Survey Information and Methodology

Introduction

Knowledge Networks conducted a study on different aspects of retirement on behalf of the Center for Retirement Research at Boston College. Specifically, the study examined:

- Retirement savings and money management before and after the economic downturn
- Key reasons for saving money
- Plans for retirement before and after the economic downturn
- Perceptions of job security and long-term financial security
- The roles of retirement age, savings and income, and how these factors relate to each other

The project consisted of two parts: A main survey followed by a one question follow-up (re-ask) fielded to a subset of the main survey respondents to obtain a clearer view of how they would deal with a loss of assets from the economic downturn.

The main survey was launched as a pretest beginning in late July 2009, with the full sample rollout occurring approximately one week later. Because the main survey did not change from the pretest, the pretest and main survey data were included as part of the final deliverable dataset. The re-ask survey was fielded in early October 2009.

To be selected for the main survey, a panel member must:

a) be between the ages of 45 and 59;
b) be in the labor force;
c) have Knowledge Networks Financial Services profile information for amount of retirement assets.

Knowledge Networks selected the sample to accommodate the "none or low", "some" and "high" retirement asset requirements based on its Financial Services profile data. Because of the economic downturn, it was possible that a person's level of assets could have changed between the time it was collected in the Financial Services profile and the fielding of the main survey. Therefore current level of retirement assets was re-screened in the field. The numbers of interviews in each group reflect calculations from the current level of retirement assets screened in field.

Participants completed the main survey in approximately 10 minutes and the Q43 Re-Ask survey in approximately 2 minutes (median survey completion time).

The breakdown by level of retirement assets (as screened in-field) for the main survey is presented below.

Age Group	Retirement Assets	N Qualified
45-52	None or Low	178
	Some	246
	High	260
53-59	None or Low	143
	Some	230
	High	260
Total		1,317

Overall Completion and Incidence Rates

While 2,221 respondents completed the main survey, 1,317 were qualified. The remaining 904 respondents screened out according to the termination criteria in the main survey. Data for all 2,221 records were provided as part of the final deliverables.

The field periods, completion rates and qualification rates are presented below.

Survey	Field Start	Field End	N Fielded	N Completed	% Completed	N Qualified	% Qualified
Main	8/1/2009	8/10/2009	2,941	2,221	75.5%	1,317	59.3%
Q43 Re-Ask	10/2/2009	10/18/2009	406	358	88.2%	n/a	n/a

Data File Deliverables and Descriptions

Knowledge Networks prepared SPSS and Stata datasets containing the main and re-ask survey data. An additional Stata dataset containing people who were screened out was provided to supplement the main survey data.

Several supplemental variables are provided to assist the principal investigators in identifying cases that could potentially be of interest. For instance, a DURATION variable shows the number of minutes of self-administration.

The strata variable was created for use by SUDAAN for variance estimation.[1] We replicate, as closely as possible, the implicit stratification used by MSG (the RDD sample frame vendor) in selecting RDD samples to create this variable. We sort all numbers in the sample by region (9 categories), urban/rural, county, area code, and telephone exchange. We then collapse neighboring exchanges so that there are at last five exchanges in each stratum.

In addition to the survey variables from the main and re-ask surveys, Knowledge Networks' standard profile and a series of data processing variables created by Knowledge Networks are provided in the data

[1] This is an option variable useful in epidemiological projects.

file for the condition eligible cases (n=2,221). The table below shows the name and description of each of the supplemental variables.[2]

Name	Label
CaseID	Serial number
weight2	Total Qualified Respondents
weight3	Total Qualified Respondents By Group
strata	Strata
tm_start	Interview start time
tm_finish	Interview finish time
duration	Interview duration in minutes
QFLAG	DATA ONLY: Qualification Flag
QUOTA	Quota
PPAGE	Age
ppagecat	Age -7 Categories
ppagect4	Age -4 Categories
PPEDUC	Education (Highest Degree Received)
PPEDUCAT	Education (Categorical)
PPETHM	Race / Ethnicity
PPGENDER	Gender
PPHHHEAD	Household Head
PPHHSIZE	Household Size
PPHOUSE	Housing Type
PPINCIMP	Household Income
PPMARIT	Marital Status
PPMSACAT	MSA Status
PPREG4	Region 4 -Based on State of Residence
ppreg9	Region 9 -Based on State of Residence
PPRENT	Ownership Status of Living Quarters
PPT01	Presence of Household Members -Children 0-2
PPT25	Presence of Household Members -Children 2-5
PPT612	Presence of Household Members -Children 6-12
PPT1317	Presence of Household Members -Children 13-17
PPT18OV	Presence of Household Members -Adults 18+
PPWORK	Current Employment Status
PPNET	HH Internet Access

[2] Supplemental Variables are selected as part of each projects design.

Knowledge Networks Methodology
Introduction

Knowledge Networks has recruited the first online research panel that is representative of the entire U.S. population. Panel members are randomly recruited by probability-based sampling, and households are provided with access to the Internet and hardware if needed.

Knowledge Networks selects households using random-digit dial (RDD) and address-based sampling methods. Once a person is recruited to the panel, they can be contacted by e-mail (instead of by phone or mail). This permits surveys to be fielded very quickly and economically. In addition, this approach reduces the burden placed on respondents, since e-mail notification is less obtrusive than telephone calls, and most respondents find answering Web questionnaires to be more interesting and engaging than being questioned by a telephone interviewer.

Panel Recruitment Methodology

Beginning recruitment in 1999, Knowledge Networks (KN) established the first online research panel (now called KnowledgePanel®) based on probability sampling that covers both the online and offline populations in the U.S. The panel members are randomly recruited by telephone and by self-administered mail and web surveys. Households are provided with access to the Internet and hardware if needed. Unlike other Internet research that covers only individuals with Internet access who volunteer for research, Knowledge Networks surveys are based on a dual sampling frame that includes both listed and unlisted phone numbers, telephone and non-telephone households, and cell-phone-only households. The panel is not limited to current Web users or computer owners. All potential panelists are randomly selected to join the KnowledgePanel; unselected volunteers are not able to join.

RDD and ABS Sample Frames

Knowledge Networks initially selects households using random digit dialing (RDD) sampling and address-based sampling (ABS) methodology. In this section, we will describe the RDD-based methodology, while the ABS methodology is described in a separate section below.

KnowledgePanel recruitment methodology uses the quality standards established by selected RDD surveys conducted for the Federal Government (such as the CDC-sponsored National Immunization Survey).

Knowledge Networks utilizes list-assisted RDD sampling techniques based on a sample frame of the U.S. residential landline telephone universe. For efficiency purposes, Knowledge Networks excludes only those banks of telephone numbers (a bank consists of 100 numbers) that have less than 2 directory listings. Additionally, an oversample is conducted among a stratum of telephone exchanges that have high concentrations of African-American and Hispanic households based on Census data. Note that recruitment sampling is done without replacement, thus numbers already fielded do not get fielded again.

A telephone number for which a valid postal address can be matched occurs in about 70% of the sample. These address-matched cases are all mailed an advance letter informing them that they have been selected to participate in KnowledgePanel. For efficiency purposes, the unmatched numbers are under-sampled at a current rate of 0.75 relative to the matched numbers. Both the oversampling mentioned above and this under-sampling of non-address households are adjusted appropriately in the panel's weighting procedures.

Following the mailings, the telephone recruitment begins for all sampled phone numbers using trained interviewer/recruiters. Cases sent to telephone interviewers are dialed for up to 90 days, with at least 14 dial attempts on cases where no one answers the phone, and on numbers known to be associated with households. Extensive refusal conversion is also performed. The recruitment interview, about 10 minutes long, begins with informing the household member that they have been selected to join KnowledgePanel. If the household does not have a computer and access to the Internet, they are told that in return for completing a short survey weekly, they will be provided with a laptop computer (previously a WebTV device was provided) and free monthly Internet access. All members in a household are then enumerated, and some initial demographic and background information on prior computer and Internet use are collected.

Households that inform interviewers that they have a home computer and Internet access are asked to take their surveys using their own equipment and Internet connection. Incentive points per survey, redeemable for cash, are given to these "PC" respondents for completing their surveys. Panel members who were provided with either a WebTV earlier or currently a laptop computer (both with free Internet access) do not participate in this per survey points incentive program. However, all panel members do receive special incentive points for select surveys to improve response rates and for all longer surveys as a modest compensation for burden.

For those panel members receiving a laptop computer (as with the former WebTV), 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. The Knowledge Networks Call Center contacts household members who do not respond to email and attempts to restore both contact and cooperation. PC panel members provide their own email addresses and we send their weekly surveys to that email account.

All new panel members are sent an initial survey to both welcome them as new panel members but also to familiarize them with how online survey questionnaires work. They also complete a separate profile survey that collects essential demographic information such as gender, age, race, income, and education to create a personal member profile. This information can be used to determine eligibility for specific studies, is used for weighting purposes, and operationally need not be gathered with each and every survey. This information is updated annually with each panel member. Once completed new member is "profiled," they are designated as "active" and ready to be sampled for client studies. [Note: Parental or legal guardian consent is also collected for conducting surveys with teenage panel members, ages 13-17.]

Once a household is contacted by phone—and additional household members recruited via their email address—panel members are sent surveys linked through a personalized email invitation (instead of by phone or mail). This permits surveys to be fielded quickly and economically, and also facilitates longitudinal research. In addition, this approach reduces the burden placed on respondents, since email notification is less obtrusive than telephone calls, and allows research subjects to participate in research when it is convenient for them.

Address-Based Sampling (ABS) Methodology

When KN started KnowledgePanel panel recruitment in 1999, the state of the art in the industry was that probability-based sampling could be cost effectively carried out using a national random-digit dial (RDD) sample frame. The RDD landline frame at the time allowed access to 96% of the U.S. population. This is no longer the case. We introduced the ABS sample frame to rise to the well-chronicled changes in society and telephony in recent years. The following changes have reduced the long-term scientific viability of the landline RDD sampling methodology: declining respondent cooperation to telephone surveys; do not

call lists; call screening, caller-ID devices and answering machines; dilution of the RDD sample frame as measured by the working telephone number rate; and finally, the emergence and exclusion of cell-phone-only households (CPOHH) because they have no landline phone.

According to the Center for Disease Control, approximately 25% of U.S. households cannot be contacted through RDD sampling: 22% as a result of CPOHH status and 3% because they have no phone service whatsoever. Among some segments of society, the sample non-coverage is substantial: more than one-third of young adults, ages 18-24, reside in CPOHHs.

After conducting an extensive pilot project in 2008, we made the decision to add an address-based sample (ABS) frame in response to the growing number of cell-phone only households that are outside of the RDD frame. Before conducting the ABS pilot, we also experimented with supplementing our RDD samples with cell-phone samples. However, this approach was not cost effective for you our clients and raised a number of other operational, data quality, and liability issues (e.g., calling people's cell phones while they were driving).

The key advantage of the ABS sample frame is that it allows sampling of almost all U.S. households. An estimated 98% of households are "covered" in sampling nomenclature. Regardless of household telephone status, they can be reached and contacted via the mail. Second, our ABS pilot project revealed some other advantages beyond the expected improvement in recruiting adults from CPOHHs:

- Improved sample representativeness for minority racial and ethnic groups
- Improved inclusion of lower educated and low income households
- Exclusive inclusion of CPOHHs that have neither a landline telephone nor Internet access (approximately 4% to 6% of US households).

ABS involves probability-based sampling of addresses from the U.S. Postal Service's Delivery Sequence File. Randomly sampled addresses are invited to join KnowledgePanel through a series of mailings and in some cases telephone follow-up calls to non-responders when a telephone number can be matched to the sampled address. Invited households can join the panel by one of several means:

- by completing and mailing back a paper form in a postage-paid envelope;
- by calling a toll-free hotline maintained by Knowledge Networks; or
- by going to a designated KN web-site and completing an online recruitment form.

After initially accepting the invitation to join the panel, respondents are then "profiled" online answering key demographic questions about themselves. This profile is maintained using the same procedures established for the RDD-recruited research subjects. Respondents not having an Internet connection are provided a laptop computer and free Internet service. Respondents sampled from ABS frame, like those from the RDD frame are provided the same privacy terms and confidentiality protections that we have developed over the years and have been reviewed by dozens of Institutional Review Boards.

Large-scale ABS sampling for our KnowledgePanel recruitment began in April, 2009. As a result, KnowledgePanel will be improving its sample coverage of CPOHHs and young adults.

Because we will have recruited panelists from two different sample frames – RDD and ABS – we are taking several technical steps to merge samples sourced from these frames. Our approach preserves the representative structure of the overall panel for the selection of individual client study samples. An advantage of mixing ABS frame panel members in any KnowledgePanel sample is a reduction in the

variance of the weights. ABS-sourced sample tends to align more true to the overall population demographic distributions and thus the associated adjustment weights are somewhat more uniform and less varied. This variance reduction efficaciously attenuates the sample's design effect and confirms a real advantage for study samples drawn from KnowledgePanel with its dual frame construction.

Survey Administration

For client surveys, samples are drawn at random from among active panel members. Depending on the study, eligibility criteria will be applied or in-field screening of the sample will be carried out. Sample sizes can range widely depending on the objectives and design of the study.

Once assigned to a survey, members receive a notification email letting them know there is a new survey available for them to take. This email notification contains a link that sends them to the survey questionnaire. No login name or password is required. The field period depends on the client's needs, and can range anywhere from a few hours to several weeks.

After three days, automatic email reminders are sent to all non-responding panel members in the sample. Additional email reminders were sent out one week into the field period for both the main and re-ask surveys. To maximize the survey completion rate, one additional email reminder was sent out after two weeks in the field for the re-ask survey. If email reminders do not generate a sufficient response, an automated telephone reminder call may be initiated. The usual protocol is to wait at least three-four days after the email reminder before calling. To assist panel members with their survey taking, each individual has a personalized "home page" that lists all the surveys that were assigned to that member and have yet to be completed.

Knowledge Networks also operates an ongoing, modest, incentive program to encourage participation and create member loyalty. Members can enter special raffles or can be entered into special sweepstakes with both cash and other prizes to be won.

The typical survey commitment for panel members is one survey per week or four per month with a duration of 10-15 minutes per survey. Some client surveys exceed this time and in the case of longer surveys an additional incentive may be provided.

Survey Sampling from KnowledgePanel

Once Panel Members are recruited and profiled, they become eligible for selection for specific client surveys. In most cases, the specific survey sample represents a simple random sample from the panel, for example, a general population survey. Customized stratified random sampling based on profile data may also be conducted as required by the study design.

The general sampling rule is to assign no more than one survey per week to members. Allowing for rare weekly exceptions, this limits a member's total assignments per month to 4 or 6 surveys. In certain cases, a survey sample calls for pre-screening, that is, members are drawn from a subsample of the panel (such as, females, Republicans, grocery shoppers, etc.). In such cases, care is taken to ensure that all subsequent survey samples drawn that week are selected in such a way as to result in a sample that remains representative of the panel distributions.

For the main survey, a nationally representative sample of U.S. adults in the labor force aged 45-59 with varying levels of retirement assets was selected.

Sample Weighting

The design for a KnowledgePanel® sample begins as an equal probability sample that is self-weighting with several enhancements incorporated to improve efficiency. Since any alteration in the selection process is a deviation from a pure equal probability sample design, statistical weighting adjustments are made to the data to offset known selection deviations. These adjustments are incorporated in the sample's **base weight**.

There are also several sources of survey error that are an inherent part of any survey process, such as non-coverage and non-response due to panel recruitment methods and to inevitable panel attrition. We address these sources of sampling and non-sampling error using a **panel demographic post-stratification weight** as an additional adjustment.

Lastly, a set of **study-specific post-stratification weights** are constructed for the study data to adjust for the study's sample design and survey non-response.

A description of these types of weights follows.

The Base Weight

In a KnowledgePanel sample there are seven known sources of deviation from an equal probability of selection design. These are corrected in the Base Weight and are described below.

1. Under-sampling of telephone numbers unmatched to a valid mailing address

An address match is attempted on all the Random Digit Dial (RDD) generated telephone numbers in the sample after the sample has been purged of business and institutional numbers and screened for non-working numbers. The success rate for address matching is in the 60-70% range. The telephone numbers with valid addresses are sent an advance letter, notifying the household that they will be contacted by phone to join KnowledgePanel. The remaining, unmatched numbers are under-sampled as a recruitment efficiency strategy. Advance letters improve recruitment success rates. Under-sampling stopped between July 2005 and April 2007. It was resumed in May 2007 with a sampling rate of 0.75.

2. RDD selection proportional to the number of telephone landlines reaching the household.

As part of the field data collection operation, information is collected on the number of separate telephone landlines in each selected household. A multiple line household's selection probability is down weighted by the inverse of its number of landlines.

3. Some minor oversampling of Chicago and Los Angeles due to early pilot surveys

Two pilot surveys carried out in Chicago and Los Angeles when the panel was first being built increased the relative size of the sample from these two cities. With natural attrition and growth in size, the impact is disappearing over time. It remains part of our base adjustment weighting because of a small number of extant panel members from that nascent panel cohort.

4. Early oversampling the four largest states and central region states

At the time when the panel was first being built, survey demand in the four largest states (California, New York, Florida, and Texas) required over-sampling during January-October 2000.

Similarly, the central region states were over-sampled for a brief period. These now diminishing effects still remain in the panel membership and thus require weighting adjustments for these geographic areas.

5. Under-sampling of households not covered by the MSN® TV service network

Certain small areas of the U.S. are not serviced by MSN® , thus our MSN®TV units cannot be used for recruited non-Internet households. In some of these cases, we use other Internet Service Providers for Internet access via the member's personal computer. Overall, the result is a small under-sample of these geographic areas thus requiring a minor weighting adjustment.

6. Oversampling of African- American and Hispanic telephone exchanges

As of October 2001, we began over-sampling telephone exchanges with a higher density of minority households (specifically African American and Hispanic) to increase panel membership for those groups. These exchanges are oversampled at approximately twice the rate of other exchanges. This over-sampling is corrected in the base weight.

7. Address-based sample phone match adjustment

Towards the end of 2008, Knowledge Networks began recruiting panel members using an address-based sample (ABS) frame in addition to RDD recruitment. Once recruitment through the mail, including follow-up mailings to ABS non-respondents was completed, a telephone recruitment was added. Non-responding ABS households where a landline telephone number could be matched to an address were subsequently called and a telephone recruitment initiated. This effort resulted in a slight overall disproportionate number of landline households being recruited in a given ABS sample. A base weight adjustment is applied to return the ABS recruitment panel members to the sample's correct national proportion of phone-match and no phone match households.

The Panel Demographic Post-stratification Weight

To reduce the effects of any non-response and non-coverage bias in the overall panel membership, a post-stratification adjustment is applied using demographic distributions from the most recent data from the Current Population Survey (CPS). Benchmark distributions for Internet Access among the U.S. population of adults are obtained from KnowledgePanel recruitment data since this measurement is not collected as part of the CPS.

The post-stratification variables include:

- Gender (Male/Female)
- Age (18-29, 30-44, 45-59, and 60+)
- Race/Hispanic ethnicity (White/Non-Hispanic, Black/Non-Hispanic, Other/Non-Hispanic, 2+ Races/Non-Hispanic, Hispanic)
- Education (Less than High School, High School, Some College, Bachelor and beyond)
- Census Region (Northeast, Midwest, South, West)
- Metropolitan Area (Yes, No)
- Internet Access (Yes, No)

This weighting adjustment is applied prior to the selection of any client sample from KnowledgePanel. These weights constitute the starting weights for any client survey selected from the panel.

Study-Specific Post-Stratification Weights

Once all the study data are returned from the field, we proceeded with a post-stratification process to adjust for any survey non-response and also any non-coverage due to the study-specific sample design. Demographic and geographic distributions for the population ages 45-59 from the most recent Current Population Survey (CPS) are used as benchmarks in this adjustment.

The following benchmark distributions are utilized for this post-stratification adjustment:

- Gender (Male/Female)
- Age (45-48, 49-52, 53-55 and 56-59)
- Race/Hispanic ethnicity (White/Non-Hispanic, Black/Non-Hispanic, Other/Non-Hispanic, 2+ Races/Non-Hispanic, Hispanic)
- Education (Less than High School, High School, Some College, Bachelor and beyond)
- Census Region (Northeast, Midwest, South, West)
- Household income (Under $25,000, $25,000-$49,999, $50,000-$74,999, $75,000 and above)
- Metropolitan Area (Yes, No)
- Internet Access (Yes, No)

All survey respondents were weighted to look like all eligible respondents using these benchmark distributions. Comparable distributions are calculated using all completed cases from the field data. Since study sample sizes are typically too small to accommodate a complete cross-tabulation of all the survey variables with the benchmark variables, an iterative proportional fitting is used for the post-stratification weighting adjustment. This procedure adjusts the sample data back to the selected benchmark proportions. Through an iterative convergence process, the weighted sample data are optimally fitted to the marginal distributions.

After this final post-stratification adjustment, the distribution of the calculated weights are examined to identify and, if necessary, trim outliers at the extreme upper and lower tails of the weight distribution. The post-stratified and trimmed weights are then scaled to the sum of the total sample size of all eligible respondents (weight1).

Based on weight1, benchmarks based on those in the six age by retirement assets groups were calculated. The weights per group were then trimmed and scaled to sum to the sample size of total qualified respondents (weight2) and qualified respondents by group (weight3).