

‘Poor’ Browsers and Internet Surveys

Gaurav Sood
Stanford University

July 31, 2011

Given,

- older browsers are likelier to display the survey incorrectly.
- type of browser is a reasonable proxy for respondent’s proficiency in using computers (and possibly speed of the Internet connection).

Respondents using older browsers may abandon surveys at higher rates than those using more modern browsers.

Using data from a large Internet survey, we test whether people who use older browsers abandon surveys at higher rates, and whether surveys taken using old browsers contain greater missing data.

We control for some known correlates of extent of missing data, and propensity to abandon surveys - race, age, and education - in some of the analyses. It is possible that these known correlations are artifactual, a result of confounding variables like speed of Internet connection, etc.

DATA

Respondents come from 1.4 million members of Internet panel. Panel members were recruited in several ways. Initially, RDD phone calls were made to invite American adults to sign up to receive email invitations to participate in surveys, yielding about 2,500 panel members. Additional phone calls were made to professionals working in the information technology sector who were on professional lists; these calls yielded about 2,500 more panel members. These initial 5,000 panel members were offered a chance to win cash or gift certificates if they would refer friends or family who might sign up to complete online surveys. Referred panel members were offered the same incentives to refer other people. Panel members received a chance to win a prize each time they completed a survey, when someone they referred completed a survey, and each time the referral’s referral completed a survey. Panel members were also recruited through online ads (on the firm’s own website, news sites, blogs or search engines) and through emails from businesses or non-profit organizations with which the panelist had an affiliation. 72,000 members of a stratified random sample of panel received an email inviting them to participate in our survey, and 2,632 did so (4% response rate with a 2% completion rate). Probability of selection within demographic strata was specified so that invitees resembled the distribution of gender, age, and Census region in the U.S. adult population, as estimated by the 2000 Census. No quotas were used to restrict who completed the survey. Respondents were entered into a sweepstakes in return for completing our survey. Email invitations were sent on December 15th, 21st, 2009, and January 10th, 2010. The resulting sample

is highly representative with average error on observable sociodemographic characteristics less than 2%.

MEASURES

Dependent Variables

- Whether the person completed the survey or not
- Number of missing fields in the survey

Independent Variables

- Browser - dichotomized into old and new. Old - Internet Explorer 7 or lower, Firefox 3.0 or lower.
- Age in years
- Education - years of education
- Race - 4 categories: White NH, Black NH, Hispanics, Other

RESULTS

- People using older browsers are less likely to complete the survey ($p < .05$ for both datasets).
- There is some evidence that respondents using older browsers have more missing data ($p < .1$ for both datasets).

Table 1: Predicting Whether Respondent Completed the Survey

	Substantive Survey	Demographic Survey
(Intercept)	4.76*** (1.17)	2.26*** (0.10)
I(browx == "old")TRUE	-0.75* (0.38)	
ppage	-0.02 (0.01)	
ppeduc	0.06 (0.09)	
as.factor(race2)2 black nh	15.50 (714.58)	
as.factor(race2)3 hispanic	-0.01 (0.62)	
as.factor(race2)other	-0.40 (0.75)	
I(browy == "old")TRUE		-0.35** (0.13)
<i>N</i>	2013	2327
AIC	311.79	1618.14
BIC	468.80	1664.16
log <i>L</i>	-127.90	-801.07

Standard errors in parentheses

† significant at $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

Table 2: Among Completed, Predicting Number of Missing

	Substantive Survey	Demographic Survey
(Intercept)	240.93*** (0.80)	244.48*** (0.22)
I(browx == "old")TRUE	0.46† (0.26)	
ppage	0.02** (0.01)	
ppeduc	0.28*** (0.06)	
as.factor(race2)2 black nh	-0.10 (0.43)	
as.factor(race2)3 hispanic	-7.03*** (0.46)	
as.factor(race2)other	0.52 (0.62)	
I(browy == "old")TRUE		0.79* (0.32)
N	1778	1764
AIC	11090.07	11743.55
BIC	11243.60	11787.35
$\log L$	-5517.04	-5863.77

Standard errors in parentheses

† significant at $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

DISCUSSION

The poor, the old, the less educated, and racial minorities abandon Internet surveys at higher rates. It is not clear why. In fact with incentive based surveys, one expects the pattern among the poor to be roughly opposite of what we see. One possible explanation for the counter-intuitive empirical observation is that respondents from these demographics have slower Internet connections, use older browsers (which pose other problems), and are less proficient in using computers.

Inference is hampered by limitations of the measures, and the sample -

- Type of browser is a proxy for user’s proficiency in using the Internet, bandwidth of her Internet connection etc. Direct measures of variables of interest - speed on Internet connection, data on how survey is displayed, etc. - would be vastly superior.
- Since the data are from a self-selected sample recruited online, there is every chance that it over-represents those who are better at using computers and the Internet. This bias in sample likely means that effects reported here are not representative of population at large.