



Effects of Panel Attrition on Survey Estimates

Authors:
J. Michael Dennis
Rick Li
Knowledge Networks, Inc.

**For Presentation at the 2003 Annual Meeting of the American
Association for Public Opinion Research in Nashville, TN**

Session on 'Response Rates,' Saturday May 17, 2003 at 8:15am-9:45am, Salon D

Contact Author's Information:
J. M. Dennis
Knowledge Networks, Inc
1350 Willow Road, Suite 102
Menlo Park, CA 94025
mdennis@knowledgenetworks.com
(650) 289-2160 (office phone)
(650) 289-2001 (fax)

Contents

Introduction.....	3
Data and Methods	4
Table 1: Data Source Description: Background Surveys	5
Results	6
Table 2: KN Panel Characteristics, Active versus Lost Panelists, Compared to Census Benchmarks.....	7
Table 3: Frequency Distribution of Background Variable by Absolute Percentage Point Difference between Estimates by Absolute Differences.....	8
Table 4: Background Questions by Panel Status (Active vs. Lost), Unadjusted and MCA Adjusted, for 20 Questions with the Largest Absolute Differences between Estimates based on Active versus Lost Panelists.....	10
Table 5: Political/Public Affairs Background Variables by Panel Status (Active vs. Lost), Unadjusted and MCA Adjusted.....	12
Conclusions	13
Appendix A: Knowledge Networks' Panel Recruitment Methodology	14
Appendix B: Technical Note on Weighting.....	16
References	17

Introduction

One advantage research panels have for longitudinal research is that the costs of the initial recruitment can be spread out over multiple waves of data collection. Another is that the statistical power for detecting differences in experimental designs and for improving predictive reliability is enhanced by repeated same-subject measurements. But these advantages are offset to the extent that panels are affected by the loss of panelists prior to second and later followup surveys. A research panel that contains a mix of recruits from fresh cross-sectional samples and more tenured research subjects provides some insurance against nonresponse bias resulting from panel losses. However, the risk remains even then that lost panel members may have self-reported attitudes, knowledge, and behaviors that are different than those of the panelists that actually participate to the wave 2 and later surveys, and this possibility can call into question the validity of these data.

This paper examines the effects of panel attrition on substantive survey estimates regarding political attitudes and behavior, general values, health status, lifestyle, and other areas. As such, it is an attempt to go beyond the search for demographic correlates of attrition to answering the question of the extent and ways in which panel attrition affects the validity of substantive survey results.

This research comes under the heading of outcome research on unit-level survey nonresponse because much of the total nonresponse in panel surveys is attributable to panel attrition. Please note that the research question does not address the determinants of unit nonresponse, which has been addressed separately. We shall touch upon the subject as a byproduct.¹

The distinct context for this research is the Knowledge Networks (KN) web-enabled panel, a multipurpose online panel based on probability sampling of U.S. telephone households.² This panel differs from government-sponsored panels such as SIPP, NES, and PSID in important ways that should be considered if attempting to generalize the results of this paper to these established panel studies or other studies. The KN panel members, compared to panelists for these three government-sponsored projects, tend to participate in surveys of shorter duration, in surveys covering a wider variety of topics, and tend to participate more frequently (approximately three times a month) in self-administered surveys (as opposed to face-to-face surveys). All of these differences can have an impact on the rate of panel attrition as well as the impact of attrition on the representativeness of the panel.

Author Affiliations: J. Michael Dennis is Vice President and Managing Director, Government and Academic Research, Knowledge Networks; Rick Li is Senior Research Analyst, Government and Academic Research, Knowledge Networks.

Acknowledgements: The authors wish to thank Mick Couper for comments on an early draft of this paper, and for the insights of our colleagues Daniel Slotwiner, Bill McCready, and Vicki Huggins.

¹ See Dennis (2001) for analysis of predictors of unit-level nonresponse to online panel surveys. Paper available upon request. See Lepkowski and Couper (2002) for empirical evidence and a theoretical framework for a study of the determinants of nonresponse in followup study waves.

² For a fuller description of the KN panel recruitment methodology, please see Appendix A.

Data and Methods

The survey data are from an internal program of background surveys conducted by KN to facilitate sampling and analysis. All large-sample background surveys were selected for analysis, and within these surveys, all survey variables, with few exceptions, were included in the analysis.³ The background surveys took place between 2000 and midyear 2002.

For this study, the completed background interviews were sorted into two categories: those completed by panel members that were still active on the panel as of June 2002, and those no longer on the panel as of June 2002. The former are designated ‘active’ panelists and the latter are ‘lost’ panelists. Of course, all panelists were part of the KN panel at the time of completing the survey, and there was no way to know at the time of survey participation which panelists would ultimately leave or be asked to leave the panel.

The reason that it is useful to sort panel interviews into active and lost categories is that it provides a view into what the survey results would look like for a followup survey among persons that leave a panel versus those that stay on the panel (including baseline plus fresh replacement recruits). ***In other words, nonresponse bias is not the result of panel attrition if lost panelists and active panelists share similar responses to the same questions.*** Secondly, if there are differences, it is helpful to know if they can be addressed through statistical adjustments using variables that are responsible for the differences, such as demographic, Internet and computer usage and experience, and geographical variables.

During the study period, the active panel sample was continuously supplemented with injections of fresh cross-sectional samples that in effect replaced lost panelists. Therefore, the group of active panelists is made up of initial recruits plus subsequent recruitment cohorts defined by when they joined the panel. The gross differences detected between active and lost panelists do not take into account that these subsequent recruitment cohorts might be different from the earlier recruits in their attitudes and behaviors, due to, for instance, underlying changes in public opinion over time. We analyze the differences controlling for recruitment cohort, and we also analyze a specific cohort to address this layering within the web-enabled panel.

Table 1 lists the covered survey topics and the number of questions and interviews donated by each background survey. The ‘N Active’ column shows the number of interviews completed by ‘active’ panel members, while the ‘% Lost’ column shows the percentage of interviews from ‘lost’ panelists. The surveys with higher rates of loss (e.g., political/public Affairs) are those that were fielded relatively early to KN panelists, while those with lower rates of loss were first fielded more recently. In total, the data set

³ Some background variables were excluded from survey questions that asked about rare diseases, specific TV program viewing, and other highly specialized topics that we considered too arcane (even for an AAPOR audience).

covers 323 questions from 11 subject areas and includes 397,032 completed interviews. The distribution of interviews is 57.5% from active panelists and 42.5% from lost panelists, providing a large sample for the detection of differences. The average number of interviews donated per panelist is five. The sample sizes in all the background specific analyses were large; the minimum group size (N cases for active or lost panelists) was 4,725 cases and the mean was 20,751 for interviews from active panelists and 15,342 from lost panelists.

Table 1: Data Source Description: Background Surveys

	N Survey Questions	N Interviews	N Interviews with Active Panelists	% of Interviews with Lost Panelists
Profile Surveys				
Automobile	19	29,250	20,661	29.4%
Values/Psychographics	69	32,055	25,769	19.6%
Lifestyle	98	37,979	22,555	40.6%
Political/Public Affairs	19	44,552	18,333	58.9%
News Consumption	34	44,250	21,923	50.5%
Fast Food	8	30,788	23,237	24.5%
Magazine Reading	18	32,618	16,537	49.3%
Health	23	48,368	25,232	47.8%
Cable & TV	7	25,105	18,792	25.1%
Movies	2	19,992	11,381	43.1%
Computer/Internet	26	52,075	23,846	54.2%
Total	323	397,032	228,266	42.5%

Our analysis consisted of the following steps. First, we computed the absolute percentage point difference in survey results between the active and lost cases for each of 323 survey questions. To simplify the presentation of results, we selected a single response category for each question and in some cases combined categories (“agree” and “strongly agree”). We identified the twenty questions with the largest differences and grouped the political affairs/public affairs questions as they are likely to be of interest to the AAPOR audience, and averaged the absolute differences (both weighted and unweighted) for subgroups of items and the whole.

Our second step was to assess whether the effects of panel attrition on survey results are substantial after controlling for other factors such as age, race, and gender. The raw unadjusted survey results from active and lost panelists are juxtaposed to survey results after their adjustment by MCA using gender, education, household income, race/ethnicity, household size, and urban/rural as covariates.⁴ The purpose of employing MCA is to determine whether differences observed between active and lost panelists hold

⁴ For more on uses of the MCA statistical technique, see Robinson (2002).

up after controlling for common demographic variables.⁵ The differences might be reduced, for instance, because the lost panelists tend to be younger persons or those with a college degree. The MCA output shows the effect of panel status (active compared to lost) after removing the effects of covariates.⁶ The data are weighted by the inverse probabilities of selection resulting from the panel recruitment and to correct for nonresponse and noncoverage error resulting from panel recruitment.

Please see Appendix B for a technical note on the statistical weighting used for this paper.

Results

Demographic Characteristics of Active and Lost Panelists

Compared to active panelists, the lost panelists tend to be younger (age 19-29), while persons age 60 and over are less likely to drop off the panel, as shown in Table 2, and overshadow the much smaller drop-off observed in college educated persons or those residing in the West Region. This finding for younger persons has been confirmed in a University of California Irvine longitudinal study on the psychological effects of the 9/11 terrorists attacks that is using the web-enabled panel (Silver et al, 2002).

⁵ The SPSS code used for this analysis is below, where Y is the dependent variable and X is Panel Status as follows:

```
Anova Y BY X (0,1)
WITH gender education hhincome ethnicity region age hhsz msa
/COVARIATES=WITH
/MAXORDERS=NONE
/METHOD=EXPERIMENTAL
/STATISTICS=MCA
/MISSING=EXCLUDE.
```

(Note the categorical variables used as covariates have between 2 and 5 categories.)

⁶ The employed list of predictor variables is not exhaustive. Further analysis is to be conducted for identifying other predictors that could be useful for refining the nonresponse adjustment optimal for the web-enabled panel.

Table 2: KN Panel Characteristics, Active versus Lost Panelists, Compared to Census Benchmarks

Characteristics		U.S Census (Weighted)	Panel Distribution (Unweighted)	
			Active	Lost
Gender	Male	48.0%	48.7%	49.9%
	Female	52.0%	51.3%	50.1%
Age	18-29	21.7%	13.5%	25.1%
	30-44	31.1%	34.1%	38.5%
	45-59	25.8%	32.4%	26.8%
	60+	21.4%	20.0%	9.7%
Race/Ethnicity	White	72.7%	78.0%	73.4%
	Black	11.6%	9.9%	10.5%
	Other	4.7%	4.7%	8.3%
	Hispanic	10.9%	7.3%	7.8%
Education	Less than HS	16.7%	7.1%	7.8%
	HS	32.3%	25.6%	24.1%
	Some college	27.1%	39.7%	37.6%
	College +	24.0%	27.6%	30.5%
Region	Northeast	19.1%	15.7%	16.3%
	Midwest	22.8%	32.4%	28.6%
	South	35.6%	31.3%	32.3%
	West	22.6%	20.5%	22.8%

Data Sources: Census Current Population Survey and KN Panel as of June 2002

Unadjusted Differences in Survey Results

The median difference in answers between active and lost panel members is small, 1.6 percentage points weighted and 2.1 percentage points unweighted, when aggregated across the 323 background variables. Table 3 on the next page displays a count of the number of background variables according to the absolute difference in percentage points between the survey estimates for active versus lost panelists. Because of the skew introduced largely by eleven computer/Internet variables, the mean difference is higher than the median (2.3 versus a median of 1.6 for weighted variables). Simple poststratification weighting (see Appendix A for details) reduced the median difference by 24%, confirming that much of the variance can be reduced by weighting to common demographic variables and correcting for unequal selection probabilities.

Table 3: Frequency Distribution of 323 Background Variable by Absolute Percentage Point Difference between Estimates by Absolute Differences

% Point Difference	Weighted		Unweighted	
	N Variables	Col %	N Variables	Col %
0 to < 1	101	31%	82	25%
1 to < 2	89	28%	65	20%
2 to < 4	82	25%	102	32%
4 to <10	46	14%	68	21%
10 or more	5	2%	6	2%
Total	323	100%	323	100%
Mean Diff = 0.023		Mean Diff=0.028		
Median Diff = 0.016		Median Diff=0.021		

Lost panel members tend to report having more computer and Internet experience, usage, and access than active members; consequently, most of the variables with the largest differences are related to computing and the Internet. Eleven of the 20 questions exhibiting the largest differences in responses between active and lost panelists relate to computing, the Internet, or computer-related ‘gadgets.’ Two of the remaining 9 questions could be interpreted as relating to technology (i.e., renting and purchasing of DVDs). The median difference between the 20 ‘worst’ measures for the active and lost interviews is 7.2 percentage points weighted and 9.1 percentage points unweighted.

All of the ‘largest-20’ differences regarding computing and the Internet are in the direction of lost panelists having more access, use, and experience with computers and the Internet.

Despite this dynamic, surveys using the web-enabled panel produce estimates of computer and Internet access similar to the U.S. Census. The higher rate of drop off of Internet households from the web-enabled panel has offset the higher propensity of such households to join the panel in the first place, resulting in the web-enabled panel having an approximate representation of the population of Internet households. The point estimate for year 2000 for the web-enabled for U.S. households with Internet access, 47%, is slightly higher than the CPS point estimate of 42% for August 2000 (Census, 2001). The KN estimate does not include a poststratification on Internet/computer benchmarks.

A battery of 69 questions regarding values and ‘psychographics’ were asked of 32,055 respondents, focusing on valuations of the ‘American way’, tolerance for racial and sexual differences, immigration, economic equality, socialization, and the role of religion in society and politics. On these measures, the difference between active and lost panelists averaged 1.4 percentage points (weighted), providing evidence of a broad similarity in values between active and lost panelists.

Adjusted Differences in Survey Results

Table 4 on the next page shows the survey estimates for active and lost panelists for the twenty questions with the largest differences, but this time the raw unadjusted numbers are shown first, followed by the results after their adjustment by MCA using gender, education, household income, race/ethnicity, household size, and urban/rural as the covariates.

The MCA adjustment reduced the average difference from 10.7 to 5.6 percentage points for the 20 questions as a group. For three of the computer/Internet variables the difference remains above 10 percentage points after the MCA adjustment. If the narrowing of the difference resulting from the MCA adjustment for the 20 questions with the largest differences were to hold for all 323 background variables, the median difference would be reduced to 1 percentage point between active and lost interviews.

Table 4: Background Questions by Panel Status (Active vs. Lost), Unadjusted and MCA Adjusted, for 20 Questions with the Largest Absolute Differences between Estimates based on Active versus Lost Panelists

Questions	Unadjusted			MCA Adjusted		
	Active	Lost	Active - Lost	Active	Lost	Active - Lost
Internet access at home (yes)	0.40	0.57	-0.17	0.42	0.55	-0.13
Length of time using the Internet (more than a year)	0.46	0.62	-0.16	0.50	0.58	-0.08
Ever made credit card purchase over the Internet (yes)	0.41	0.55	-0.14	0.43	0.53	-0.10
Length of time using email (more than a year)	0.46	0.61	-0.15	0.50	0.58	-0.08
Computer access at home (yes)	0.56	0.69	-0.13	0.57	0.68	-0.11
Never has made purchase over the Internet	0.60	0.48	0.12	0.58	0.50	0.08
Comfort level in using computers (comfortable)	0.53	0.64	-0.11	0.57	0.60	-0.03
Gadget owner (late adopter)	0.37	0.27	0.10	0.35	0.30	0.05
Bought DVDs in last 12 months (yes)	0.21	0.30	-0.09	0.23	0.28	-0.05
Reader of Readers' Digest in last six months (yes)	0.57	0.47	0.10	0.54	0.51	0.03
Rented pre-recorded video tapes in last six months (yes)	0.52	0.62	-0.10	0.54	0.58	-0.04
Watches national news (every week/almost every week)	0.58	0.49	0.09	0.54	0.54	0.00
Searched for news on AOL in last six months (yes)	0.29	0.38	-0.09	0.29	0.37	-0.08
Gadget owner (early adopter/innovator)	0.09	0.17	-0.08	0.11	0.16	-0.05
Rented DVDs in last 12 months (yes)	0.11	0.18	-0.07	0.13	0.17	-0.04
Exercises by walking (yes)	0.72	0.64	0.08	0.70	0.66	0.04
Interest level in public affairs (very/somewhat)	0.70	0.60	0.10	0.67	0.62	0.05
Length of time using computers (more than a year)	0.67	0.76	-0.09	0.70	0.73	-0.03
Exercises by running/jogging (yes)	0.11	0.18	-0.07	0.13	0.16	-0.03
Interest level in national news (some/very)	0.60	0.50	0.10	0.58	0.56	0.02
Avg Difference	NA	NA	0.107	NA	NA	0.056

Note: Covariates for the MCA procedure are: Gender, Education (4 categories), HH Income (5 categories), Race/Ethnicity (4 categories), HH size (4 categories), MSA (2 categories)

Table 5 on the next page displays raw unadjusted and MCA-adjusted results for the questions on the KN political/public affairs background survey. Two of the questions from this background were listed in the “top 20” in Tables 4 and 5 (i.e., ‘watches national news,’ and ‘interest in public affairs’). The MCA adjustment reduced the average difference from 4.4 to 1.8 percentage points for the 16 questions as a group, a reduction of 59%. For the five attitudinal questions in the survey, the MCA adjustment reduced the average difference from 2.4 to 1.0 percentage points, a decrease of 58%. On political party ID, a one-point difference remained after the MCA adjustment.

Controlling for the Effect of Recruitment Cohort on Survey Results

The MCA procedure was also used with an additional covariate indicating the calendar quarter in which the political/public affairs background survey was completed. This covariate was added to control for the changes in attitudes and behavior evident in the time period 2000-2001. The second and primary motivation for adding the measure is to determine whether some of the differences between active and lost panelists is a function of differences between recruitment cohorts resulting, for instance, from underlying changes in public opinion and political behavior over time. Adding this covariate resulted in a change to only one MCA-adjusted measure shown in Table 5.⁷ The remaining measures did not change, which is suggestive that the effects of panel attrition on survey response are similar across the recruitment cohorts.

Within a specific recruitment cohort, survey estimates between active and lost panelists were different to the same extent observed across all the cohorts. The MCA procedure was employed for a specific recruitment cohort of 6,000 panelists completing the political/public affairs background in the same calendar quarter in 2000, and then sorting the cohort into active and lost categories based on their panel status in midyear 2002. The MCA-adjusted difference for this cohort across all the questions was 0.02, which is comparable to the 0.018 difference observed across cohorts.

⁷ The percentage of active panelists reporting approval of Congress increased from 30% to 31% as a result of adding the covariate for calendar quarter in which the background survey was completed.

Table 5: Political/Public Affairs Background Variables by Panel Status (Active vs. Lost), Unadjusted and MCA Adjusted

Questions	Unadjusted			MCA Adjusted		
	Active	Lost	Active - Lost	Active	Lost	Active - Lost
Watches local TV news (every week/almost every week)	0.76	0.66	0.10	0.73	0.68	0.05
Interest level in public affairs (very/somewhat)	0.70	0.60	0.10	0.67	0.62	0.05
Watches national news (every week/almost every week)	0.58	0.49	0.09	0.54	0.54	0.00
Current voter registration (registered)	0.83	0.76	0.07	0.81	0.77	0.04
Religion (Protestant)	0.39	0.33	0.06	0.36	0.35	0.01
Ever served in the U.S. Armed Forces, Reserves, or National Guard (never)	0.19	0.14	0.05	0.16	0.16	0.00
Overall impression of George W. Bush (favorable)	0.39	0.36	0.03	0.38	0.37	0.01
Member of a labor union (yes)	0.14	0.11	0.03	0.14	0.12	0.02
Watches cable news networks (every week/almost every week or more)	0.36	0.33	0.03	0.34	0.36	-0.02
Political party ID (Democrat)	0.59	0.56	0.03	0.58	0.57	0.01
Overall impression of Bill Clinton (favorable)	0.41	0.41	0.00	0.41	0.41	0.00
Direction of country (wrong direction)	0.52	0.53	-0.01	0.53	0.53	0.00
Approve or disapprove of the way the U.S. Congress has been handling its job (approve)	0.30	0.31	-0.01	0.30	0.31	-0.01
Religious service attendance (never/seldom)	0.29	0.30	-0.01	0.29	0.30	-0.01
How often search for news on the Internet(Every week/almost every week)	0.13	0.14	-0.01	0.12	0.15	-0.03
Political ideology (liberal)	0.40	0.46	-0.06	0.42	0.44	-0.02
Avg Difference	NA	NA	0.044	NA	NA	0.018

Note: Covariates for the MCA procedure are: Gender, Education (4 categories), HH Income (5 categories), Race/Ethnicity (4 categories), HH size (4 categories), MSA (2 categories)

Conclusions

The conclusions drawn from this paper are specific to the Knowledge Networks context; however, general themes are likely to emerge that apply to other panels constructed with probability sampling and that follow appropriate panel management techniques.

1. Overall, panelists that have dropped off the panel have similar attitudes, values, and behaviors compared to panelists that remain on the panel. The median difference across the 323 variables is 1.6 percentage points (weighted). In general, the differences are much smaller than the contribution normally made by sampling error to total survey error. The largest observed differences involve computer/Internet usage and experience questions.
2. The computer and Internet usage and access questions serve to highlight differences between active and lost panelists, with the latter having more access and experience. This suggests a line of future research for statistical nonresponse adjustments. However, these differences are not large enough to damage the representativeness of the panel with regard to Internet access.
3. On questions of values and 'psychographics,' active and lost panel members show very similar orientations (1.4 percentage point or 0.014 difference).
4. When controlling for demographic factors using MCA, the differences between active and lost panelists decreased by one half or more, demonstrating that many of the differences can be accounted for by demographic factors.
5. On political attitude questions, the active and lost panelists showed similar response patterns, with the MCA-adjusted differences averaging less than 1 percentage point (0.008) for the five political attitude questions.
6. In analyzing a specific recruitment cohort versus all cohorts combined, similar effects of panel attrition on survey estimates were found (on the order of 0.02 to 0.018 differences). Also, since controlling for panel recruitment cohort had no effect on the estimation of the effects of panel attrition, we conclude that replenishing cross-sectional samples are similar to initial recruits on this dynamic.

Appendix A: Knowledge Networks' Panel Recruitment Methodology

Knowledge Networks' panel recruitment methodology uses the quality standards employed by several gold-standard Random Digit Dialing (RDD) surveys conducted for the Federal Government.

Knowledge Networks utilizes list-assisted RDD sampling techniques on the sample frame consisting of the entire United States telephone population. The sample frame is updated quarterly. Knowledge Networks excludes only those banks of telephone numbers (consisting of 100 telephone numbers) that have zero directory-listed phone numbers. Knowledge Networks' telephone numbers are selected from the 1+ banks with equal probability of selection for each number. Note that the sampling is done without replacement to ensure that numbers already fielded by Knowledge Networks do not get fielded again.

Having generated the initial list of telephone numbers, the sample preparation system excludes confirmed disconnected and non-residential telephone numbers. Prior to August 2002, the sample was screened to identify telephone numbers that cannot use a local telephone call to connect to the MSN® TV Internet Service Provider network. Approximately 8% of the telephone numbers are of this type. Beginning in August 2001, these telephone numbers were subsampled for panel recruitment in order to reduce the potential for noncoverage error.

Telephone numbers for which Knowledge Networks is able to recover a valid postal address (about 65%) are sent an advance mailing informing them that they have been selected to participate in the Knowledge Networks Panel. In addition to information about the Knowledge Networks Panel, the advance mailings have contained a monetary incentive of various size or no incentive, and includes news articles and other materials highlighting studies that have been based on research using the web-enabled panel.

Following the mailing, the telephone recruitment process begins. The numbers called by interviewers consist of all numbers sent an advance mailing, as well as a random sample of telephone numbers not sent an advance mailing. The resulting cost efficiency more than offsets the decrease in precision caused by the need for sample weights. Cases sent to telephone interviewers are dialed up to 90 days, with at least 15 dial attempts on cases where no one answers the phone, and 25 dial attempts on phone numbers known to be associated with households. Extensive refusal conversion is also performed.

All recruitment interviews are conducted by experienced interviewers. An interview, which typically requires about 10 minutes, begins with the interviewer informing the household member that they have been selected to join the Knowledge Networks Panel. Prior to August 2002, they were told that in return for completing a short survey weekly, the household will be given a WebTV system. Starting in August 2002, households that reported to have computer-based Internet access at home were qualified to be enrolled into the web-enabled panel. All members in the household are then enumerated, and

some initial demographic variables and background information of prior computer and Internet usage are collected.

Each household recruited prior to August 2002 was provided with identical hardware, even if they currently own a computer or have Internet access. Microsoft's MSN® TV (WebTV) is the hardware platform currently used by the Knowledge Networks panel. The device consists of a set-top box that connects to a TV and the telephone. It also includes a remote keyboard and pointing device. WebTV has a built-in 56K modem that provides the household with a connection to the Internet. The base unit also has a small hard drive to accommodate large file downloads, including video files. File downloads do not require any user intervention and usually occur during off hours.

Prior to shipment, each unit is custom configured with individual email accounts, so that it is ready for immediate use by the household. Most households are able to install the hardware without additional assistance, though Knowledge Networks maintains a telephone technical support line and will, when needed, provide on-site installation. The Knowledge Networks Call Center also contacts household members who do not respond to e-mail and attempts to restore contact and cooperation.

All new panel members are sent an initial survey to confirm equipment installation and familiarize them with the WebTV unit. Demographics such as gender, age, race, income, and education are collected for each participant to create a member background. This information can be used to determine eligibility for specific studies and need not be gathered with each survey.

Beginning in August 2002, Knowledge Networks began recruiting households that already have home-based Internet access, and these recruited households download a utility from KN that facilitates their participation in the web-enabled panel, such as notification of survey invitations and managing multimedia survey content.

Appendix B: Technical Note on Weighting

In analyses using statistical weighting in this paper, separate weights were used for analysis of interviews from ‘lost’ panelists versus interviews from ‘active’ panelists.

Weights for Interviews from Lost Panelists. The weights employed for the analysis of interviews from ‘lost’ panelists were not calculated to correct for nonresponse caused by panel attrition or by nonresponse to any web-enabled panel surveys. The employed weight is a gross weight applied to the entire research panel ever recruited by Knowledge Networks (i.e., current plus lost panelists) to account for the probabilities of selection as well as nonresponse and noncoverage error resulting from panel recruitment. These poststratification adjustments are applied by cell ratio adjustments to the CPS benchmarks for age, race, sex, ethnicity, region, urban/rural, and education, not by raking.

Weights for Interviews from Active Panelists. The weights used for the analysis of ‘active’ cases are calculated for all panelists active at the time the analysis data file was constructed in midyear 2002. The weights were calculated to account for the probabilities of selection and nonresponse and noncoverage error due to panel recruitment, and nonresponse error resulting from panel attrition. These post-stratification adjustments are applied by cell ratio adjustments to the CPS benchmarks for age, race, sex, ethnicity, region, urban/rural, and education, not by raking.

For neither set of statistical weights did we employ poststratification weights specific to the background survey datasets that would have corrected for nonresponse to the specific background survey. This is a refinement that would undoubtedly narrow the differences observed between active and lost panelists.

For the analyses conducted using the MCA procedure, the same weight that was used for the analysis of lost panelists was employed.

References

Dennis, J. Michael. 2001. Response Timing and Coverage of Non-Internet Households: Data Quality in an Internet-enabled Panel. Presented at the May 2001 Conference of the American Association for Public Opinion Research.

Lepkowski, James M. and Couper, Mick P. 2002. Nonresponse in the Second Wave of Longitudinal Households Surveys, in *Survey Nonresponse*, edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, pp. 259-2774. NY: Wiley.

Robinson, J. 2002. Multiple Classification Analysis: Purpose and an Example. The document is located at http://www.webuse.umd.edu/handouts/tutorial/mca/MCAwebshop_Alan.pdf

Silver, R. C., E. A. Holman, D. N. McIntosh, M. Poulin and V. Gil-Rivas. 2002. Nationwide Longitudinal Study of Psychological Responses to September 11. *Journal of the American Medical Association*, 288(10): 1235-1244.

U.S. Bureau of the Census. 2001. *Home Computers and Internet Use in the United States: August 2000*. A Current Population Report by Eric C. Newberger.