dialing (RDD).

In this scenario, a few approaches are outlined.

(a) Intercept surveys

These generally use systematic sampling to invite every k^{th} visitor to a site to participate in the survey. Multiple invitations to an individual may be blocked through the use of cookies. Customer satisfaction surveys provide common example of application.

(b) WBS with e-mail request for response

This typically begins with a list of those with Web access. E-mail invitations are sent to the members of the list to participate, with access controlled to avoid multiple or proxy submissions.

(c) Mixed Mode surveys

This approach views web as an option for a respondent, among a few alternatives like telephone or mail. This method is particularly popular in panel surveys of establishments (firms or businesses) involving repeated contacts over a time span. Reducing response burden and time are major concern in this category.

(d) Pre-recruited panels of internet users

This differs from web panel creation of volunteers (under nonprobability approach) where panel members are recruited using random sampling such as RDD telephone surveys. Telephone is used to get background information check for internet access and then recruit into the survey panel. The motive is to obtain a probability sample of internet users. After consent to participate, the panel members are e-mailed a request for survey participation. Electronic controls are used to ensure that only those who are invited to do so participate and do so just once.

(e) Probability samples of full population

This approach has potential for obtaining a random sample from the target population, by augmenting Web penetration through providing the necessary equipment and tools to respondents, who do not otherwise have this facility. Thus this approach allows generalization beyond the current population of internet users. The high cost of recruitment is a deterrent factor of this approach. Nonetheless this approach has the potential to be an alternative to the traditional methods of random sampling.

3. Advantages of WBS

As compared to the traditional approaches, WBS have the following features:

a. High potential and a low cost.
b. Faster responses; no physical distances.
c. Ease of sending reminders to participants
d. Easier process data, since responses could be downloaded to a spread sheet, data analysis package or a data base.
e. Dynamic error checking capability
f. Option for posing questions in random order.
g. Scope for making complex skip pattern questions easy to follow.
h. Scope for pop-up instructions for select questions and manipulations that are possible in the electronic medium.

4. Concerns with WBS

WBS present a double edged sword for the survey industry. On a positeve note, WBS has democratized survey facility. On the negative side, the risk with WBS is that their proliferation makes it very difficult to monitor the quality aspect, leading to a mix up of the good, bad and the ugly. Ensuring near randomness of the samples is also a challenge with WBS. A few specific shortcomings of WBS are listed below:

a. The sample drawn is not really a random sample, unless special care is taken to do so.
b. Privacy concern of the data entered by the respondents, internet access being worldwide and easy.
c. Data security concerns of the surveyor on the server.
d. Lack of computer skill of respondents as a source of error.
e. The questionnaires do not look alike in different browsers and monitors. This may be a source of response bias.

5. Sources of Error

The use of WBS has increased dramatically, but the growth has not focused on survey error reduction.[4][5][6] The chief error sources in survey are sampling method, coverage of population, nonresponse and what is actually measured[2]. In fact coverage is a big concern in WBS. Coverage error is the biggest block to infer from WBS data, at least to m groups beyond access to or use of the web. The coverage error is a function of the mismatch between the target population and the sampling frame, as demonstrated by the following simple result.

5.1 Bias Due to Non-coverage

Let *P* be the web penetration (proportion of units with web access in the target population) and *t1* be an unbiased estimate of the mean of the study variate in the sampled population segment. The bias in *t1* as an estimate of the target population mean *M,* is given by

$$B(t1) = M1 - M$$
$$= M1 - [P(M1) + (1 - P)M2]$$
$$= (1-P)(M1 - M2)$$

where, the suffixes 1 and 2 refer to sampled and non-sampled population segments. Thus the magnitude of bias depends on two factors:

a. Extent of non-penetration *(1-P)*
b. The difference between the means of the two segments *(M1-M2)*

This result is indicative of the magnitude of bias due to non-access of web to the respondents, when *(M1-M2)* is substantial. When web penetration is low and non penetrated population segment differs drastically in the study characteristic, the non coverage bias can be serious. For instance, with 30% web penetration the bias due to non-coverage alone will be *0.7(M1-M2).*

Table 1 gives a numerical picture of Bias.

P	*1-P*	10	30	50	70	90
.20	.80	8	24	40	56	72
.40	.60	6	18	30	42	54
.50	.50	5	15	25	35	45
.60	.40	4	12	20	28	36
.80	.20	2	6	10	14	18

⇐---------- I M2-M1 I ---------- ⇐

Table 1 Absolute Bias for Typical Cases

6. Nonresponse

Nonresponse is due to unwillingness of sample respondents to complete the survey. Generally WBS has a

lower response rate than mail surveys. [2][12] Abandonment or drop out is also a major concern with WBS.[3] There are several reasons for respondents not completing a WBS. Researchers have identified some of these reasons as follows:

a. Use of open ended questions.
b. Questions arranged in tables and fancy designs.
c. Unclear instructions and want of navigational aids
d. Using a complex grid of questions and responses.

7. Types of Responders

A WBS respondent is supposed to understand the question, recall necessary information to answer it, make a judgment and then select a response. Redline and Dillman [10] cited work by Krosnick (1991) noted two respondent types:

1. *The optimizers:* Who devote their full attention to survey completion.
2. *The satisfiers:* Who go through the motions of answering the questions but look for ways to expend as little effort as possible for the survey.

7.1 A Classification of Responders

An empirical analysis by Bosnjak and Tuten[1], shows seven distinct response types in WBS.

a. *Complete responders:* View all the questions and respond to them.

b. *Item nonresponders:* View the entire questionnaire but only answer some of the questions.

c. *Item non responding drop outs:* View some of the questions but answer some but not all the questions viewed and also quit prior to the end of the survey.

d. *Answering drop outs:* provide answers to those questions displayed but quit prior to completing the survey.

e. *Lurkers:* view all the questions of the survey but do not answer any of them.

f. *Lurking drop outs :* view some of the questions without answering and also quit prior to reaching the end.

g. *Unit nonresponders :* Do not participate in the survey.

For any specified target population, estimating the proportions in the seven categories and taking corrective steps to reduce the proportions in the non responding groups will be helpful to tackle the incompleteness of the sample data. This type of online monitoring as an advantage is not available with traditional survey methods like a mail survey.

8. Uniqueness of WBS

As compared to a traditional survey, WBS have the following special features:

a. *Use of self administered questionnaires:* This is the so called self reporting method.[13]

b. They are a visual stimulus, and respondents have control over how and even whether they read and comprehend each question.[6]

c. Participants in WBS are less likely to take extreme positions in their responses than those in telephone survey.[11]

d. WBS provide opportunities for variety in question structure, layout and design unlike in paper surveys[2][3]. It has been demonstrated that there are various ways to manipulate both the verbal and auxiliary languages of self – administered questionnaires to improve design of skip instructions and thereby improve the response rate for skip pattern questions.

9. Business Applications

WBS have high potential for use in the business set up. A few typical situations are the following:

a. *Employee satisfaction surveys:* These are mostly used in corporate offices to measure their employee pulse, voice of work force , service quality etc. WBS as a mode for employee satisfaction survey has the advantage of giving the respondents dynamic options. For example, in a multiple choice question, the next set of questions, appear depending on the response chosen by the responder. Thus every option in the question leads to different sets of subsequent questions.

b. *Customer satisfaction surveys:* The pressure to improve customer satisfaction is always high. Selecting the

best medium for customer survey is not easy. WBS as a mode for evaluating customer satisfaction has the advantage that once set up , there are far fewer manual resources involved. The customers are directed to the relevant web page where they complete the questionnaire online and the results are collected and analyzed electronically. WBS are the least expensive and least impacted by volume of all survey methods. After the initial set up costs (which are low) the schemes are fully automated and these surveys cost little to maintain.

c. *Patient satisfaction surveys:* Employed by hospital managements to monitor the patient satisfaction level with hospital services periodically for in- house as well as out- patients.

d. *Electronic marketing:* This is employed by leading market brands and market research firms especially in the urban areas. Product or a concept is advertised by means of a WBS.

e. *Opinion polls:* In countries like the USA where the web penetration is high, WBS is used to conduct opinion polls.

10. Addressing Some Key Issues

The shortcomings of WBS can be tackled in several ways. They are briefed here.

The proponents of WBS point to the fast growth rate of the internet in support for their optimism. In developed countries like the USA the web penetration is assessed to be over 40% for the adult population. In India it is about 4% (48 million active internet users, 2009) and the optimism regarding the potential of WBS is therefore based on the projected trajectory of future web penetration, which is expected to grow by a factor of *three* by the year 2012.

10.1 Noncoverage

The following ideas may be useful to reduce the noncoverage.

a. *Augmented penetration:* This is done by providing the necessary access or tools to responders who are not having this facility. However this is bound to escalate the costs.

b. *Augmented reach:* The web sampling reach may be augmented by using a mixed-mode design where the responders have a choice of completion method. For example, through mail, email, telephone etc.

10.2 Nonrandomness of Samples

There are misguided presumptions behind several WBS that

1) large samples imply more valid responses,
2) the sample size is the only determinant of the sampling error, and
3) statistically valid inference is possible with nonrandom samples too.

Unless the population is perfectly homogeneous, that is, the units are alike with respect to the study characteristics (which is an unlikely event in survey practice), these presumptions do not hold. Thus valid generalization of sample results to the target population crucially hinges on randomness (at least near randomness) of the sample. The options available for improving randomness in WBS scenario include the following

a. *Dual- frame sampling (Kalton,2001):* The dual-frame Web – phone designs are cost effective for sampling rare groups, say no more than 20%, within a general population. The set of all web users may be treated as a rare group in the Indian set up now.

b. *River sampling:* River sampling has been positioned as the best option for reaching a random, less surveyed online audience. River sampling recruits using banner ads , pop up ads and similar instant "capture promotions". Individuals who volunteer to participate are screened for their reported demographic characteristics and then "randomly assigned" to the appropriate survey. Hence, the metaphor of" being captured" from the flowing river of online persons is used. (DiSorga,2008)

10.3 Handling nonresponse:

Non response in WBS can be handled slightly differently as compared to a traditional survey .Several concomitant steps can be used to improve the response rate . Some of these are as follows;

a. In view of the classification of responders (refer Sec.7), a pilot study may be used to assess the

proportions of responders in the seven categories. In turn, monitoring steps may be taken to improve the response proportions in the "desirable" categories, (*e.g.* category of *complete responders).*

b. Simple formats and plain design of the questionnaires.

c. Pre-notification of the survey purpose, clarifying confidentiality of the responders and generality of the final inference.

d. Personalized e-mail covering letters and follow-up reminders.

e. Incentives for satisfactory participation by the responders.

f. Use of web as a communication medium is constrained by the asynchronous and relatively impersonal communication common with the internet. This implies that on- line monitoring and giving a personal touch to WBS will help in reducing nonresponse.

11. Conclusion

WBS has added a new dimension to survey modes, which is now being used with increasing frequency. Reduced time and cost for data collection and efficient analysis are the major advantages of this electronic mode, while the issue of randomness of the sample, privacy for respondents, data security and nonresponse are serious concerns. A few seed ideas for addressing these issues have been outlined in Sec.10. There is scope for further work, both theoretical and empirical, centered around these ideas. This will improve the quality aspect of derived estimates and make WBS a valuable addition to the kit of quality minded survey statistician. The research effort in this context must come in the form of joint work by persons with a strong technology background and experts in survey methodology, since the two aspects are intertwined here. This can eventually provide a faster survey mode with reliable end results, a very desirable outcome indeed.

References

1. Bosnjak, MM and Tuten, TL (2001). Classifying response behaviours in web based surveys. Journal of computer mediated communication, Vol6(3). http://www.ascusc.org
2. Couper MP. (2000): A review of issues and approaches, Public opinion quarterly, Vol.64(4) pp 464-481.
3. Couper MP, Traugott MW and Lamias, MJ (2001). Web survey design and Administration, Public opinion quarterly, Vol 65(2), pp 230-253.
4. Dillman, DA (2000). Mail and internet surveys: The tailored design Methods, 2nd edition, Wiley , New York.
5. Dilman, DA and Bowker, DK (2001). The web questionnaire challenge to Survey methodologists.: http//survey.sesrc.wsu.edu
6. Dilman, DA et al (2001). Response rate and measurement Differences in mixed mode surveys using Mail, Telephone, Interactive Voice Response and Internet: http//survey.sesrc.wsu.edu.
7. DiSorga,C (2008).River samples: A good catch for researchers? www.knowledgenetworks.com.
8. Holly Gunn (2008). Web-based Surveys: Changing the survey Process. First Monday journal on the internet, Vol 7(12) pp 1-17.
9. Kalton,G(2001) Practical methods for sampling rare and mobile populations: Proceedings of American statistical Association, Section on Survey Research methods.
10. Redline, CD and Dilman , DA (1999). The influence of Auxiliary, Symbolic, Numeric and Verbal language on navigational Complience in self administrated questionnaires: http//www.survey.sesrc.wsu.edu.
11. Satmetrix, (2001). Investigating validity in Web surveys. http//www.satmetrix.com.
12. Solomon, DJ (2001). Conducting Web-based surveys: Practical assessment, Research and Evaluation, Vol 7(19). http//www.ericae.net.
13. Srivenkataramana, T (2007). "Self reporting method in business surveys", Dharana- Bhavan's International journal of Business, Vol.1 (1) Pg 38-40.