

# A New Online Sample Frame Concept: Passive Data Collection

Presented by:  
Elaine Trimarchi, Mktg, Inc.

Presented at:  
Federal Committee on Statistical Methodology  
Washington, D.C.  
November 6, 2013

Authors:  
Steve Gittelman, Ph. D.  
Phil Garland, Ph. D.

# The State of Survey Research

- \* Government surveys (like the BRFSS) have taken place by phone interviewing for decades
- \* However, phone costs have risen as response rates are dropping. Even harder when attempting to reach cell phones
- \* Some agencies have begun to investigate online research as an alternative
- \* In those circumstances where there are no lists available, commercial online panels are an alternative.

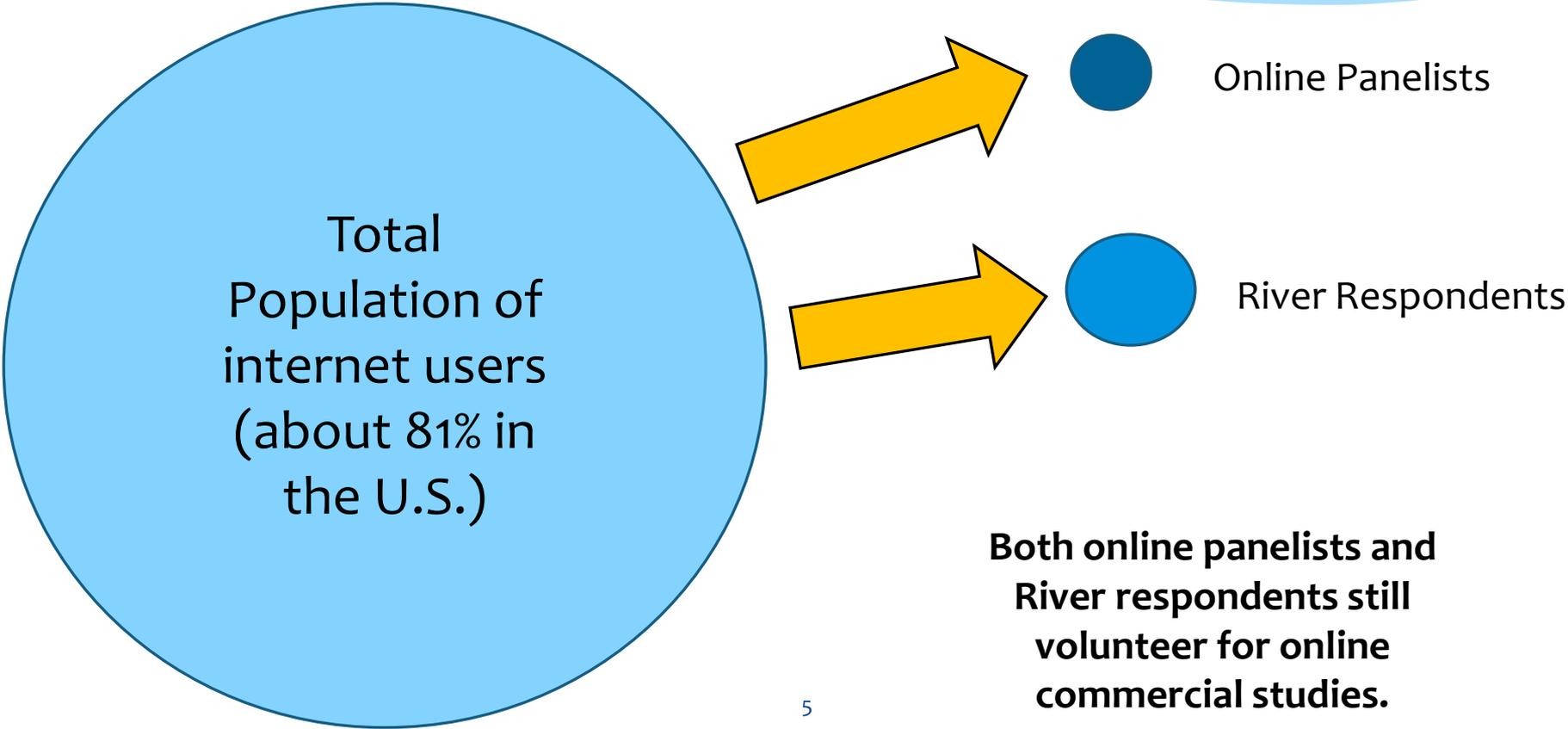
# The State of Online Sample

- \* Commercial online panelists consist of respondents who are often attracted to pop-up invites.
- \* They join a panel, are then sent to a survey and are given an incentive for their time.
- \* At one time, it was estimated that there were 7 million unique online panel respondents. We don't know what that number currently is but we believe it has dropped dramatically.

# The State of online sample—Cont'd

- \* River sampling is another method of obtaining respondents.
- \* Respondents are solicited to do surveys through a stream of pop-up invites.
- \* These respondents agree to do a survey but they are different from online panelists in that they do not agree to be on a panel.

# Online Sourcing



# Routers

- \* A router is an automated system to direct this traffic.
- \* A router has a large number of studies in a queue waiting for respondents to be sent to them.
- \* Respondents are directed to studies based upon their qualifications and the needs of the study.
- \* It is the general school of thought that the broader the selection of sample sources deployed, the closer to a “representative” sample achieved.

# So What's the problem?

- \* The amount of traffic from online panel and River sources is not enough.
- \* There are feasibility issues especially in smaller markets.
- \* We are finding it increasingly difficult to use just one panel source to complete a study.

# A Study in Georgia

- \* As an example of feasibility issues, we conducted a study in the state of Georgia in January-March of 2013.
- \* We collected 795 demographically balanced surveys using questions from the BRFSS.
- \* A full set of demographic quotas were applied – sex, age, income, ethnicity, and education.
- \* It required ~40 days and the resources of 7 panels to complete the study.

# Alternatives to the traditional online Sample?

- \* We need a new source of sample.
- \* The traditional way is becoming harder and harder to complete our surveys.

We need to look for alternative ways to find respondents.  
*Our Concept is called “Grassroots sampling”.*

# Grassroots sampling

- \* Working definition:
  - \* Generating a sample from non-commercial traffic.
  - \* Without an invitation to join or participate in an online panel.
  - \* No financial incentive.
  - \* Traffic is diverse in nature and is driven to a surveying site with the sole motive of offering information.
  - \* They are driven to doing a survey to support their community or organization .

# Proof of Concept

- \* As proof of concept, we did a test using Survey Monkey.
- \* Survey Monkey is a do-it-yourself tool of which the users of their tool are very diverse.
- \* They see 30 million unique respondents per month
- \* They are not panelized and are not incentivized.

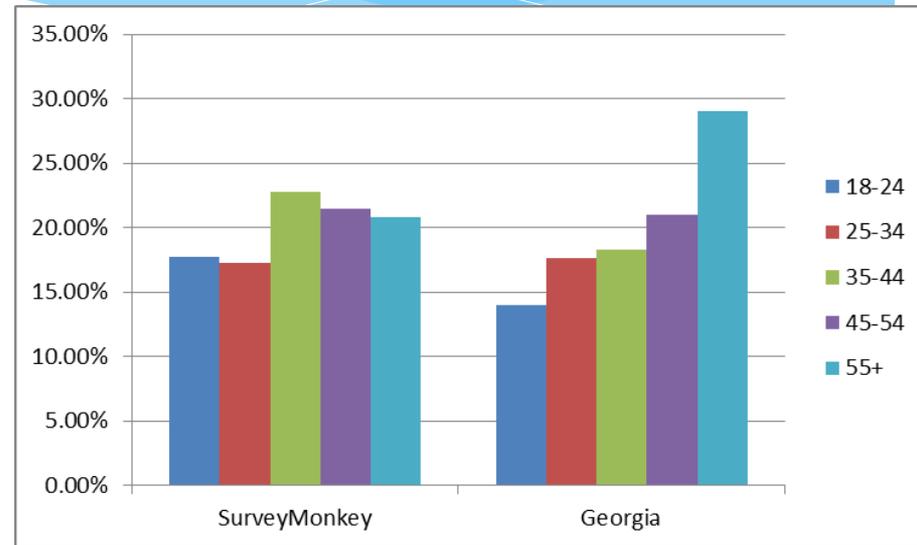
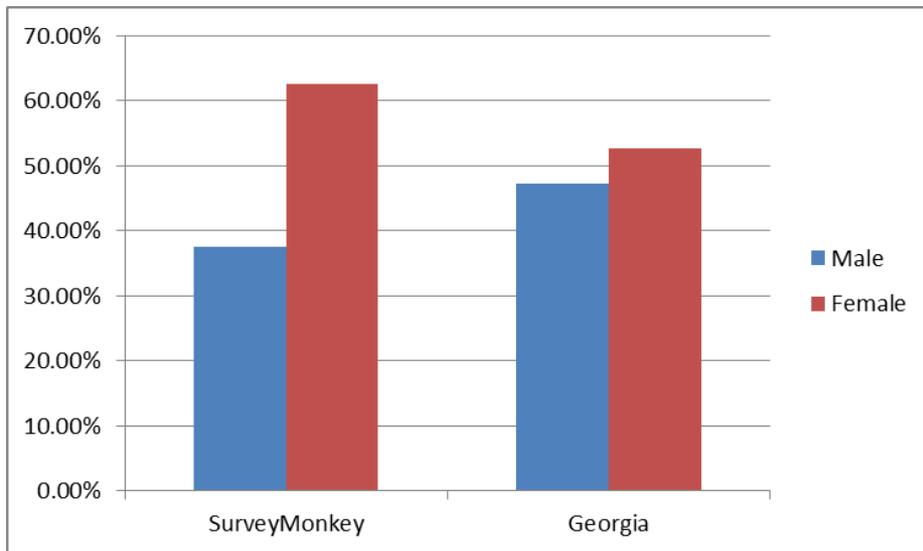
*These respondents are different from the normal commercial pool and are perhaps behaviorally closer to a more representative selection of respondents. They fit the “grassroots” type of respondents.*

# Methods

- \* 9,469 completes in Georgia.
- \* Data was collected over ten days.
- \* 9 questions taken from the BRFSS
- \* We added an end page to all surveys being conducted by Survey Monkey during that ten day time period inviting respondents to participate in a second survey on health.
- \* No demographics quotas.
- \* Weighted to census targets

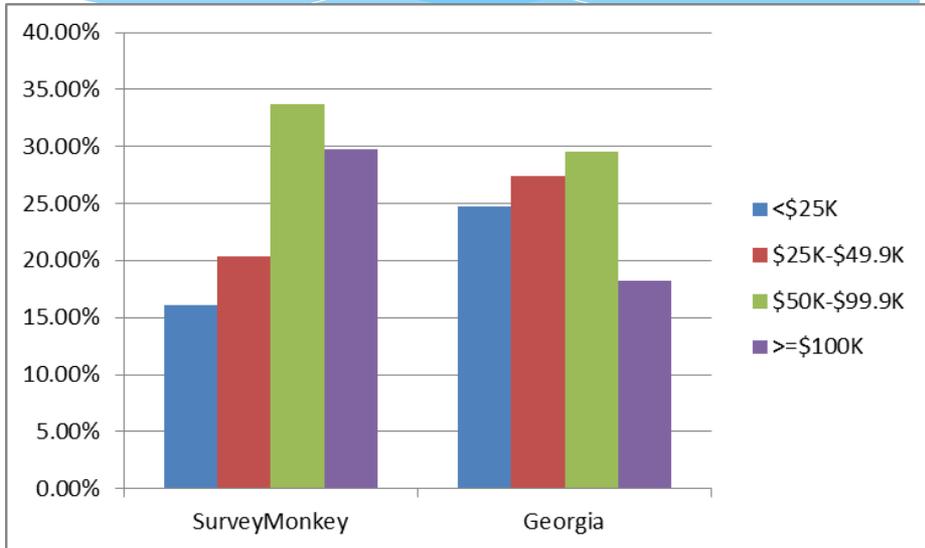
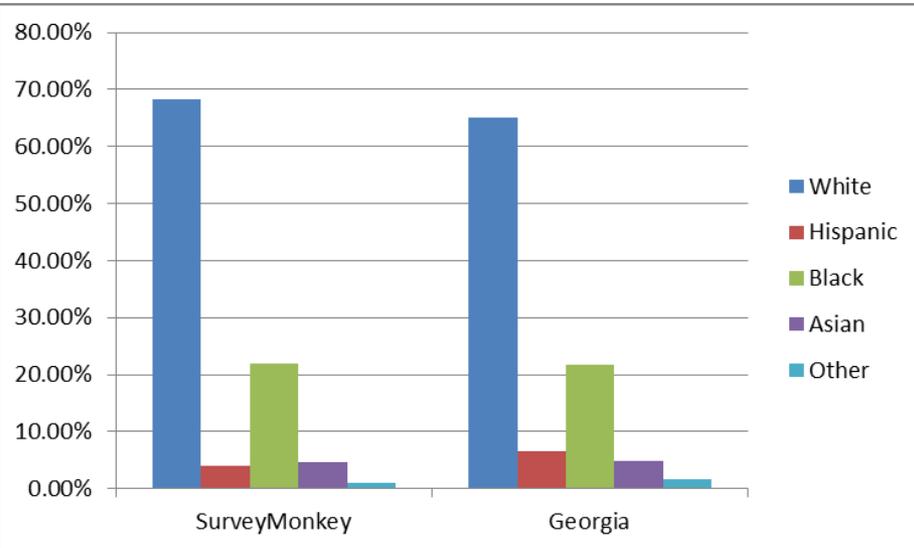
Compared against 2010 BRFSS data as well as the data collected in Georgia using online panels in early (January-March) 2013.

# Results – Demographics I



SurveyMonkey provided a somewhat demographically unbalanced sample compared to census-based targets.<sup>13</sup>

# Results – Demographics II



Despite some shortfalls, SurveyMonkey provided enough completes in all demographics for the purposes of post-stratification weighting.

# Results

	Online Panel	BRFSS	SurveyMonkey
In general, would you say that your health is... (Poor or Fair)	13.29%	17.27%	12.48%
Was there a time in the past 12 months you did not receive medical treatment due to cost?	26.78%	18.39%	21.55%
Do you smoke cigarettes?	22.33%	17.0%	14.80%
During the past month, other than your regular job, did you participate in any physical activities or exercises?	63.26%	74.21%	80.03%
About how long has it been since you last visited a doctor for a routine checkup? (Within the past year)	57.70%	74.90%	68.53%
How many days in the past month was your physical health not good?	4.77	3.86	3.81
Do you have any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare?	79.79%	82.74%	80.44%
During the past 30 days, have you had at least one drink of any alcoholic beverage such as beer, wine, a malt beverage, or liquor?	48.0%	47.59%	54.63%
About how long has it been since you last visited a dentist or dental clinic for any reason? (Past Year)	-	68.2%	61.3%

# The BEST Sources

- \* The Government has a unique resource to provide grassroots type of sample.
- \* Government websites, at all levels, represent a unique opportunity. They receive high volume, and diverse traffic.
- \* Traffic to government websites probably offers a more representative sample of the online population than online panels can provide
- \* This traffic has sufficient volume to complete studies in far smaller areas than is possible with online commercial research today
- \* The opportunity to be surveyed could be offered when potential respondents complete their intended tasks at these sites.

Examples: DMV, IRS, CDC, National Parks , all the way down to the local town hall , library and the school district., to name a few.

# A Plan to Consider

- \* Consider that traffic from government websites can be driven to a router .
- \* This router acts as the traffic manager for government research.
- \* The router will assure the proper distribution of respondents to fulfill the final sample.

A commercial panel partner could be used to back fill hard to obtain respondents.

# Conclusions

- \* The cost of phone sampling is very high.
- \* Online panels and river sampling are more economical but....
- \* Panel resources may not be sufficient, especially for large government studies.
- \* Routers are being used to direct this traffic with the hopes to maximize their yield.
  
- \* We need a new source of sample traffic
- \* Grassroots sampling approach brings a broad spectrum of non-panelized, non incentivized respondents.
- \* “Grassroots” recruitment, tapping community institutions may be an alternative sourcing.
- \* Preliminary results showed that “Grassroots” sources appear to have provided greater accuracy than panel resources.



# QUESTIONS?

Elaine Trimarchi

Mktg, Inc.

631-277-7000

[elaine@mktginc.com](mailto:elaine@mktginc.com)