

Adjusting responses in a Non-probability Web Panel Survey by the Propensity Score Weighting

The National Center for University Entrance Examinations

Osamu YOSHIMURA

Keywords

Web survey, non-probability online panel, probability online panel, the propensity score, weighting

Abstract

As is often the case with non-probability Web surveys, responses differ from other traditional surveys. Difference in response to the same question in two different surveys may be caused by differences in survey modes, survey administrators, survey periods, design and(or) layout of questionnaire etc, and sample unbalance. If two surveys differ in only sample, a survey result could be adjusted to another survey by weighting. The propensity score weighting is a useful technique when the number of covariates to be considered is large.

In order to examine the availability and effectiveness of the propensity score weighting in practice, an experimental Web survey was conducted on two different panels, and then tried to adjust responses in non-probability panel to that probability one. The response variables were owner rates of fourteen kinds of information equipments, and the covariates adopted in estimating the propensity weights were ten demographic variables. The possibility of a general weight that could adjust responses in a non-probability Web survey and the effective usage of the propensity score weighting in Web surveys were discussed.

Introduction

Recently in Japan, it has become hard to conduct traditional sample surveys. With the deterioration of the survey environment, attempts to collect data on the Web, known as Web surveys, have increased dramatically. However most of them are 'non-probability surveys' (Couper 2000, 2001).

It is well known that one of the biggest problems in current Web surveys is selection bias of respondents. Moreover, it is pointed out that most respondents in Web surveys are professional

respondents. This means that the surveys are not scientific.

In a scientific survey, a sample should be representative of a population. A self-selected sample, such as a volunteer online panel which is typically used in a Web survey, has no way of ensuring the representation of the population. Therefore any statistical inference from the sample cannot be drawn. The best solution is to randomly select a sample from the population, but it would require incredible resources and be of huge cost.

The second solution is to adjust the selection bias of the sample. Recently, there is an argument (e.g. Taylor 2000) that the problem of selection bias in Web surveys can be solved by adjustments using a weighting technique, known as, the propensity score weighting (the PSW) (Rosenbaum and Rubin 1983, Rosenbaum 1987). This claim could be possible. However, the claim of the propensity score weighting is, usually, made based on studies which succeeded in adjusting results of Web surveys to RDD survey results. Is it possible that the adjustments could be done despite different survey modes?

In general, differences in response to the same question in two different surveys may be caused by not only sample unbalance, but also by various other factors. These may be survey mode, survey administrator, survey period, design of questionnaire, layout of pages, and so on. When there is no difference in all these factors, we may say that the differences in responses would be caused by sample unbalance. In order to examine the efficacy of the weighting technique, the adjustment should be adopted to responses in two surveys under the same conditions.

In this study, we conducted a probability and a non-probability Web survey and adjusted the responses in the non-probability survey to the probability survey in order to examine (1) to what extent bias in responses in the non-probability panel will be reduced, (2) whether the covariates used to balance samples will be easily identifiable, (3) what kind of variables will be adopted as covariates.

Method

Since 1998, our research group has conducted a series of experimental Web surveys (Ohsumi and Yoshimura 1999, Yoshimura and Ohsumi 2000). One of these, our most recent surveys in 2002, gave us an opportunity to hold two Web surveys with a non-probability panel and a probability panel, respectively. The surveys were administered with the same format questionnaire, page design, Web server, and at the same periods, and were conducted by the same administrators, but on two different panels. Consequently, a difference in the results between these two panels can only be caused by sample unbalance. Therefore, adopting the PSW to the adjustment of the responses in the non-probability survey, we can know the efficacy of the technique.

The non-probability panel was a volunteer online panel and the size of the panel was 13,201. The probability panel was a pre-recruited Internet user panel and the size of the panel was 10,852. We conducted a survey of two samples randomly selected from each panel. The sample size and response rate were 849 (59.1%) and 693 (56.6%), respectively.

The question was "Which information technology items do you own?", and respondents checked all the items that they owned from the 14 listed items: cellular phone, PHS, PDA, notebook PC, desktop PC, laser printer, non-laser printer, Hi-vision TV, facsimile, game console, car navigation system, digital camera, digital video camera and PC communication. We will refer the selection rate of each item to a 'response variable'.

In this study, we will approximate the values of the response variables in the probability survey by weighting the response in the non-probability survey

using the PSW. We can examine the efficacy and practical applications of the PSW by comparing the approximation with observed data.

Now, let G_1 be the respondents from the non-probability panel (size n_1) and G_0 be the respondents from the probability panel (size n_0 , $n_1+n_0=N$). Y_i ($i=1,,N$) refer to a response variable of a respondent i , and Y_{1i} shows the respondent is from G_1 , and Y_{0i} shows i is from G_0 . Let g be a binary variable and if the respondent i is from the non-probability panel, g_i will be 1, else will be 0. Additionally, let \mathbf{x} be a covariate vector of Y . Then $E(Y)$ can be calculated by the following equation (Rosenbaum 1987). This means that we could know $E(Y)$ even if Y_0 is not observed.

$$E(Y) = \frac{1}{N} \sum_{i=1}^N \frac{g_i Y_i}{\hat{e}(\mathbf{x}_i)} \dots\dots(1)$$

where $e(\mathbf{x}_i) = \Pr(g_i = 1 | \mathbf{x}_i)$

In the context of this study, we need the estimation of the mean of response variables in the probability survey, that is $E(Y_0)$ instead of $E(Y)$. Yoshimura and Shojima (2004) showed that $E(Y_0)$ is obtained by the following equation. We can examine the efficacy and practical applications of the PSW by comparing the observed mean of Y_0 . with it's approximation obtained by equation (2).

$$E(Y_0) = \frac{1}{n_0} \sum_{i=1}^N \frac{g_i Y_i (1 - \hat{e}(\mathbf{x}_i))}{\hat{e}(\mathbf{x}_i)} \dots\dots(2)$$

$$N = n_0 + n_1$$

The following 10 variables were considered for covariates of the response variables: gender, age, marital status, occupation, housing, family structure,

household income, Internet access locations, Internet connection type and length of Internet experience. Among these, a set of covariates for each response variable were selected by the following way.

- (1) Select a set of likely variables,
- (2) Estimate the propensity score,
- (3) Calculate the approximation of $E(Y_0)$,
- (4) Evaluate the difference between the approximation and observation of $E(Y_0)$,
- (5) Repeat (1) to (4) until the root squared difference between approximation and the observation would be less than 1%.

Results

Despite that both of the two samples consist of Internet users, they differ from each other in demographics and other variables. Comparing with the probability panel, we found the following characteristics of the non-probability panel.

- The proportion of male is higher in the non-probability panel than that in the probability panel (64.7% and 49.3%, respectively).
- Half of the respondents in the non-probability panel are in their 30s.
- The distribution of marital status doesn't differ from each other.
- In the non-probability panel, the rate of technicians is considerably high (22.1%).
- As to Internet access locations, the rate of offices in the non-probability panel is higher than that in the probability panel(23.2% and 10.7%, respectively).
- Respondents in the non-probability panel

have a better Internet access environment.

- Respondents in the non-probability panel have much more experience than those in the probability panel.

These differences between the two panels mean that the non-probability panels obviously do not represent the Internet user. As we saw above, since the two samples differ from each other in demographics and other variables, the survey results would also differ from each other.

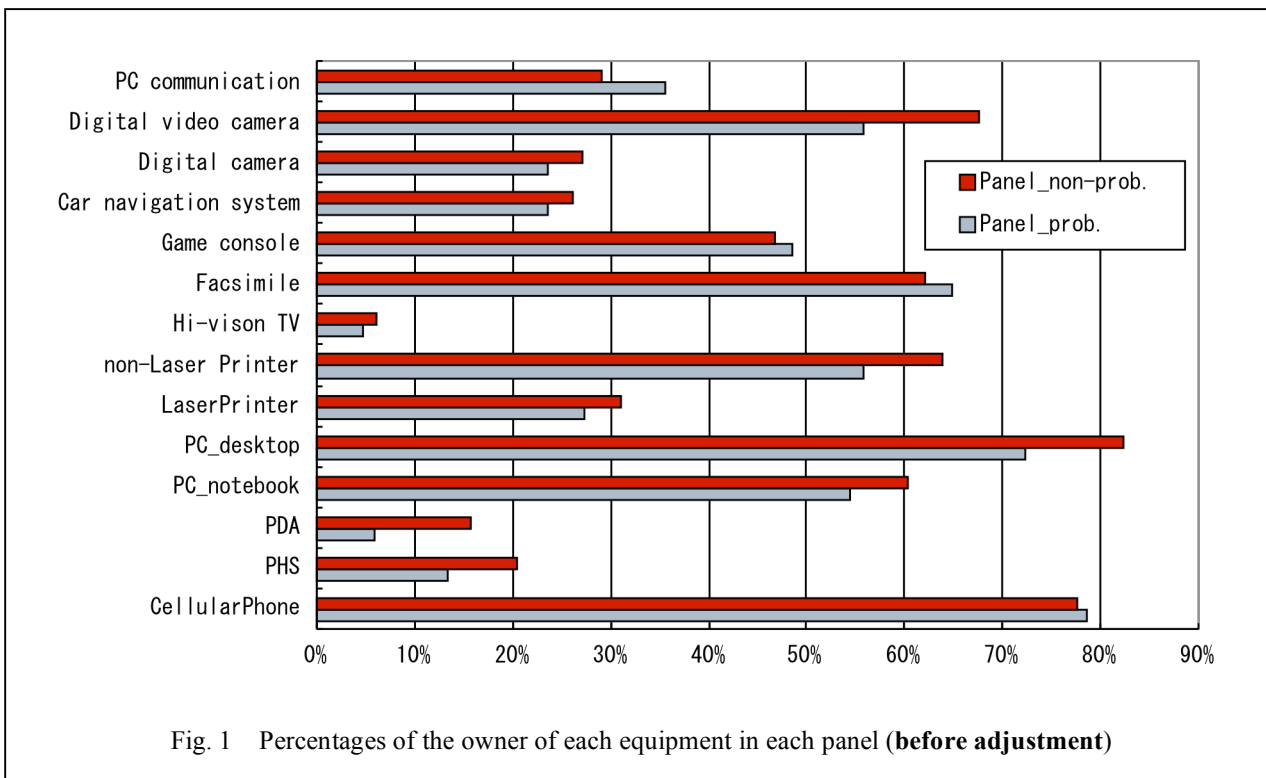
Fig. 1 shows the percentages of the owner in respondents in each panel. As we can see, some of the differences can not be ignored. It is obvious that the results in the non-probability panel cannot be substituted with those in the probability panel. This implies that the results of the Web survey with a non-probability panel could mislead us to a wrong conclusion.

Fig. 2 shows the rate of owners after adjustment of the responses in the non-probability panel. In this figure, we can see the differences between the two panels are greatly reduced.

The root squared difference (RSD) of selection rate of each item in the two panels was calculated. On average, 81% of the differences between the two panels were reduced. We can say that the adjustment has been successful in most of the variables. However, the adjustment of the responses to Hi-vision TV ended in failure. It is likely that we failed in identifying and selecting covariates of the variable.

The sets of covariates of response variables were as follows:

- Digital Video Camera: gender, age, age by gender, marital status
- Cellular Phone: gender, age, occupation



- Car Navigation System: age, age by gender, occupation
- Laser Printer: age, age by gender, occupation, family structure
- Notebook PC: age, age by gender, occupation, family structure, household income
- Facsimile, Game Console, PC Communication: gender, age, occupation, family structure, access location, connection type
- non-Laser Printer: age, age by gender, occupation, family structure, housing, household income, length of Internet experience
- PHS, PDA, Desktop PC, Digital Camera: age, age by gender, occupation, family structure, housing, household income,

access location, connection type, length of Internet experience.

As shown above, a set of covariates used to balance the two samples varied with every response variable. This means that there is no one single set of covariates we could use to adjust any response variables.

Discussion

In this study, we adopted the PSW to a non-probability Web survey data, and approximated the data in a probability Web survey. Based on the results of the study, we can conclude that the PSW works well, if and only if the covariates of the variables are observed and identified. It is easy to say that identifying covariates is crucial for the PSW, however, it is impossible to identify all the covariates before carrying out surveys. We can only empirically

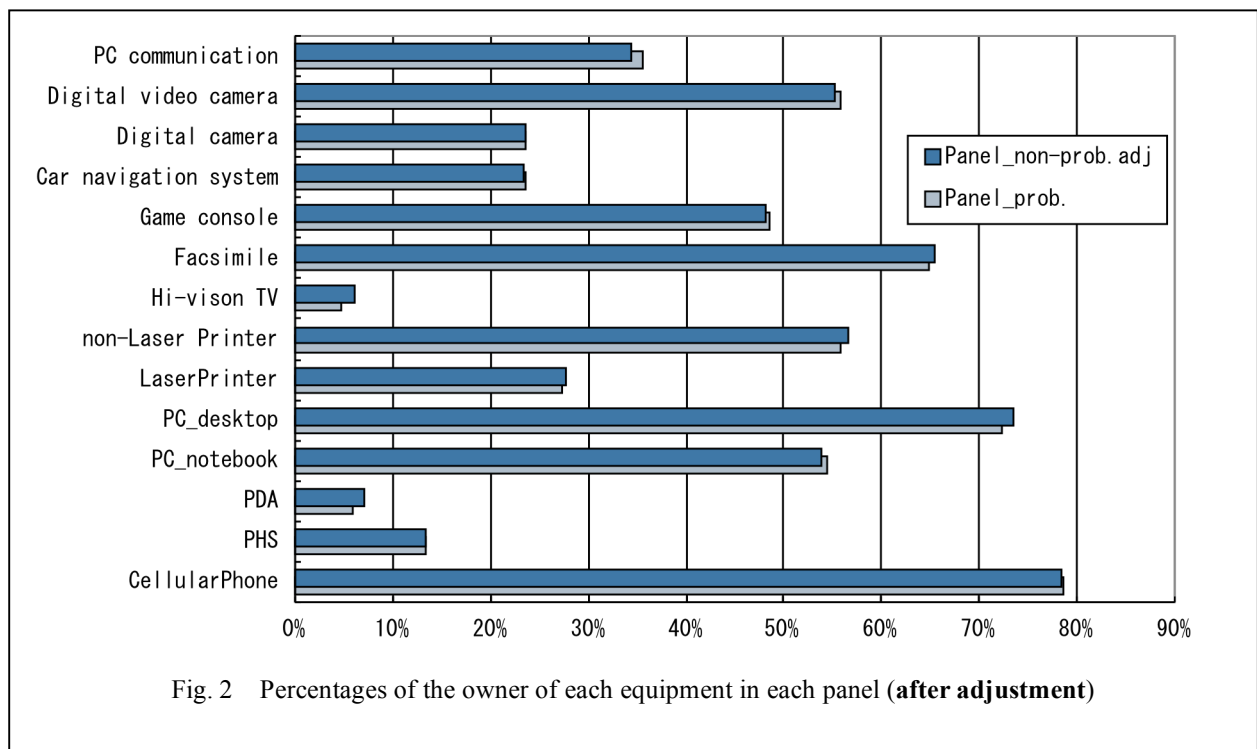


Fig. 2 Percentages of the owner of each equipment in each panel (after adjustment)

know what the covariates are by conducting surveys repeatedly. The PSW is undoubtedly a powerful tool for adjusting Web survey data. However, before adopting the tool to real Web survey data, we have to examine the stability of each relation between a response and its covariates, and also we have to confirm to what extent the stability is affected by survey methods. Moreover development of efficient methods to identify and select covariates would be needed.

Without such verifications, no one can claim the efficacy of the technique. We should be worried that mercenary motivated Web surveys would result in the collapse of the reliance on surveying itself. It is to be desired that misuse of the PSW would not accelerate the collapse of the reliance on surveys.

References

- Couper M.P (2000) Web Surveys: A Review of Issues and Approaches. *Public Opinion Quarterly*, 64 (4), 464-494.
- Couper M.P (2001) The Promises and Perils of Web Surveys. In A. Westlake et al. (eds), *The Challenge of the Internet*. London: Association for Survey Computing, 35-56.
- Ohsumi, N. and Yoshimura, O. (1999) The online survey in Japan: An evaluation of emerging methodologies, *Bulletin of the International Statistical Institute 52nd Session, Book2*, 171-174.
- Rosenbaum,P.R. and Rubin.D.B. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70 (1), 41-55.
- Rosenbaum,P.R. (1987) Model-Based Direct Adjustment. *Journal of American Statistical Association*, 82, 387-394.
- Vehovar, Vasja, Batagelj, Zenel, and Lozar, Katja (1999) Web Surveys: Can the Weighting Solve the Problem? *Proceedings of the Section on Survey Research Methods*. American Statistical Association, Alexandria.
- Taylor, H. (2000) Does internet research work? *International Journal of Market Research*, 42(1) 51-53.
- Yoshimura, O. and Ohsumi, N. (2000) Some experimental surveys on the WWW environments in Japan, *Data Analysis, Classification, and Related Methods* (eds. H. Kiers, etc.), 353-358, Springer, Berlin.
- Yoshimura, O. and Shojima, K. (2004) Comparing difficulties of a main test and it's supplementary by the propensity score weighting. *Research Bulletin*, 33, 19-28, The National Center for University Entrance Examination, (in Japanese).

E-mail: osamu@rd.dnc.ac.jp