

Social Participation and Intent to Participate in Internet Surveys

U-Seok Seo
Department of Urban Sociology
University of Seoul
usseo@uos.ac.kr

Gihong Yi
Department of Sociology
Hallym University
gihongyi@hallym.ac.kr

Abstract

A Statistics Korea survey of 2009 asked a unique question about intent to participate in social surveys via the Internet in the future. Analyses first show that demographic variables such as sex, age, education, employment status, and marital status explain the intent. In addition, the frequency of reading Internet newspapers and some of social participation factors, such as non-political donation and social group participation, also do so. This suggests that Internet surveys, as they increasingly continue to complement traditional offline surveys, may include or exclude certain characteristics in the general population. This in turn may imply that sampling needs to be more carefully planned and executed to minimize potential bias issues, especially for government surveys.

Introduction

Social surveys via the Internet have greatly increased in Korea over the last decade. At first, market analysis surveys using online panels proliferated. Afterwards, web-based social surveys started to spread in various areas. Much of government statistics are nowadays collected through Internet surveys. Statistics Korea (formerly the National Statistical Office of South Korea) intends to increasingly replace what used to be face-to-face interviews with Internet surveys for the census and many other government surveys.

The idea behind increasing Internet surveys is, most of all, based on the fact that most Koreans already have convenient access to the Internet. In 2011, the rate of Internet use in Korea is 78.0% (KISA, 2012)¹. Particularly remarkable is that the figure reaches 99% in the

¹ This KISA and many other Korean surveys estimate Internet use based on the time frame of the previous month. Such a measurement of Internet use rate is different from that of the ITU, which refers to the previous year. According to the ITU criterion, the rate of Internet use in Korea is 83.8%.

age groups 10s through 30s. Potential coverage bias due to the digital divide is practically of little or no concern at least for these generations. Although a certain level of age gaps still exist in Internet usage, they are rapidly decreasing. Even constructing senior online panels has become quite feasible unlike in the past.

The rapid rise of Internet surveys may be explained by increasing difficulties from the traditional ways of survey conducting. For instance, as the usage of landline phones dramatically drops with an explosive increase in the mobile phone population, young people, who definitely prefer using mobile phones, have become difficult to contact based on the conventional know-hows of telephone surveys. Skepticism on the conventional data collection grows due to this kind of inevitable changes in the communications environment. Even before factors like this, face-to-face surveys began to suffer from an increasing refusal rate due to lifestyle diversification and privacy concerns. As newly built apartment home complexes strengthen security measures to attract privacy-sensitive residents, survey interviewers sometimes cannot at all obtain access to quite a large portion of sample in certain areas.² In some cases, interviewers conducting the official government census are refused at the gate of apartment home communities. Such changes clearly make the Internet an attractive alternative in diverse data collection situations.

Academia of course pays attention to the peculiarity of Internet surveys. After extensive discussions including special issues devoted to relevant topics in the *Sociological Methods and Research* 2009 and the *Public Opinion Quarterly* 2008, advantages of Internet surveys achieve strong recognition (Frippiat and Marquis, 2010). Internet surveys enable researchers to complete data collection with less cost and time, in addition to the fact that certain respondents are otherwise inaccessible. Even monitoring the ongoing process of data collection can be done much more effectively if the Internet is chosen as the medium.

Some issues, however, remain controversial with Internet surveys. Some sociologists are reluctant to use online survey data (Farrell and Petersen, 2010) while marketing research extensively uses Internet surveys with pre-constructed panels. It is most likely because researchers slightly disagree in evaluating sample representativeness.

As far as Internet surveys are concerned, one of the most prominent issues is sample representativeness. Although some online panels are based on probability sampling³, much more Internet surveys rely on 'volunteers' or 'convenience' sampling (Taylor, 2000).

Two points require attention with regards to the representativeness of Internet surveys. One is coverage error which stems from the fact that some people have no access to the Internet at all. In the case of South Korea, this is relatively unimportant as such a vulnerable population is decreasing rapidly. The other is self-selection bias. Most Internet surveys are conducted in the formality of self-administration (Couper et al., 2007). Those who participate in Internet surveys mostly volunteer, meaning that the refusal rate, among other things, is difficult to estimate or control.

² A 59% of South Koreans live in apartments, which are far less accessible to than houses, as of 2010 according to Statistics Korea.

³ Couper (2000) introduces eight types of web surveys, among which five belong to probability-based models.

Many attempts have been made to improve sample representativeness for Internet surveys. Noteworthy is the success of Harris Interactive in the 2000 US presidential election. Their usage of online panelists and propensity score adjustment results in quite an accurate prediction. Conventionally, demographic variables are analyzed for weighting and/or post-stratification; Harris Interactive does more. To obtain sufficient information for propensity score adjustment, Harris Interactive conducts a reference survey based on probability sampling towards the same target population. Since this kind of success, many studies revisit diverse issues on propensity score adjustment for Internet samples (e.g., Schonlau et al., 2009; Lee & Valliant, 2009; Valliant & Dever, 2011). Controversies, perhaps only inevitably, appear to persist regarding the effectiveness and validity of this method (Baker et al., 2010). South Korea also produces a delicate set of studies examining the usefulness of adjustment methods (Kim and Lee, 2003; Lee and Jang, 2009; Huh and Cho, 2010).

Detailed information on the selection of variables for adjustment purposes often remains undisclosed as in the case of Harris Interactive. Even if such information is provided for public discussion, replicating it to other research situations might be unfeasible because every single sampling effort is subject to many, often-unsortable factors. A few success cases remain insufficient for theoretical justification and generalization.

The selection of variables for the adjustment of Internet samples deserves deliberation from diverse perspectives. We suggest a sociological one to better understand who shows intent to participate in Internet surveys based on a South Korean government survey. As a truly major method on equalizing face-to-face and Internet samples is yet to emerge, we suggest obtaining more empirical evidence that would eventually enable further comparison between those who accept future social surveys via the Internet and those who do not.

Social Participation and Intent to Participate in Internet Social Surveys

The increase of Internet surveys can be considered a hallmark in “the societal trend toward self-administration (Dillman, 2000)” in the area of survey research. One of the most important features of Internet surveys is the self-selection of respondents. As coverage error due to the digital divide is getting less of an issue in many countries, self-selection bias require so much more attention for the time being. Necessary is further explanations on respondent’s decision to “voluntarily” participate in Internet surveys.

From this point of view, we move to think about the sociological meaning of public opinion survey and raise a question about who is more willing to participate in social surveys and who is less or not so. An interesting and important starting point in this discussion is a strong belief in the relationship between public opinion and democracy, shared by key figures in survey research such as George Gallup and the founding editorial team of the *Public Opinion Quarterly* too (Shapiro, 2011). In this vein, theories on survey participation stress connections between civic duty and survey participation alongside the influence of economic incentive (Groves, Cialdini and Couper, 1992; Couper, Singer and Kulka, 1998; Groves, Singer and Corning, 2000). Motivation to participate in the US decennial census, for example, reflects a sense of civic obligation, which affects other forms of social participation including voluntary service and voting.

Given the role of topic interest in survey participation decisions (Groves, Presser and Dipko, 2004), it is realistic to assume that a sense of civic duty will be more salient in social surveys than in other kinds of surveys, of other topics. Furthermore, we can raise another question about if this kind of connection exists with respect to Internet surveys as well. As major theoretical perspectives on Internet survey participation are yet to emerge, empirical analyses on this issue need to be accumulated to a further extent than what now seems available. We hope that this study contributes as one of empirical grounds in clarifying sample representativeness issues regarding Internet surveys.

As Internet surveys require a stronger volunteering attitude than traditional face-to-face surveys, the sense of civic duty may even play a greater role in explaining intent to participate in Internet surveys. The relationship between civic duty and survey participation might even be strengthened in the case of Internet surveys considering that the Internet is viewed as a very popular channel for everyday political expression and a potentially very effective tool for mass mobilization as in discussions of “collective intelligence”. Particularly in Korea, the impact of online activities on social/political participation has been repeatedly highlighted in the last several years, e.g., the 2002 World Cup street festival, the 2008 candlelight cultural festival, etc. This study intends to ultimately contribute to discuss the relationships among civic duty, social participation, and intent to participate in social surveys via the Internet.

Analysis: Predicting Intent to Participate in Social Surveys via the Internet

This study utilizes a 2009 South Korean government survey data set including a question on future intent to participate in social surveys via the Internet. This data set is adequate for research that deals with diverse factors possibly relating to the intent as it also includes questions about culture, leisure, welfare, income/consumption, work, social participation, etc. This study investigates the influences of demographic variable (e.g., sex, age, education, job status, and marital status), frequency of Internet newspaper reading as a proxy of Internet usage, and social participation factors such as non-political donation, group participation, and volunteering. Descriptive statistics, crosstabulations, and logistic regression analysis are to be shown in this 2012 WAPOR conference paper.

< Table 1 about here >

< Table 2 about here >

Tables 1 and 2 show descriptive statistics of major variables used in this study. This data set is a survey of 37,049 South Koreans, who are of age 15 or older. The two tables tell basic information about the demography of the sample, and also basic distributions of the variables used in the analyses as follows.

< Table 3 about here >

Table 3 summarizes crosstabulations between intent to participate in social surveys via the Internet and each of the independent variables already mentioned. 18% of men show the intent while only 14% of women do so. Men are more interested in Internet surveys than women. When educational attainment is divided into the following 4 categories, elementary school or less, middle school, high school, and college or more, the best educated group shows the strongest interest (25%) in Internet surveys whereas the least educated group shows almost no interest (3%) therein. Internet survey interest shows difference in terms of job status too. The employed show a stronger interest (18%) than the unemployed (12%). Within the category of the unemployed, students show the strongest interest (24%) whereas the retired show the least interest (3%). This seems to suggest that workers and students better understand the benefit of the Internet as they have access thereto practically all the time at their work and school compared with the rest. By marital status, the married (14%) seem less interested in Internet surveys than the unmarried (18%), within which the never married are most interested (24%). This seems to at least suggest that marital status may be an explaining factor for the interest in Internet surveys for further analysis.

The crosstabulation of Internet newspaper reading frequency and the intent to accept Internet surveys shows that those who read Internet newspapers often are more likely to be interested in Internet surveys. The ones who read Internet newspapers almost every day show the strongest interest (29%) whereas the never group shows the least intent (7%). This appears to tell that those who are on the Internet more often for information collection are more likely to do other things, e. g., surveys, online too. Experiences such as non-political donation, group participation, and volunteering, which respectively could be considered an aspect of social participation, turn out to influence intent to participate in Internet surveys. Those who donate (21%), participate in various social group(s) (19%), or volunteer (24%) show stronger interest in Internet surveys than those who do not. This may mean that socially active people are so much more interested in voicing their opinions so that they are more willing to do Internet surveys. The above results suggest that the variables used herein are worth more exploring as independent variables in further analyses to see how they could affect intent to participate in social surveys via the Internet.

< Table 4 about here >

Table 4 shows the result of logistic regression analysis of intent to participate social surveys via the Internet, which is the dependent variable, versus 9 independent variables including logit coefficients (β), p-values (p), odds ratios (e^{β}), and model test results. This sort of logistic regression analysis is fit in discussing if the independent variables indeed statistically significantly influence the dependent variable and also, if they do, to what extent. The 3 model tests and the goodness-of-fit test above suggest such logistic regression is valid as an analytical model. All the independent variables except for sex turn out statistically significant.

Considering the independent variables altogether, the following briefly summarizes the result of the above logistic regression of intent to participate in social surveys via the Internet. Age negatively affect the intent, meaning that the older one gets, the less likely s/he participates in Internet surveys. This seems to suggest a generational divide issue. As old people are less used to doing things via the Internet than younger people are, they are less likely to participate in Internet surveys. Institutional education increases such chance, which appears to tell that education makes people more clearly understand the convenience and efficiency of Internet surveys. According to the analysis, having a job and being married also increase the likelihood of agreeing to Internet surveys. Working people seem to prefer Internet surveys as they are more time-saving and freely schedulable than face-to-face meetings with interviewers while working or resting. When marriage is considered as one of the independent variables like in this logistic regression analysis, its effect turns out different from the previous crosstabulation where married respondents show less interest than the unmarried category. Married people seem to prefer Internet surveys because, among many reasons, they are less privacy-invasive.

Internet newspaper reading frequency turns out to increase intent to participate in Internet surveys. This suggests that familiarity with the Internet as a medium positively affects intent to accept Internet surveys. Many everyday Internet users tend to transfer everyday tasks to online, e. g., banking, shopping, etc., and survey participation may be next for some of them. Non-political donation, group participation, and volunteering are all alike in increasing the likelihood of wanting to be surveyed over the Internet. Considering these three factors, each as a proxy of social participation, this result tells that those who are socially active tend to prefer Internet surveys than those who are not.

Overall, the above results suggest that intent to participate in social surveys via the Internet varies not only by demographic factors such as sex, age, education, employment status, and marital status. Familiarity with the Internet and social participation-related factors such as donation, group participation, and volunteering also positively affect intent to accept Internet surveys.

Conclusion and Discussion

Above analyses show that intent to participate in Internet surveys is dependent on not only many demographic characteristics, but also Internet familiarity and social participation. Considering that most Internet surveys so far collect information from volunteers, Internet samples, while they show merits in terms of time, cost, etc., need to be more carefully designed through further studies. Especially when it comes to government data, the best level of sample representativeness needs to be maintained as they function as basic evidence for many public decisions. In so doing, researchers should go beyond considering just demographic variables as social aspects such as Internet familiarity and social participation positively affect intent to participate in Internet surveys.

As far as the result of this study is concerned, Internet surveys should strive to retain those who are less Internet-familiar and/or less socially-participating in the respondent pool to minimize sampling bias. More research needs to follow to discover what other social factors may influence intent to participate in Internet surveys.

Bibliography

- Barker et al. 2010. Research Synthesis. AAPOR Report on Online Panels. *Public Opinion Quarterly*, 74(4): 711-780.
- Couper, M.P. 2000. Web Surveys: A Review of Issues and Approaches. *Public Opinion Quarterly*, 64(4): 464-494.
- Couper, M.P., Kapteyn, A., Schonlau, M. and Winter, J. 2007. Noncoverage and Nonresponse in an Internet Survey. *Social Science Research*, 36: 131-148.
- Couper, M.P., Singer, E. and Kulka, R.A. 1998. Participation in the 1990 Decennial Census. *American Political Quarterly*, 26(1): 59-80.
- Dillman, D.A. 2000. *Mail and Internet Surveys. The Tailored Design Method*. Second Edition. John Wiley & Sons.
- Farrell, D. and Petersen, J.C. 2010. The Growth of Internet Research Methods and the Reluctant Sociologist. *Sociological Inquiry*, 80(1): 114-125.
- Frippiat, D. and Marquis, N. 2010. Web Surveys in the Social Sciences: An Overview. *Population-E*, 65(2): 285-312.
- Groves, R.M., Cialdini, R.B. and Couper, M.P. 1992. Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly*, 56: 475-495.
- Groves, R.M., Presser, S. and Dipko, S. 2004. The Role of Topic Interest in Survey Participation Decisions. *Public Opinion Quarterly*, 68(1): 2-31.
- Groves, R.M., Singer, E. and Corning, A. 2000. Leverage-Saliency Theory of Survey Participation. *Public Opinion Quarterly*, 64: 299-308.
- Huh, M. and Cho, S.K. 2010. Propensity Adjustment Weighting of the Internet Survey by Volunteer Panel. *Survey Research*, 11(2): 1-28 (In Korean).
- Kim, W.Y. and Lee, H.C. 2003. Testing the Validity of Web Survey Using Propensity Scores in 2002 Korean Presidential Election. *Broadcasting Research*, Summer: 143-166 (In Korean).
- KISA(Korea Internet & Security Agency). 2012. *2011 Survey Report on Internet Use*.
- Lee, K. and Jang, D. 2009. Applying Propensity Score Adjustment on Election Web Surveys. *Survey Research*, 10(3): 21-36 (In Korean).
- Lee, S. and Valliant, R. 2009. Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment. *Social Methods & Research*, 37: 319-343.
- Schonlau, M., van Soest, A., Kapteyn, A. and Couper, M. 2009. Selection Bias in Web Surveys and the Use of Propensity Scores. *Social Methods & Research*, 37: 291-318.
- Shapiro, R.Y. 2011. Public Opinion and American Democracy. *Public Opinion Quarterly*, 75(5): 982-1017.
- Taylor, H. 2000. Does Internet Research Work? *International Journal of Market Research*, 42(1): 51-63.
- Valiant, R. and Dever, J.A. 2011. Estimating Propensity Adjustment for Volunteer Web Surveys. *Social Methods & Research*, 40(1): 105-137.

Table 1. Descriptive Statistics of Major Categorical Variables
(N=37,049)

Variables		frequency	%
Sex	Male	17,662	47.67
	Female	19,387	52.33
Educational attainment	Elementary school or less	6,939	18.73
	Middle School	5,641	15.23
	High School	13,954	37.66
	College or more	10,515	28.38
Job status	Hired currently	20,544	55.45
	Else	16,505	44.55
Within Else	<i>Student</i>	4,122	11.13
	<i>Retired</i>	3,586	9.68
	<i>Homemaker</i>	6,379	17.22
	<i>Jobless</i>	2,418	6.53
Marital status	Married currently	23,894	64.49
	Else	13,155	35.51
Within Else	<i>Never married</i>	8,809	23.78
	<i>Widowed</i>	3,114	8.41
	<i>Divorced</i>	1,232	3.33
Internet newspaper reading frequency	Almost every day	8,048	21.72
	3 or 4 times a week	4,318	11.65
	1 or 2 times a week	3,878	10.47
	Once every 2 weeks	1,641	4.43
	Never	19,164	51.73
Non-political donation experience of the past year	Yes	12,262	33.10
	No	24,787	66.90
Group participation experience of the past year	Yes	15,081	40.71
	No	21,968	59.29
Volunteering experience of the past year	Yes	7,392	19.95
	No	29,657	80.05
Intent to participate in social surveys via the Internet	Yes	5,759	15.54
	No	31,290	84.46

Table 2. Descriptive Statistics of Major Ratio Variables (N=37,049)

Variables	Total	Intent to participate in social surveys via the Internet	
		No	Yes
Age	<i>mean</i>	44.64	<u>35.17</u>
	<i>S.D.</i>	17.34	13.51
	<i>min.</i>	15	15
	<i>max.</i>	109	87
Years of education	<i>mean</i>	11.33	<u>13.51</u>
	<i>S.D.</i>	4.23	2.86
	<i>min.</i>	0	0
	<i>max.</i>	23	23
Non-political donation frequency	<i>mean</i>	1.77	<u>3.06</u>
	<i>S.D.</i>	5.42	8.73
	<i>min.</i>	0	0
	<i>max.</i>	320	320
Group participation degree [*]	<i>mean</i>	.65	<u>.90</u>
	<i>S.D.</i>	.89	1.02
	<i>min.</i>	0	0
	<i>max.</i>	3	3
Hours of volunteering activity	<i>mean</i>	.21	<u>.33</u>
	<i>S.D.</i>	4.38	5.04
	<i>min.</i>	0	0
	<i>max.</i>	440	206

*: In the questionnaire, the following categories of social groups were given: 1) socializing organization, 2) religious organization, 3) hobby, sports, or leisure organization, 4) NGO, 5) academic organization, 6) interest group, 7) political organization, and 8) etc. This variable refers to the number of categories that respondents chose on a 0 to 3 scale as the question limited the number of categories up to 3.

Table 3. Crosstabulation of Major Variables and Intent to Participate in Social Surveys via the Internet (%)*

Variables		No	Yes	Total
Sex	Male	82.43	<u>17.57</u>	100.00
	Female	86.31	13.69	100.00
Educational attainment	Elementary school or less	96.89	3.11	100.00
	Middle School	88.23	11.77	100.00
	High School	84.14	15.86	100.00
	College or more	74.65	<u>25.35</u>	100.00
Job status	Hired currently	81.83	<u>18.17</u>	100.00
	Else	87.72	12.28	100.00
Within Else	<i>Student</i>	75.59	<u>24.41</u>	100.00
	<i>Retired</i>	97.38	2.62	100.00
	<i>Homemaker</i>	90.50	9.50	100.00
	<i>Jobless</i>	86.72	13.28	100.00
Marital status	Married currently	85.76	14.24	100.00
	Else	82.09	<u>17.91</u>	100.00
Within Else	<i>Never married</i>	75.85	<u>24.15</u>	100.00
	<i>Widowed</i>	97.08	2.92	100.00
	<i>Divorced</i>	88.80	11.20	100.00
Internet newspaper reading frequency	Almost every day	70.66	<u>29.34</u>	100.00
	3 or 4 times a week	75.24	24.76	100.00
	1 or 2 times a week	80.04	19.96	100.00
	Once every 2 weeks	81.84	18.16	100.00
	Never	93.44	6.56	100.00
Non-political donation experience of the past year	Yes	78.80	<u>21.20</u>	100.00
	No	87.25	12.75	100.00
Group participation experience of the past year	Yes	80.63	<u>19.37</u>	100.00
	No	87.08	12.92	100.00
Volunteering experience of the past year	Yes	76.03	<u>23.97</u>	100.00
	No	86.56	13.44	100.00

* All crosstabulations above show statistically significant χ^2 results.

Table 4. Logistic Regression Analysis of Intent to Participate in Social Surveys via the Internet (N=37,049, Pseudo R²=.1313)

Predictor	β	SE β	Wald's χ^2	df	p	e ^{β}
Constant	-2.403	.099	-24.21	1	.000	N/A
Sex (female=1; male=0)	-.062	.031	-1.96	1	.050	.940
Age	-.038	.002	-23.62	1	.000	.962
Years of education	.083	.006	14.05	1	.000	1.087
Hired (hired=1; else=0)	.408	.036	11.30	1	.000	1.504
Married (married=1; else=0)	.199	.041	4.88	1	.000	1.220
Internet newspaper reading	.196	.011	18.16	1	.000	1.216
Non-political donation experience of the past year (yes=1; no=0)	.280	.035	8.01	1	.000	1.323
Group participation experience of the past year (yes=1; no=0)	.198	.034	5.83	1	.000	1.219
Volunteering experience of the past year (yes=1; no=0)	.336	.037	9.05	1	.000	1.400
Test			χ^2	df	p	
Overall model evaluation						
Likelihood ratio			4203.03	9	.0000	
Score test			6119.88	9	.0000	
Wald test			2850.37	9	.0000	
Goodness-of-fit test						
Hosmer & Lemeshow			26.92	8	.0001	