

## **KnowledgePanel®: Processes & Procedures Contributing to Sample Representativeness & Tests for Self-Selection Bias**

**March 2010**

Prepared by **J. Michael Dennis**, EVP, Government & Academic Research, Knowledge Networks  
([mdennis@knowledge networks.com](mailto:mdennis@knowledge networks.com))

We often are asked by reviewers of manuscripts submitted to scientific journals to describe the sampling and data collection procedures as well as the results of statistical tests for self-selection bias.

This brief research note is a summary of some of the key process-based factors – besides the probability basis that is the sampling foundation for KnowledgePanel® – responsible for the representativeness of KnowledgePanel, and the highlights of statistical tests on the same.

### **Table of Contents**

<b>Processes and Procedures Contributing to Sample Representativeness.....</b>	<b>2</b>
<b>Statistical Tests for Sample Representativeness and Self-Selection Bias.....</b>	<b>2</b>
<b>Analyses of Sample Representativeness.....</b>	<b>3</b>
<b>Tests for Comparability of Survey Results to Benchmarks.....</b>	<b>4</b>
<b>Direct Measurement of Non-Response Bias .....</b>	<b>6</b>
<b>Measurement of Self-Selection Bias.....</b>	<b>7</b>
<b>References .....</b>	<b>9</b>

## Processes and Procedures Contributing to Sample Representativeness

First, we describe the processes and procedures that serve to prevent self-selection bias and achieve sample representativeness. Here are highlights:

**Sample Frame:** Our sample frame is the universe of U.S. residential addresses from the U.S. Postal Service Delivery Sequence File, providing sample coverage for approximately 99% of U.S. households (DiSogra, 2009). The USPS DSF includes cell-phone-only households.

**Sample Selection for the Panel Recruitment Sampling:** Our panel recruitment starts with drawing probability-based random samples from the USPS DSF. From time to time, we modify the sampling protocol in order to oversample (on a modest level) local areas that are high in population density in non-white groups or Hispanic ethnicity.

**Panel Recruitment Methodology for Gaining Cooperation:** Our mail-based panel recruitment methodology also includes an intensive non-response refusal conversion component that is conducted by professional telephone interviewers.

**Inclusion of Non-Internet Households and Technically Challenged Households:** KnowledgePanel households that do not have Internet access are provided a laptop computer and Internet access; in addition, our panel case managers provide telephone support for all technically challenged households that need help connecting their computers to the Internet and accessing their email and Internet surveys.

**Inclusion of Spanish-Language Households:** The panel recruitment mail-out materials and telephone-based recruitment are conducted in both English and Spanish; Spanish-language households are recruited as well by oversampling metro areas that are high-density Latino, with recruitment conducted both by telephone and by mail.

**Low Burden of Panel Membership:** The burden of panel membership is managed to avoid respondent fatigue by selecting respondents for surveys no more than once a week (and frequently less often) and typically for short surveys.

**Panel Maintenance & Panel Loyalty Programs:** Our panel maintenance and loyalty programs help respondents feel that they are contributing and are part of a larger feedback system for decision-makers in government and the private sector; moreover, our panel loyalty program rewards consistent and long-term panel participation with in-kind giveaways and monetary payouts tied to long-term loyalty.

Those are a few of the key processes and procedures that facilitate our providing representative samples for statistical surveys. For more information, please read the "Design Summary" available at <http://www.knowledgenetworks.com/knpanel/index.html>.

## Statistical Tests for Sample Representativeness and Self-Selection Bias

With respect to sample representativeness and self-selection bias, a number of methodological studies have been conducted, most of which are available for downloading at our Web site at <http://www.knowledgenetworks.com/ganp/reviewer-info.html>. KN researchers have been active in conducting our own independent research on this topic or helping our customers conduct their own research on the topic. Prior to their publication, many of the peer-reviewed articles based on KnowledgePanel data were first subject to extensive inspection by scientific reviewers, often with specific tests for sample representativeness.

The relevant statistical studies can be grouped into one of the following types:

- Analyses of Sample Representativeness
- Tests for Comparability of Survey Results to Benchmarks
- Direct Measurement of Non-Response Bias
- Measurement of Self-Selection Bias

Below is a brief description of examples for each of these types of studies.

### **Analyses of Sample Representativeness**

KnowledgePanel survey samples are well known for resembling the U.S. Census benchmarks for primary demographics. Later, I will show evidence from the Yeager and Krosnick paper that KN secondary demographics are also closely in line with the Census Benchmarks. In the table below, the weighted distribution of Adult KN Members (in total) is displayed for primary demographics; the survey data are weighted for non-response and non-coverage. Interview survey data for general population surveys often approximate this high level of sample representativeness. This level of representativeness is found in RDD and in-person surveys only after substantial investments have been made in refusal conversion.

**Table 1: Knowledge Networks Panel: Demographic Characteristics Compared to U.S. Census Benchmarks (*Weighted*)**

<b>Demographics</b>		<b>Adult Panel Members</b>	<b>Adult U.S. Population (June 2009 CPS)</b>
<b>Gender</b>	Male	47.3%	48.3%
	Female	52.7%	51.7%
<b>Age</b>	18–24	10.4%	12.6%
	25–34	17.7%	17.8%
	35–44	19.1%	18.1%
	45–54	18.9%	19.6%
	55–64	18.3%	15.3%
	65 or over	15.7%	16.7%
<b>Race</b>	White	79.5%	81.2%
	Black (African American)	12.4%	11.8%
	American Indian, Alaskan Native	1.1%	0.8%
	Asian	1.8%	4.6%
	Hawaiian or Pacific Islander	0.4%	0.3%
	2+ Races	4.7%	1.3%
<b>Hispanic Ethnicity</b>	Hispanic	14.0%	13.8%
	Non-Hispanic	86.0%	86.2%
<b>Employment Status</b>	In the Labor Force	67.4%	67.6%
	Not in the Labor Force	32.6%	32.4%
<b>Marital Status</b>	Married	53.4%	55.5%
	Not Married	46.6%	44.5%
<b>Housing Ownership</b>	Own	72.9%	71.0%
	Rent/Other	27.1%	29.0%

## Tests for Comparability of Survey Results to Benchmarks

### 1. Yeager, Krosnick et al Comparison Survey

A landmark study has been completed recently that is the most comprehensive comparison of KnowledgePanel survey results to RDD telephone and to opt-in non-probability Web panels. The citation for the paper is: David Yeager, Jon Krosnick, LinChiat Chang, Harold Javitz, Matthew Levindusky, Alberto Simpser, Rui Wang. "Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples," August 2009, available at <http://communication.stanford.edu/faculty/krosnick/>.

The authors administered the same survey instrument to multiple samples: Seven non-probability Internet survey platforms were compared to two probability-based survey platforms (telephone survey using RDD and KnowledgePanel®). The authors' main conclusion was that the non-probability Internet surveys were less accurate, and customary weighting adjustments did not uniformly improve them.

As shown below, the KnowledgePanel sample was the most representative in terms of primary demographics, even more representative than RDD. The results below are not weighted. On average, the KN interview cases were only 2.47 percentage points off from the Census benchmarks.

**Table 2: Average Error from Benchmarks**  
**Comparison Table: Six Primary Demographics<sup>c</sup>**  
**(Without Post-Stratification)**

Probability Samples		Non-Probability Samples						
Telephone (RDD)	Internet (KnowledgePanel)	1	2	3	4	5	6	7
3.43 <sup>b</sup>	2.47	4.14 <sup>b</sup>	4.96 <sup>ab</sup>	6.44 <sup>ab</sup>	6.35 <sup>ab</sup>	7.01 <sup>ab</sup>	6.05 <sup>ab</sup>	12.82 <sup>ab</sup>

<sup>a</sup> Significantly different from telephone survey at  $p < .05$

<sup>b</sup> Significantly different from probability Internet survey at  $p < .05$

<sup>c</sup> Age, Gender, Race, Ethnicity, Education, and Region

The authors also compared the platforms' sample representativeness in terms of secondary demographics such as marital and employment status. As shown below, the RDD and KnowledgePanel representativeness were comparable (less than four percentage points off the Census benchmarks), while the non-probability samples were far less representative. These data are not weighted.

**Table 3: Average Error from Benchmarks**  
**Comparison Table: Thirteen Secondary Demographics and Non-Demographics<sup>c</sup>**  
**(Without Post-Stratification)**

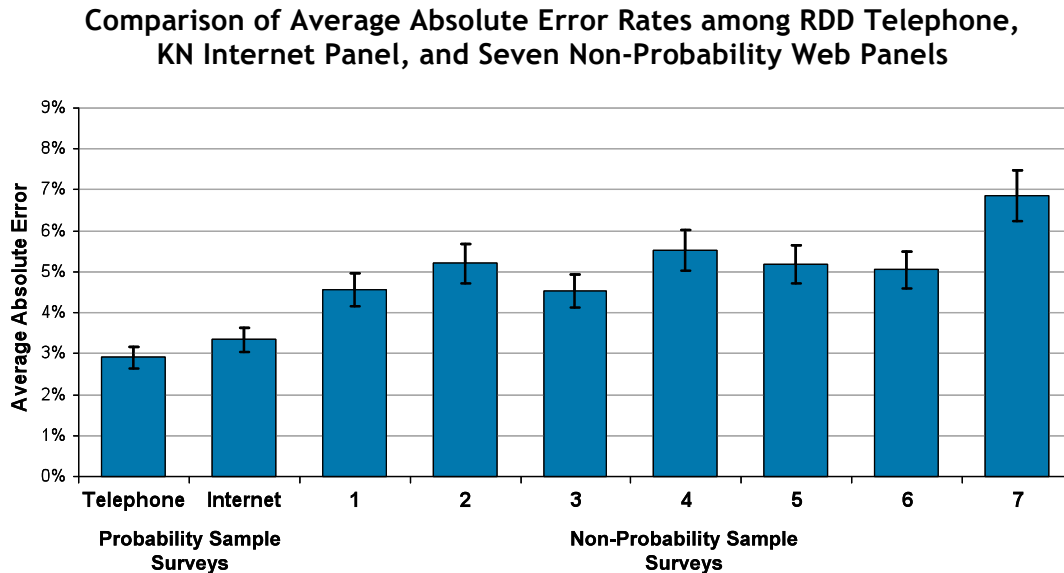
Probability Samples		Non-Probability Samples						
Telephone (RDD)	Internet (KnowledgePanel)	1	2	3	4	5	6	7
3.64	3.96	5.25 <sup>ab</sup>	5.79 <sup>ab</sup>	6.05 <sup>ab</sup>	4.79 <sup>a</sup>	4.81 <sup>a</sup>	5.38 <sup>ab</sup>	8.93 <sup>ab</sup>

<sup>a</sup> Significantly different from telephone survey at  $p < .05$

<sup>b</sup> Significantly different from probability Internet survey at  $p < .05$

<sup>c</sup> Secondary Demographics: Marital Status, Number of People in Household, Employment Status, Number of Bedrooms in Home, Number of Vehicles Owned, Home Ownership, and Household Income. Non-Demographics: Frequency of Smoking Cigarettes, 12 Drinks of Alcohol or More Lifetime, Average Number of Drinks of Alcohol on Days of Drinking, Ratings of Quality of Their Health, Possession of U.S. Passport, and Possession of a Driver's License.

The authors summarized the error rates for each survey platform. Error rates were calculated as the difference between the objective benchmark (often from U.S. government statistics) and the survey estimate. Again, as shown above for telephone, KN Internet, and the seven non-probability Web panels, the RDD and KnowledgePanel estimates were comparably accurate, far more accurate than the results from the non-probability Web panels.



The authors' key findings are:

- **Probability sample surveys** are consistently more accurate whether they are on phone or Internet platforms.
- **Non-Probability sample Internet surveys** were always less accurate.
- **Post-stratification with demographics** sometimes improved and sometimes reduced the accuracy of non-probability surveys - therefore this method cannot be relied upon to repair sampling deficiencies.
- **False claim made by advocates of non-probability samples:** *"Optimizing methods of conducting non-probability Internet surveys can maximize their accuracy."* One of the non-probability surveys was strikingly inaccurate (#7), and the rest were evenly inaccurate at a lower level.
- **False claim by advocates of non-probability samples:** *"A higher completion rate is an indication of higher accuracy."* Completion rates of non-probability surveys were slightly negatively correlated with their accuracy.
- **Low response rates with probability samples do not necessarily mean low levels of accuracy**, if substantial efforts are made to interview as many respondents as possible (i.e., callbacks, follow-ups, incentives, etc.).

Dennis, Osborn (et al, 2009) similarly found significant differences between probability-based and non-probability-based Web panels on attitudinal questions and on measures of Internet usage and frequency of survey taking.

## 2. Comparing the General Social Survey Findings to the KnowledgePanel Survey

One of the more comprehensive tests of KnowledgePanel was a comparison of the KnowledgePanel results to those obtained for the General Social Survey, which is the gold-standard Federally funded survey that tracks social attitude trends in the U.S. KN staff members have co-authored papers and articles based on a series of studies based on a comparison of KnowledgePanel methodology to the in-person survey conducted by NORC. The 2007 article by Dennis and Li is available for downloading at <http://ijor.mypublicsquare.com/view/more-honest-answers>; the article supports and replicates the findings from the 2004 comparison of the KnowledgePanel and GSS data (see Smith and Dennis, 2005). This study examines the role of interviewer effects in accounting for differences observed between the telephone and in-person administrations of the GSS items versus online administration of the same items.

In the Dennis and Li article, an experimental design is employed that provides empirical support for the pattern documented elsewhere in the literature that interviewer-administered surveys can be affected by “social desirability bias.” The authors concluded the following based on the experiment:

These observations lead us to conclude that there are important differences in the survey results that are attributable to the presence of an interviewer for the in-person and telephone modes, and to the absence of an interviewer in the web mode. The direction of the differences in the survey results, as seen in how respondents are more likely to report in the web mode that the country spends “too much” on certain problems in society, is consistent with the conclusion that web panel respondents are more honest and exhibit more candor in their responses, compared to interviewer-administered surveys. This conclusion is reinforced by the experimental design of our study, which controlled for the source of the sample. To be clear, we are not indicating that we know the “true” measure for public opinion, nor are we suggesting that the online mode survey results are closer to the “truth” about U.S. public opinion. However, we do believe that the differences we observe in the survey results are consistent with the hypothesis that online respondents feel less potent pressure to answer questions in socially desirable ways (Dennis and Li, 2007).

### **Direct Measurement of Non-Response Bias**

#### 1. The Boston University Survey Replicating NESARC Using KnowledgePanel

One of the most thorough examinations of KnowledgePanel and its usefulness for public health research was conducted by researchers from the Youth Alcohol Center at Boston University. In research supported by the National Institute of Alcohol and Alcoholism of NIH, the researchers commissioned Knowledge Networks to conduct an epidemiological survey in a replication test of the gold-standard survey conducted by the Census Bureau. The authors’ methodological findings were published in Alcoholism: Clinical and Experimental Research (Heeren, 2007). The article compares results from the KN survey to results from the National Epidemiologic Study on Alcohol and Related Conditions (NESARC), a face-to-face probability sample survey of 43,093 adults, with a focus on associations between demographics, age of drinking onset, and alcohol dependence. In their conclusion, the authors stated that KnowledgePanel, as it is based on probability sampling, provides an alternative to random-digit-dial telephone surveys and in-person surveys for studies of factors associated with alcohol-related problems.

The Boston University survey featured a non-response follow-up survey of non-respondents. These non-respondents were randomly selected from the pool of research subjects that refused to join the KnowledgePanel at the time of panel recruitment or joined KnowledgePanel but later dropped off the panel. This non-response follow-up survey (over 600 interviews) provided a means for the direct measurement of the attitudes and behaviors of non-respondents. The interviews were conducted by telephone. The authors compared the findings from the non-response follow-up survey to those from the KnowledgePanel Web survey. They found no associations between the type of sample (non-responders

versus KnowledgePanel members) and the risk factors for alcohol dependence (such as family history, risky drinking category, and age of onset on alcohol drinking).

## 2. Civic Attitudes After 9/11: Comparison of KnowledgePanel and KnowledgePanel Non-Responders

A substantial study of non-response bias and the effects of mode of data collection was conducted by RTI and Knowledge Networks researchers a few months after 9/11 (see Dennis, Chatt, et al, 2005). The full paper and instrumentation are available at: <http://www.knowledgenetworks.com/ganp/rtimode.html>.

At the time, KnowledgePanel surveys were relatively new and unproven. The purpose of the research was to explore the potential for Internet panel-based survey research by conducting an experiment to investigate survey error that could hinder the validity of Internet-based survey results. In this experiment, the KnowledgePanel methodology was compared to traditional telephone surveys through an experimental design that controls for sample origin. Although previous research done on telephone-Internet surveys had addressed data collection mode effects, none had been done which controls for sample origin. The experiment was embedded in the design of the *Survey of Civic Attitudes and Behaviors After 9/11*, a study sponsored by RTI International and co-designed by RTI and the Odum Institute at the University of North Carolina.

Three randomly selected sample groups completed the *Survey of Civic Attitudes and Behaviors After 9/11*: i) an Internet survey of active KnowledgePanel members, ii) a telephone survey of active KnowledgePanel members, and iii) a telephone survey of persons refusing to join KnowledgePanel and those KnowledgePanel members who did not respond to the Web survey. The first two random samples were drawn from active KnowledgePanel members, but differed in the mode of data collection (Internet versus telephone). The second and third samples overlapped in terms of mode of data collection (both are telephone), but the two groups differed in terms of sample origin (active KnowledgePanel members versus refusals). The design, therefore, provides a control group of KnowledgePanel members who participated using the telephone mode of data collection. Various univariate and multivariate statistical tests were conducted in order to measure differences associated with mode of data collection and sample origins. The sources of error investigated are sample representativeness, mode effects, sample effects, panel experience effects, primacy and recency effects, the effects of visual versus aural survey administration, and non-differentiation in survey answers.

Differences among sample groups were found to be due primarily to mode of data collection and panel experience, and somewhat due to sample origin. Basic differences between Internet surveys and telephone interviews could be traced back to mode of data collection: The telephone interview data collected from the KnowledgePanel members were very similar to the telephone interview data from the KN non-respondents. These data indicated that the attitudes and opinions of KnowledgePanel members and those that refused to join KnowledgePanel are very similar. The differences found between the mode of data collection in this telephone versus Internet study were strikingly similar to the telephone versus mail mode effects found in civic attitude studies by Tarnai and Dillman and in telephone versus face-to-face mode effects by Krysan. Both studies found a tendency for telephone respondents to answer at the extreme positive end of the scale. In addition, this study found that Internet respondents were more likely than the telephone sample to use the full range of response option scales; therefore, non-differentiation of survey responses was more prevalent in the telephone sample groups.

### **Measurement of Self-Selection Bias**

Knowledge Networks has participated in statistical tests using the Heckman selection correction technique for the detection of self-selection bias into the panel and self-selection into actual KnowledgePanel surveys. The statistical techniques, the ecological databases constructed, and the analysis itself are involved and require no modest amount of effort to understand their intricacies, and even more effort to evaluate the work that has been done. The most comprehensive test to date was conducted by Professor Trudy Cameron of the University of Oregon. Her research is detailed as Appendix D entitled "The Knowledge Networks Panel and Sample Selection Corrections" for the paper

T.A. Cameron and J.R. DeShazo (2008) "[Demand for Health Risk Reductions](http://www.uoregon.edu/~cameron/vita/wpabstracts.htm)" (currently in revise resubmit status). The full paper and appendix are available at <http://www.uoregon.edu/~cameron/vita/wpabstracts.htm>.

Cameron's research found that application of the Heckman selection correction procedure using the RDD sample frame as the base did not support the hypothesis that attitudes towards regulatory issues are correlated with propensity to participate in a KnowledgePanel survey. This test supports the hypothesis that self-selection bias is not an important factor in KnowledgePanel surveys on the subject area of attitudes towards government regulation.

A second but less sophisticated implementation test of Heckman selection correction procedure also did not support the hypothesis that valuations of water quality are highly correlated with the propensity to participate in a KnowledgePanel survey. The use of the Heckman selection correction procedure resulted in an adjusted estimate of -6.16%. This test supports the hypothesis that self-selection bias is not an important factor in KnowledgePanel surveys on the subject area of valuations of public goods such as water quality. (Source: Joel Huber, W. Kip Viscusi, Jason Bell, "The Value of Regional Water Improvements: Further Evidence," Presented at the Valuation of Ecological Benefits Conference, EPA, October 2004.) In the same report, the authors also concluded that valuation results of an ecological public good were independent of KN respondent panel characteristics such as time in panel, time to complete survey, and high likelihood of quitting panel.



## References

**Cameron, Trudy A. and J.R. DeShazo. 2008.** The Knowledge Networks Panel and Sample Selection Corrections. See the full paper and technical appendix available at <http://www.uoregon.edu/~cameron/vita/wpabstracts.htm>.

**Dennis, J. Michael. 2009.** Summary of KnowledgePanel® Design. Available at <http://www.knowledgenetworks.com/ganp/reviewer-info.html>.

**Dennis, J. M., C. Chatt, R. Li, A. Motta-Stanko, and P. Pulliam. 2005.** Data Collection Mode Effects Controlling for Sample Origins in a Panel Survey: Telephone versus Internet. Paper presented at the 2005 Annual Meeting of the American Association for Public Opinion Research. Available at <http://www.knowledgenetworks.com/ganp/rtimode.html>.

**Dennis, J.M. and R. Li. 2007.** More Honest Answers to Surveys? A Study of Data Collection Mode Effects. Interactive Marketing Research Organization's (IMRO) Journal of Online Research. Available at <http://www.knowledgenetworks.com/ganp/reviewer-info.html>.

**Dennis, J. Michael, Larry Osborn, and Karen Semans. March 2009.** Comparison Study of Early Adopter Attitudes and Online Behavior in Probability and Non-Probability Web Panels. Available at <http://www.knowledgenetworks.com/ganp/reviewer-info.html>.

**DiSogra, Charles, Mario Callegaro, and Erlina Hendarwan. 2009.** Recruiting Probability-based Web Panel Members using an Address-based Sample Frame: Results from a Pilot Study Conducted by Knowledge Networks. Presented at the 2009 Joint Statistical Meetings. Available at <http://www.knowledgenetworks.com/ganp/reviewer-info.html>.

**Heeren, Timothy, Erika M. Edwards, J. Michael Dennis, Sergei Rodkin, Ralph W. Hingson, and David L. Rosenbloom.** A Comparison of Results from an Alcohol Survey of a Prerecruited Internet panel and the National Epidemiologic Survey on Alcohol and Related Conditions. Alcoholism: Clinical and Experimental Research 2008; 32(2):222-9.

**Smith, Tom W., and J. M. Dennis. December 2005.** Online versus In-Person: Experiments with Mode, Format, and Question Wordings. Public Opinion Pros. Available under "Past Issues" at <http://www.publicopinionpros.norc.oregon.edu/index.asp>.

**Yeager, David, Jon Krosnick, LinChiat Chang, Harold Javitz, Matthew Levindusky, Alberto Simpser, and Rui Wang. August 2009.** Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. Available at <http://communication.stanford.edu/faculty/krosnick/>.

For more information, contact [mdennis@knowledgenetworks.com](mailto:mdennis@knowledgenetworks.com).