I MPLEMENTATION OF NATIONALLY REPRESENTATIVE WEB-BASED SURVEYS

Vicki J. Huggins and Karol Krotki
Knowledge Networks and NuStats

KEY WORDS: Internet Survey, RDD Sampling, Research Panel, MSN®TV

I. Introduction

We start with the goal of selecting survey research samples that can statistically support inferences about the total population of households in the United States, and/or subsets within the population. Nationally representative Internet based sampling frames do not currently exist because every household in the U.S. cannot be accessed via the Internet. This results in serious undercoverage that can significantly affect outcome study variables. Statistics from the Current Population Survey (CPS) for August 2000 indicate that 51% of the U.S. households have computers and 41.5% of households have access to the Internet. [Newburger, 2001] Any serious attempt to do national household surveys needs to take into account this undercoverage.

The Knowledge Networks solution is to utilize standard Random Digit Dial (RDD) sampling [Lepkowski, 1988] to obtain a representative sample of U.S. telephone households and then equip those households with a MSN®TV unit for survey administration.

There is an established and ongoing methods research program for evaluating and improving the survey methods and quality of studies conducted using the Knowledge Networks panel. This research has been conducted by Knowledge Networks, the Research Triangle Institute and supported by academic researchers.[1]

Following a more detailed description of the sample design for the Knowledge Networks Panel in section II and a summary of cooperation rates in section III, survey methods research results to date are summarized and presented in section IV. Section V closes with a summary of plans for future research.

[1] Knowledge Networks and RTI are in an Alliance to competitively bid Federal research projects that involve a web-enabled panel. Knowledge Networks works independently with academic researchers on their own investigator-initiated research projects. For more information on the Alliance, go to the following Internet link: www.knowledgenetworks.com/ganp

II. Sample Design for a Web-based National Probability Sample Panel:

The sample design begins with an Random Digit Dialing (RDD) sample of households, followed by a reverse address match, and mailing of an introductory letter to every household for which we are able to obtain an address match. Households (both addressed matched and non-addressed matched) are then recruited by telephone. Once a household agrees to participate, KN delivers a MSN®TV unit that essentially transforms the television in the household into a monitor for survey administration. All household members are recruited and all adults (18 and over) are given a welcome survey to familiarize them with use of the MSN®TV. Then a profile questionnaire is assigned to each household to collect basic demographic information about the household and its members. Once we have received the profile data, the household is considered ready to receive regular surveys.

To ensure appropriate representation, panel post-stratification weights are updated after each sample selection such that the weighted panel distributions match benchmarks as determined from the most recent monthly CPS. We use a weighting cell approach where the cells are defined by cross-classifying the following variables: age, gender, region, race, ethnicity, and education.

Samples are drawn consecutively throughout the week with probabilities proportional to the panel weights using systematic sampling applied to the sorted panel members.

III. Cooperation/Response Rates

Ensuring high response and cooperation rates is one of the most challenging aspects of survey administration. Clearly in the industry, response rates through the telephone mode of data collection are more and more difficult to maintain. The overall historical cooperation rate at the recruitment stage is approximately 50%. There remain at least three more stages before the member becomes fully profiled, active, and ready for weekly surveys. And at each stage there is some attrition. Thus, the final overall cumulative response rate ranges from 25% to 50% depending on the level of efforts expended for individual projects. We consider this to be one of our

major challenges and are carrying out extensive research to maximize the cooperation rate at each stage. The current and cumulative response rates for fielding an Internet survey from the Knowledge Networks panel are found in Table 1.

IV. Panel Quality

Below, we discuss several areas where analyses have been initiated to investigate sampling and nonsampling error in the conduct of surveys from the Knowledge Networks Panel. We will present summary results in the following areas:

- *Coverage Error*
- *Benchmarking Analyses*
- *Existence of Panel Bias*
- *Nonresponse Bias*

A. Coverage Error

There are two key sources of coverage error that can affect the representative nature of the Knowledge Networks panel sample: Error arising from noncoverage of nontelephone households and error arising from noncoverage of non-MSN®TV areas. We discuss the magnitude of each of these and our approaches to reduce biases stemming from them.

Noncoverage of Nontelephone Households

According to the June 2001 CPS, approximately 5% of households in the U.S. are without a phone at the time of interview. Phone coverage differs by household income (80% for households with income less than $5,000 and 92% for households with income $15-20K), state, metro status, race, ethnicity, etc. Currently, a post-stratification weighting adjustment is made to the Knowledge Networks panel to ensure total population estimates from the RDD based sample are consistent with U.S. population estimates for the phone and non-phone population. The adjustment is made at the state level, and then further refined through post-stratification (raking) using gender, age, race/ethnicity and education level. The complete post-stratification scheme is implemented for two purposes: (1) Reduce the bias in the panel due to coverage and nonresponse error and (2) Reduce the variance for statistics highly correlated with the demographic benchmarks.

We are investigating whether a separate weighting adjustment specifically to account for nontelephone coverage error would be more accurate for reducing potential bias at lower levels in the sample. Specifically, we are investigating the methodology proposed by Frankel (2000) for reducing nontelephone bias in RDD surveys that uses survey data collected on interruption in telephone service to identify respondents more like non-telephone households for weighting purposes.

Table 2 presents comparative estimates of household characteristics of panel members who were asked whether they had an interruption in telephone service for 1 week or more in the past year. It is quite clear from the table that the group with interruption in telephone service and the group without are different. Estimates of the number of children under 18, household size, household income < 25K, and household type are all statistically different between the groups with and without telephone interruption. Approximately 3.6% of recruited households reported being without telephone service for 1 week or longer in the previous 12 months.

The next steps are to look at mean square error for selected estimates if the weighting approach is administered at the state/msa level and perform sensitivity analysis on a range for the bias. We will also investigate whether other variables such as household tenure, having access to a computer at home and/or access to the Internet or household type might well be good predictors as well.

Noncoverage of Non MSN®TV Service Areas

Initially, the Knowledge Networks panel suffered from noncoverage of households due to the fact that the Internet Service Provider – MSN®TV – does not cover all areas of the U.S. Currently though, a subsample of households in these non-covered areas are recruited using different Internet providers. Estimates computed from the panel show that non MSN®TV covered areas are more rural, concentrated more in the Midwest, less educated, more likely to be white, non-Hispanic and with a lower income distribution. Inclusion of a subsample of households from non-MSN®TV covered areas will mitigate any panel bias associated with the original noncoverage of non- MSN®TV areas.

B. Benchmarking Analyses

One method for analyzing the quality and representativeness of a study or sample is to compare a variety of estimates from the study or sample to known and/or official benchmark estimates. This section presents results of comparisons of data from the Knowledge Networks panel to several other sources including the Current Population Survey, the Behavior Risk Factor Surveillance Survey (BRFSS) 2000, and the Ohio State University RDD Survey on Public Opinion and Voting Intentions for the 2000 U.S. Presidential Elections.

Table 3 presents a comparison between the KN panel and the Current Population Survey (CPS) for selected demographics as of June 2001. Column 1 contains estimates for the panel using the entire recruited panel sample with the associated weight from the initial selection probabilities. Column 2 presents June 2001 CPS estimates.

As you can see from table 3, column 1, the Knowledge Networks panel under represents the elderly, is skewed towards the upper end of the socioeconomic scale, and under represents the African American minority. The differences in the race estimates is an primarily a difference in the way Census asks race as compared to Knowledge Networks, with Knowledge Networks offering "Other" as a race category. The panel also slightly under-represents the Hispanic population. Due to the large sample sizes associated with both the Knowledge Networks panel and the Current Population Survey, small differences are statistically detectable as asterisked. In general, the average deviations are not huge, and sample representativeness is never dramatically poor.

Benchmarking of results from several surveys conducted using the Knowledge Networks panel has also been conducted. Table 4 presents comparative results from a Study on Smoking to comparable estimates from the BRFSS. [Dennis, 2001] According to the Knowledge Networks sample, 26% of Veterans between the ages of 22 and 80 currently smoke. The BRFSS survey of year 2000 shows that 24% of Veterans between the ages of 22 and 80 currently smoke.

In an independent study conducted by Jon Krosnick and LinChiat Chang at Ohio State University [Krosnick, 2001], Knowledge Networks survey results were compared to results from both a random digit dial study conducted by the Ohio State University Center for Survey Research and the Harris Interactive Internet opt-in panel. The same questionnaire to gauge public opinion and voting intentions for the 2000 U.S. Presidential Election was administered under each of the survey modes and standard data collection methods.

Krosnick and Chang concluded that Internet based data collection represents a viable approach to conducting Random Digit Dialing surveys. And the Knowledge Networks methodology resulted in a more representative sample than the opt-in panel sample utilized by Harris Interactive. Results also suggest that Internet data collection improves the accuracy of the reports respondents provide over accuracy obtained through telephone interviews.

C. Preliminary Research on the Existence of Panel Bias

Research panels may be susceptible to two types of panel effects. The first type is the possibility of conditioning research subjects in a panel sample, turning them into "professional respondents" whose attitudes and behaviors are changed by panel participation. The second type of effect that panels are potentially vulnerable to is selection bias, which can make successive samples less representative. Preliminary research, using data from a variety of different studies, has not detected serious levels of panel effects. The discussion below presents selected results that illustrate these findings. More detail can be found in Dennis (2001).

Sensitive questions

Panel members with more tenure might be expected to be more comfortable with the survey environment and be less affected by the impulse to give socially desirable answers. Although the surveys are taken in a self-administered setting, some newer panelists might feel an urge to be more positive and conforming. However, the data from a survey of approximately 6,000 panelists provides limited support for this hypothesis.

When asked about their comfort level with a shop owner with AIDS, newer panel members were more likely to provide the socially pleasing response of comfortable – see table 5. Most of these small-scale effects evident in the other questions disappeared or were diminished when controlling for panelists' demographic characteristics within each tenure group. Overall, the effects are small and are almost certainly less serious than the social desirability effects well documented in telephone and face-to-face interviewing.

D. Nonresponse Bias

As described in section III above, nonresponse or cooperation bias can creep in at several different stages, from RDD recruitment, MSN®TV installation, profiling of members, and completion of project-specific surveys. Different levels and detail for data are available on nonrespondents at the different stages.

Currently, a weighting adjustment to reduce nonresponse bias from panel recruitment through profiling implemented with the use of post-stratification to CPS population totals prior to sample selection of weekly surveys. Then, after a survey is

fielded, a separate nonresponse adjustment to reduce nonresponse bias for individual surveys is applied.

We have been able to compare selected characteristics of responders and nonresponders at the point where recruited households are asked to complete the household profile questionnaire. The recruiting interview collects information about household decision maker, use of a computer, access to the Internet, and household composition. Our analyses show statistical differences between responders and nonresponders about whether a computer exists in the home and whether it is connected to the Internet (87.0% and 78.9% respectively for nonresponders and responders). Also, there is a slight skewness for households with a smaller number of members completing the household core profile versus not completing it. These variables can be considered for use in a nonresponse adjustment for profiled households to better adjust for non-profiled households.

The Research Triangle Institute sponsored a formal study of the effects of nonresponse on key outcome variables in the recent Survey on Health and Aging. [Wiebe, 2001] The methodology included re-sampling of nonrespondents, fielding the core survey to the nonrespondents, and weighting the nonrespondent completes using the resampling design. Implemented in 2000, telephone interviews were conducted with samples from the following nonresponse groups:

- RDD Panel Recruitment Refusers (n=71 completes)
- RDD Acceptors: Agreed to participate in the Web-enabled panel but had not yet hooked up MSN®TV (n=129 completes)
- Telephone prompting encouraged Panel Nonrespondents to complete the survey on the Internet Appliance (n=238 completes)

Data collection from the resample of nonrespondents was conducted using both telephone and web assisted methods. Where possible, nonrespondents were contacted by phone and asked to complete the Survey of Health and Aging (SHA) on the web device. If this was not possible, they were asked to complete the survey over the phone.

The weighted response rate for the study increased from 25% to 43% as a result of nonresponse follow-up by phone. Different participation groups appear to report different answers in the survey with no clear pattern in the responses.

The primary question that motivated the study was whether the follow-up would change the key study estimates. The conclusion made by the researchers was that the nonresponse follow-up did not make any significant changes in the overall representativeness of the sample. The representativeness of the sample was actually achieved through the standard procedures used by Knowledge Networks to select the sample from the full panel. Inclusion of additional recruitment groups did not affect the estimates. When all components of the nonresponse followup are included with the initial MSN®TV based estimate, no significant changes in the outcome estimates resulted.

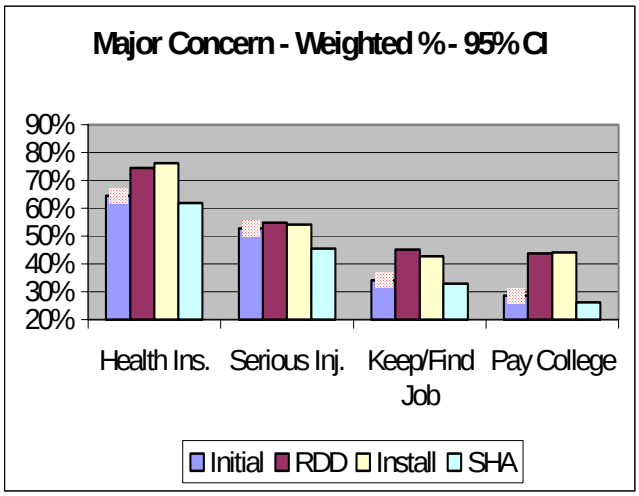
Table 6 presents the estimates for the question on coping with serious injury split out by response and nonresponse stages as well as estimates using data combining the nonresponse sample results with the original Survey of Health and Aging Results. We can see by examining the cumulative results that the additional weighted responses from NRFUS had little impact on the overall prevalence estimates.

Respondents were asked about how concerned they were with having adequate health insurance, coping with serious injury or illness, keeping a job, job hunting, or changing careers, and paying for children's college education. Figure 1 shows the responses provided on these questions for the four types of respondents:

- Those who completed during the initial study
- Those who refused the RDD recruitment but completed the nonresponse follow-up survey (NRFUS)
- Those who failed to install the MSN®TV device but completed the NRFUS
- Those who refused the initial survey but completed the NRFUS.

The results indicate that the persons who refused the RDD recruitment and those who failed to install the device provided significantly different responses on topical questions than did those who cooperated with the initial survey. This suggests that the respondents and nonrespondents may be different and should be carefully evaluated in future studies.

Figure 1: Effect of Nonresponse on Substantive Estimates?



V. Future Research

The methodological issues presented in this paper will continue to be investigated. These include teasing out panel effects, mode effects, nonresponse and noncoverage bias, and response bias. We also shall address instrument design issues on the Internet raised by Mick Couper [Couper, 2000] and others. More topical benchmarking is needed as well. Knowledge Networks and the Research Triangle Institute will jointly conduct basic research on the panel, experimenting with the use of incentives, assessing panel bias as the panel ages, and expanding nonresponse studies. As new surveys and new research related to this new survey mode for large scale panel data collection continues, we will continue to clarify the problems and pose potential solutions.

References

Dennis, Michael J., "A Study of Panel Effects", May 2001.

Dennis, Michael J. "Knowledge Networks Profile Data Compared to BRFSS 2000: Smoking Prevalence and Demographic Characteristics of U.S. Veterans", Knowledge Networks Report, May 22, 2001.

Clinton, Joshua, "InterSurvey Panel Effects Study – June 2000", Knowledge Networks Research Note, June 2000.

Couper, Mick P., "Web Surveys: a Review of Issues and Approaches", Public Opinion Quarterly, 64:464-494, 2000.

Elizabeth F. Wiebe, Joe Eyerman, and John Loft, "Evaluating Nonresponse in a Web-Enabled Survey on Health and Aging", Presented at the 2001 Meeting of the American Association for Public Opinion Research, Montreal, Quebec May 17 - 20, 2001.

Frankel, Martin R. and Michael P. Battaglia, David C. Hoaglin, Robert A. Wright, and Philip J. Smith, "Reducing Nontelephone Bias in RDD Surveys", Presented at the Meeting of the American Association for Public Opinion Research, 2000.

Krosnick, Jon A. and LinChiat Chang, "A Comparison of the Random Digit Dialing Telephone Survey Methodology with Internet Survey Methodology as Implemented by Knowledge networks and Harris Interactive", Presented at the 2001 Meeting of the American Association for Public Opinion Research, Montreal, Quebec May 17-20, 2001.

Lepkowski, J. "Telephone Sampling Methods in the United States", Chapter 5 in Telephone Survey Methodology, Wiley and Sons, 1988.

Newburger, Eric. Home Computers and Internet Use in the United States: August 2000, Current Population Reports, U.S. Census Bureau, P-23-207, September, 2001.

Table 1: Knowledge Networks Cooperation/ Response Rates

Component of Overall Response	Rate	Cumulative Response Rate
Panel Recruitment Cooperation	56%	56%
MSN®TV Installation	80%	45%
First-survey Profile Completion	88%	39%
Internet Survey Response*	85%	34%

* Varies according to design choices between 75% and 90%.

Table 2. Characteristics by Interruption in Telephone Service

Characteristic	Interruption: Yes	Interruption: No
# of Children <18	.79*	.61
HHType – Single, detached	45%*	67%
Tenure - Owner	45%*	73%
Income < 25K	37%*	16%

* Indicates statistical significance for p<.05 (2-sided)

Table 3. Knowledge Networks Panel and Current Population Survey (CPS) Demographics: Sept. 2001

U.S. Characteristic	KN Panel	CPS
Male	49.4%	47.9%
18-24	12.5%*	13.2%
25-54	42.7%*	37.0%

55-64	10.3%*	11.8%
65 or over	8.6%*	16.1%
White	79.3%*	83.2%
Black	10.5%*	11.9%
Other	5.2%	n/a
Hispanic	6.4%*	10.7%
Married	61.6%*	57.5%
Less than HS/ HS Diploma	42.8%*	48.8%
$10,000-$24,999	12.5%*	18.4%
$75,000 or more	24.8%	24.6%
Northeast	18.3%*	19.1%
West	22.5%	22.4%

* Indicates statistical significance for p<.05 (2-sided)

Table 4. Current Smoking Prevalence Rates KN and BRFSS 2000: Males Age 22-80 Years

Smoking Status	Veteran Status			
	Yes		No	
	KN	BRFSS	KN	BRFSS
Currently Smoke	**26%**	**24%**	**28%**	**24%**
No, Do Not Smoke	**74%**	**76%**	**72%**	**76%**
Total	**100%**	**100%**	**100 %**	**100%**

*p-value < .05 (two-sided)

Table 5. Attitudes on sensitive questions

	Recruitment Group				
	Web (stand-ard)	Web with Phone Prompt	Phone - Refused Web Panel Recruit-ment	Phone - Failed to Install Web Device	Total
Not at all con-cerned	4.1%	7.1%	4.2%	5.3%	4.2%
2.00	6.0%	8.4%	2.8%	6.1%	6.1%
3.00	16.9%	14.6%	15.5%	16.0%	16.8%
4.00	19.7%	22.8%	19.7%	18.4%	19.8%
Major concern	52.8%	45.5%	54.9%	54.1%	52.5%
Did Not Answer	.5%	1.6%	2.8%		.6%
Total	100%	100%	100%	100%	100%

Table 6. How concerned are you with coping with serious injury or illness? Recruitment Group with weighted percentages

Question	Panel Tenure (months)			
	<3	4-6	7-9	10-12
People with AIDS deserve it (% agree)	15	18	19	21
How likely to get AIDS from sharing same drink glass (% likely)?	25	25	23	24
How likely to get AIDS from someone coughing or sneezing (% likely)?	22	23	20	22
Is there currently a cure for AIDS (% yes)?	19	19	18	18

*p-value < .05 (two-sided)