How Far Have We Come?

The Lingering Digital Divide and Its Impact on the Representativeness of Internet Surveys

J. Michael Dennis, Ph.D.
Executive Vice President and Managing Director
Government and Academic Research
GfK CRNA

American Association of Public Opinion Research | May 16-19 | Boston, MA
Co-Authors

Curtiss L. Cobb III, Ph.D.
Senior Research Director and Director of Survey Methodology
Sampling Statistics Group
GfK CRNA

Beth Jaworski, Ph.D.
Research Analyst
Government and Academic Research
GfK CRNA
Internet adoption, 1995-2011

% of American adults (age 18+) who use the internet, over time. As of August 2011, 78% of adults use the internet.

More: http://pewinternet.org/Trend-Data/Internet-Adoption.aspx
Some subgroups of the population remain less likely to be “online” than others.

Source: Pew Research Center’s Internet & American Life Project’s August Tracking Survey, 2011
Advantages of Web Surveys

Despite the “Digital Divide”, the use of the Internet for survey data collection grows exponentially because…

<table>
<thead>
<tr>
<th></th>
<th>Web</th>
<th>In-Person/Face-to-Face</th>
<th>Mail</th>
<th>Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>Cheapest</td>
<td>Costly</td>
<td>Cheap</td>
<td>Moderate</td>
</tr>
<tr>
<td>Speed</td>
<td>Fastest</td>
<td>Slow</td>
<td>Moderate</td>
<td>Fast</td>
</tr>
<tr>
<td>Social Desirability</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Survey Complexity</td>
<td>High</td>
<td>High</td>
<td>Poor</td>
<td>Good</td>
</tr>
<tr>
<td>Chance of Interviewer Bias</td>
<td>None</td>
<td>High</td>
<td>None</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Source: Duke Initiative on Survey Methodology (http://www.dism.ssri.duke.edu/)
Motivating Question

Has Internet penetration reached a point where the “online” population can adequately represent the U.S. general population?

OR

Given the persistence of the “Digital Divide”, should survey researchers still be concerned about sampling coverage issues?
Methodology
Probability-based ABS recruitment
Recruitment takes place throughout the year
Representative of U.S. adults

Includes:
- Adults with no Internet access (22% of adults)
  - KP provides netbook and free ISP
- Cell phone only (30% of adults and growing)
- Spanish-language
- Extensive profile data maintained on each member
  - demographics, attitudes, behaviors, health, media usage, etc.

Samples from the panel are assigned to projects
  - e-mail invitations and a link to the online survey questionnaire

60,000+ members
Methodology

- Both weighted to the same demographic benchmarks
  (age, race/ethnicity, gender, Census region, Internet access, income, marital status, language and education)
- Bivariate analyses
  (t-tests, chi-squares)
- Examine 15-25 variables each across multiple topics using KnowledgePanel® profile data
  (demos, public affairs, retail, technology, health)
Results
Results: Demographics

Estimates for 5 out of 15 demographic variables not used for weighting were statistically different between the two samples.

Absolute Difference
(Average = 1.7 percentage points)
Results: Public Affairs

Estimates for 7 out of 25 public affairs variables were statistically different between the two samples.

Absolute Difference
(Average = 1.4 percentage points)
Results: Health

Estimates for 9 out of 25 health variables were statistically different between the two samples.

Absolute Difference
(Average = 2.1 percentage points)
Results: Retail

Estimates for 10 out of 25 retail variables were statistically different between the two samples.

Absolute Difference
(Average = 2.6 percentage points)
Results: Technology

Estimates for 11 out of 15 technology variables were statistically different between the two samples.

Absolute Difference
(Average = 3.4 percentage points)
Results: Relationship Between Variables

The inclusion of Non-Internet Households impacts the relationships between variables as well.
**Results: Have Things Changed Over Time?**

Differences between the offline and online populations persist over time.

<table>
<thead>
<tr>
<th></th>
<th>2008*</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recycled Newspapers in Last 12 Months</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Recycled Plastic in Last 12 Months</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Actively Participate in a Neighborhood or Community Group</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Gun Ownership</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

*Zhang, Chan, Mario Callegaro, Melanie Thomas, and Charles DiSogra. 2009. *Do We Hear Different Voices?: Investigating the Differences Between Internet and Non-Internet Users on Attitudes and Behaviors*. Presented at the 2009 Annual Conference of the American Association for Public Opinion Research.*
Conclusion

Online researchers need as clear a picture as possible of the entire U.S. population for:

- Estimating true incidence levels
- Sizing markets and opportunities
- Obtaining publishable findings for peer-reviewed journals

Despite growth in Internet penetration over the years, excluding non-Internet households can still lead to over- or under-estimations for individual variables and change the magnitude of the correlations between variables.