

SAMPLE SURVEYS BASED ON INTERNET PANELS: 8 YEARS OF LEARNING

**Dr Julian Baim, Dr Michal Galin, Dr Martin Frankel, Risa Becker, Joe Agresti
Mediamark Research & Intelligence**

Using convenience, opt-in Internet panels as sampling frames has become virtually commonplace in today's survey research world. The movement towards Web-based convenience panels is inevitable, given the relatively low cost of these samples, the shorter time frames for completing surveys and the flexibility provided by the Internet for conducting complex surveys. As with every dramatic change in sample selection and/or interviewing mode, concern arises among survey practitioners about the impact potentially radical changes have on survey results from previously employed other sampling procedures. This unease is especially felt among magazine researchers because print audience ratings services are the cornerstone of the buying and planning processes. It is equally disconcerting because many of the country-specific print rating services still employ area probability, in-person surveys with an effort to ensure every eligible respondent has a known probability of being selected.

Over the past 8 years, Mediamark Research & Intelligence (MRI) has explored the impact of using convenience Internet panels on audience ratings and has even utilized these studies in providing more granular information about magazine reading. We have approached Internet panels with a substantial degree of trepidation since we are very cognizant of the potential biases associated with these sampling frames. Beyond analyzing potential uses of Internet panels for magazine research, we have also examined the relative performances of different panels and developed some clear guidelines about the uses and misuses of these sampling frames. This paper discusses the insights gleaned from conducting almost 750,000 surveys on the Internet using convenience panels and tries to offer the proper context in which magazines can use convenience panels for very specific purposes. Although the findings reflect our experience with opt-in Internet panels in the United States, they address issues faced by researchers in many other countries.

BACKGROUND

In the spring of 2002, MRI conducted a series of tests exploring the use of convenience, opt-in Internet panels to measure magazine average-issue audiences (Frankel et al., 2003). While the findings (mentioned briefly below) strongly concluded convenience Internet panel biases precluded using these sample frames as the basis of a print ratings system, MRI continued its investigation into other uses of Internet panels for magazine research. We reported our findings on studying possible title confusion at the 2005 Worldwide Readership Symposium (Baim et al, 2005, Frankel et al, 2005) and shared our initiative on measuring issue-specific readership at the 2007 Worldwide Readership Symposium (Frankel et al, , 2007; Baim et al., 2007). Since then, we have continued the Issue-Specific Measurement service and have used convenience, opt-in Internet panels for our AdMeasures service. Our ongoing efforts also led to using more than one company's Internet panel and thereby provided insights into comparable performances between panels. In the process, we have continued to examine how best to utilize Internet panels and have drawn conclusions about these panels regarding:

- Demographic representation
- Comparable absolute audience estimates
- Using post-stratification or propensity weights to adjust for biases
- Comparable relative audience estimates

DEMOGRAPHIC REPRESENTATION

In 2003, we depicted the stark differences in demographic profile between the U.S. Bureau of the Census estimates and two separate Internet panel sources from one company. Six years later, we update these comparisons using current information from two different Internet panel sources in our Issue-Specific study. (We have omitted the names of the companies for confidentiality purposes.)

The demographic comparisons (Tables 1a and 1b) are based on over 125,000 respondents from each of the two panels. The samples are drawn on a systematic, random basis from the respective panel sample frames. With the exception of identifying 18-24 year-old panel members for special incentive treatment, we have made no effort to oversample any demographic cohort. The respondents represent aggregated weekly samples of 2,500 each from the two panels over a one-year period, with all invitations and surveying conducted simultaneously; there are no differences in sampling instructions. It is somewhat surprising that, even with the passage of six years, these panels still have substantial differences in demographic composition compared to Census data. Regardless of gender, younger adults (18-24), older adults (65+), African-Americans, Hispanics, the less affluent and those with lower levels of educational attainment are substantially underrepresented in both panels. In some cases (young adults, Hispanics, less than high school graduates), panel compositions are less than 50% of comparable Census figures. For one of the two panels, the less than high school proportion is virtually negligible. Even if the survey design called for quota sampling to match Census demographic compositions, it would be difficult to assert these demographic cohorts in the panel sample frames can somehow be representative subsets of that same group in the national population. (As is discussed later, we still see and make valuable use of Internet panels in magazine research.)

A more dramatic finding from Tables 1a and 1b are the substantial differences in demographic profiles between the Internet panels! Panel One members are significantly better educated (there are twice as many college graduates in Panel One than in Panel Two), more affluent (Panel One has thrice the number of members living in households with annual incomes of 100K+) and more likely to be employed. (If Panel One's employment proportions reflected the national population, we would be forced to conclude the recession ended some time ago!) We can only conjecture about the reasons for differences in composition: the two companies probably have clear distinctions in their website recruitment procedures and may also have different perspectives on trying to match certain Census proportions. Regardless of the explanations, these two panels do not resemble each other. Although many have contended that panels have considerable overlapping membership, these figures suggest these two particular panels are enrolling members from different subsets of the Internet population.

Even beyond these dramatic population composition differences between opt-in Internet panels and the Census estimates is the heavy Internet use of panel members compared to the behavior for those adults who use the Internet in the past month¹. Table 2 shows the comparison between the frequencies of Internet usage from each of the panels, respectively, and the past month Internet-user population (measured in MRI's National study). Over half (55.82%) of Panel One members use the Internet more than five times per day; the comparable estimate for all past month Internet users in MRI's National study is 37.71%. While Panel Two's most frequent user proportion is virtually equal to MRI's national estimate for the same category, its second most frequent usage category is approximately 60% higher than national estimates.

¹ Past month Internet user estimates are taken from MRI's Spring 2009 National study

Table 1a: Demographic Comparisons Between Convenience Panels and Bureau of the Census Estimates: Males**Table 1 a**

Demographic	Panel One	Panel Two	U.S. Census
Male			
Age	%	%	%
18-24	5.90	4.05	13.29
25-34	25.64	11.94	18.57
35-44	22.58	17.34	19.42
45-54	20.33	24.61	19.76
55-64	18.15	25.19	14.56
65+	7.40	16.88	14.41
Marital Status			
Single	25.29	21.58	29.01
Married	64.11	57.00	57.56
Other	10.61	21.42	13.42
Race			
White Only	85.33	86.01	75.12
Black Only	2.69	6.66	10.59
Other	11.98	7.32	14.29
Hispanic			
Yes	5.79	4.90	14.24
No	94.21	95.10	85.76
Education			
College grad or more	60.95	34.50	26.85
Some college	31.20	42.97	26.02
High School Grad	7.27	20.04	31.46
Less than HS Grad	0.58	2.49	15.66
HHI			
< 15K	2.58	8.02	6.98
15-24.9	3.85	11.31	8.36
25-49.9	17.19	31.85	23.80
50-74.9	23.06	22.88	20.35
75-99.9	20.61	13.30	14.32
100+	32.73	12.66	26.19
Employed			
Yes	82.69	58.94	69.25
No	17.31	41.06	30.75

Table 1b: Demographic Comparisons Between Convenience Panels and Bureau of the Census Estimates: Females

Demographic	Panel One	Panel Two	U.S. Census
Females			
Age	%	%	%
18-24	7.54	4.53	12.13
25-34	24.86	16.41	17.24
35-44	24.61	21.06	18.57
45-54	24.05	27.58	19.31
55-64	14.83	20.89	14.71
65+	4.12	9.53	18.05
Marital Status			
Single	22.59	16.78	22.37
Married	61.81	56.86	53.62
Other	15.60	26.37	24.01
Race			
White Only	86.05	87.56	76.09
Black Only	4.26	6.65	11.88
Other	9.70	5.80	12.04
Hispanic			
Yes	5.28	4.28	12.64
No	94.72	95.72	87.36
Education			
College grad or more	54.12	27.16	26.26
Some college	35.61	45.97	28.96
High School Grad	9.80	24.46	30.92
Less than HS Grad	0.47	2.41	13.86
HHI			
< 15K	3.11	9.50	11.29
15-24.9	4.69	13.00	10.51
25-49.9	21.59	32.49	23.79
50-74.9	24.44	22.98	18.61
75-99.9	19.27	11.86	12.77
100+	26.90	10.17	23.03
Employed			
Yes	74.85	54.39	57.24
No	25.15	45.61	42.76

Table 2
Frequency of Internet Usage:
Internet Panels vs. MRI National Estimates

Frequency of Internet Use	Panel One (%)	Panel Two (%)	Adult Population Having Access to the Internet (%)
5 or More Times a Day	55.82	38.14	37.71
2-4 Times a Day	31.99	39.65	24.77
Once a Day	8.12	14.42	15.13
3-6 Times a Week	2.57	5.10	10.52
1-2 Times a Week	0.72	1.36	7.15
Less Than Once A week	0.78	1.31	4.71

POST-STRATIFICATION: COMPENSATING FOR BIAS TO ESTIMATE AVERAGE-ISSUE AUDIENCES

The panel demographic and usage comparisons are certainly troubling for those who contend that Internet panels should be a microcosm of the national population and certainly indicate biases in panel recruitment procedures. At the same time, a common practice for removing bias from skewed sample frames or from probability samples with low or differential response rates is to post-stratify (i.e., weight) the demographic data to conform to nationally-accepted estimates (e.g., Census estimates). In an effort to understand whether post-stratification brings average-issue readerships estimates from these two panels close together, MRI weighted each of the panel samples to Census figures, using all the demographic variables shown in Tables 1a and 1b.²

Table 3 shows the relationship between average-issue audience estimates from Panel One compared to those in Panel Two. (We calculated average-issue audiences using “read or looked into any issue” responses over an entire year from MRI’s Issue Specific Study. Across 228 magazines, Panel One’s audiences were, on average, 15.4% higher than those in Panel Two. Audiences in Panel One were higher in more than two-thirds of the cases (157 out of 228). For those who contend that post-stratification can remove biases from Internet panels, these findings pose a potentially irresolvable problem: If post-stratification removes biases, why don’t audience estimates from the two Internet panels converge? There are clearly other factors, besides demographics, impacting magazine audience estimates in the respective Internet Panels.

Table 3: Comparison of Average –Issue Audiences between Panel One and Panel Two

Ratio of Panel One Audiences to Panel Two Audiences	No. Of Magazines	Percentage of Cases
200%+	6	2.6%
150-199%	14	6.1%
125-149%	51	22.4%
100-124%	86	37.7%
75-99%	59	25.9%
<75%	12	5.3%

Magazine researchers confront an even more disconcerting question: how can we use Internet panels to measure **absolute average-audience levels** if different panels produce such divergent estimates?

This issue becomes more acute when we examine comparable levels for four different magazine groups. Tables 3-6 show the relationship between average-issue audiences obtained from the two Internet panels, respectively, for four magazine groups. For every magazine in each of the four groups, average-issue audiences are consistently and substantially higher for one panel source compared to the other. In the first two cases (airline and travel magazines), audiences in Panel One are approximately twice as high on average as comparable estimates from Panel Two; in the latter two cases (motor enthusiasts and outdoor magazines), Panel One average-issue audiences are only 75% of Panel Two projections. There may be compelling explanations (e.g., recruitment procedures, incentive strategy) why we observe such disparate results from these two panels. It is clear, however, that basic demographic post-stratification cannot reconcile procedural differences in the way each Panel generates its membership.

² The 2003 WRRS paper already indicated that Internet panels did not produce comparable audience estimates to those found in MRI’s National study. While neither of the present panels produced any closer results to MRI estimates after post-stratification, our focus in this paper was to understand whether the two panels produced similar average-issue audiences to each other.

Table 4: Comparison of Average –Issue Audiences between Panel One and Panel Two: Airline Magazines

Airline Magazines	Ratio of Panel One Audiences to Panel Two Audiences
Magazine 1	288.70%
Magazine 2	274.07%
Magazine 3	225.63%
Magazine 4	214.25%
Magazine 5	201.05%
Magazine 6	151.78%
Average	225.91%

Table 5: Comparison of Average –Issue Audiences between Panel One and Panel Two: Travel Magazines

Magazine Type	Ratio of Panel One Audiences to Panel Two Audiences
Magazine 1	190.01%
Magazine 2	189.16%
Magazine 3	173.66%
Magazine 4	170.66%
Average	180.87%

Table 6: Comparison of Average –Issue Audiences between Panel One and Panel Two: Outdoor

Magazine Type	Ratio of Panel One Audiences to Panel Two Audiences
Magazine 1	80.53%
Magazine 2	78.96%
Magazine 3	78.72%
Magazine 4	68.38%
Magazine 5	65.51%
Magazine 6	62.95%
Magazine 7	61.12%
Magazine 8	57.32%
Average	69.19%

Table 7: Comparison of Average –Issue Audiences between Panel One and Panel Two: Motor Car Enthusiasts

Magazine Type	Ratio of Panel One Audiences to Panel Two Audiences
Magazine 1	87.42%
Magazine 2	86.95%
Magazine 3	85.11%
Magazine 4	83.65%
Magazine 5	81.89%
Magazine 6	77.25%
Magazine 7	75.59%
Magazine 8	72.83%
Magazine 9	69.98%
Magazine 10	57.00%
Magazine 11	52.28%
Average	75.45%

A substantial volume of literature has been produced about the use of propensity weights to correct for frame or respondent bias in Internet studies (Schonlau et al., 2006; Smith et al., 2001 Taylor, 2000). Many of the proponents of this approach were concerned with analyzing political variables or, in other cases, single variables. On the other hand, magazine audience research is concerned with producing estimates for two hundred or more publications, including subsets with unique or niche appeal. For a similar type of study of health and financial assets (Schonlau et al., 2009), the authors found propensity weighting works for some variables and not for others:

We find that the corrections generally work well for health variables, but not for past health behavior (smoking and drinking) or, particularly, financial assets.

Given the substantial diversity of magazine publications, it is difficult to envision that any single set of propensity variables (if they could even be identified) would perform any better in a magazine ratings study than it did in the Health and Retirement Study.

The above analysis strongly suggests that demographic and behavioral biases in Panel membership create substantial difficulties in using opt-in, convenience panels, alone, to estimate **absolute** incidence estimates, including print ratings measures. Our conclusion is echoed in a recently released study of using Internet panels to provide absolute estimates (Langer, 2009). Gary Langer's assessment of the study concluded:

“Non-probability research often is done to assess relationships between variables – but not to measure the magnitude of such associations, much less population values, such as how many people think or do X, Y or Z. If that is a researcher's aim, Yeager and Krosnick say, “non-probability sample surveys appear to be considerably less suited to that goal than probability sample surveys.”

Using Opt-In Panels for Analysis of Relative Audience Estimates:

While there are evident issues with using opt-in, convenience panels to provide absolute estimates of behavior, we have adopted their use in providing assessment of **relative** issue-by-issue performance. All of the above-mentioned benefits (e.g., cost, timeliness, flexibility) of using Internet panels commend careful consideration of these frames for valuable research purposes. In MRI's Issue Specific Study, we make use of the **relative** ratings generated for a particular issue against a baseline of all issues of the same magazine over a year's period. However, the underlying or implicit justification for using these panels for this particular purpose is based on the belief that Internet panels provide valid information about relative patterns of behavior in the general population.

MRI evaluated this assumption in two different analyses:

- A comparison of relative audience changes between the Internet panels and MRI's National study
- A comparison of relative issue changes between Panel One and Panel Two samples

Our ability to apply issue-to-issue changes found in the Internet-panel frame Issue Specific study to audiences estimates from our National (area probability) study is predicated on the (tacit) assumption that there is a reasonably strong correspondence between respective magazine audience changes in the two studies. In order to examine this hypothesis, MRI compared the direction of audience changes for each magazine in the Issue Specific study to the similar measure of change in the National study.³ Of 179 comparisons, we found that 68 % of the audience changes for the same magazine from the two respective studies were in the same direction. Using a standard sign test, the proportion of agreement is statistically significant at the .001 level/ Even more compelling is the finding that the agreement rate was 76% among magazines with relative audience changes in the National study of 10% or more.

Having established a strong justification for applying relative findings from the opt-in convenience Internet panels to the National study, MRI further examined the question whether the two different Internet panels would provide consistent information about **relative** audience changes. We have documented above that the two Internet panels failed to generate equivalent absolute ratings. Despite this incompatibility, it was still possible that issue-specific variations within a magazine could be consistent, albeit at very difficult overall levels.

Since MRI measured issue -specific audiences for common magazines in each of the two Internet panels separately (before aggregating the results), it was possible to compare relative audiences (measured as indices) generated for each issue of a magazine in the two panels. Prior to making these comparisons, MRI used the same one-year sets of data from the above analyses and standardized the demographic weights for the panels to ensure consistency in profile. We then ran a series of correlations, comparing indices for each issue of a specific magazine. For monthlies, MRI's comparisons consisted of a set of 12 matched issues; for weeklies, MRI correlated approximately 50 matched issues for each title.

Tables 8a and 8b show the distribution of correlation coefficients for the 181 magazines released in our Issue-Specific study. (Men and women results are shown separately.) For either gender, the correlation coefficients are remarkably high. For men, approximately 40% of correlations are .8 or higher; the comparable figure for women is almost 50%. The median coefficients for men and women are .74629 and .79834, respectively. Instances where the correlations are low or negative are generally cases based on the off-sex readership of particular magazines, where readership levels are usually low and highly unstable.

The consistently strong correspondence between issue-specific indices generated from surveys using two separate opt-in Internet panels reinforced our confidence in using these samples to provide measures of relative issue-to-issue audiences. Even though the two samples generated substantially different absolute audience estimates, they showed very similar issue-to-issue variation within a magazine. The analysis underlines the distinction between using convenience Internet panels to provide absolute magazine readership incidence levels and using these samples to inform the industry about relative magazine performance.⁴

Table 8a. Correlation of Issue-Specific Audiences From Two Opt-In Panels

Base: Males

Range of Coefficients	No. of Cases	Percent of cases
.9 or greater	37	20.44%
.8 - .89	36	19.89%
.7 - .79	33	18.23%
.6 - .69	19	10.50%
.5 - .59	16	8.84%
.2 - .49	22	12.15%
0 - .19	9	4.97%
-1 - -.01	9	4.97%
Median Correlation	0..74629	100.00%

³ MRI averaged the last six months of issue-specific audiences for each magazine against the previous year's worth of aggregated issue-specific data for the same magazine and noted whether the audiences over that period of time had increased or decreased. We conducted a similar analysis of the past six months of individual magazine audience data from the National study compared to the previous year's audience estimates.

⁴ We acknowledge the potential value of integrating opt-in panel absolute results with findings from high quality, strict probability samples, an approach employed in MRI's AdMeasures study

Table 8b. Correlation of Issue-Specific Audiences From Two Opt-In Panels

Base: Females

Range of Coefficients	No. of Cases	Percent of cases
.9 or greater	59	32.60%
.8- .89	30	16.57%
.7- .79	28	15.47%
.6- .69	13	7.18%
.5- .59	16	8.84%
.2- .49	15	8.29%
0- .19	9	4.97%
-1 - -.01	11	6.08%
Median Correlation	0.79834	100.00%

Conclusions

Over the past 8 years, MRI has utilized opt-in convenience Internet panels for magazine research. Our studies have ranged from examining title confusion under test-control conditions to applying variation between issue audiences to magazine ratings from our National study. We have acknowledged concerns regarding the biases inherent in recruitment strategies of Internet panels, but we have also recognized the desirability to incorporate their use in the proper context for magazine research. Our studies, which have surveyed approximately 750,000 panel members, have provided substantial information about how best to utilize convenience Internet panels. These data show:

- Panels remain demographically unrepresentative of the U.S. adult population
- Panel demographics are not consistent between frames.
- Post-stratification, alone, cannot reconcile differences in measures of absolute incidence levels (i.e., magazine audience ratings).
- There are unique differences in niche magazine ratings between panels that propensity weighting is unlikely to address effectively
- Notwithstanding the difficulties in measuring absolute levels, Internet panels are valuable in reflecting overall relative trends in magazine readership
- Magazine issue-to-issue variations within one Internet panel are mirrored in a second Internet panel

We do recognize that panel companies make continuous efforts to redress some of the deficiencies discussed in this paper and we also believe that thoughtful, model-based adjustments can be used effectively (but selectively) to reap the benefits of these sampling frames. It is also evident that the appeal of using Internet panels for survey research will continue to grow and it is foolhardy to ignore that trend. At the same time, careful, continuous thought and analysis must be employed before we move full-board to Internet panel frames.

References

Baim, Julian, Martin Frankel, Michal Galin, Joseph Agresti and Kerry Zarnitz. "Measuring Issue Specific Audiences." Paper presented at the Worldwide readership Symposium, Vienna, Austria: 2007.

Baim, Julian, Michal Galin, Martin Frankel and Scott McDonald. "Title Confusion: The Impact of Response Error on Competitive Pairs." Paper presented at the Worldwide readership Symposium, Prague, Czech Republic: 2005.

Frankel, Martin, Julian Baim, Michal Galin and Joseph Agresti. "Issue Specific Estimation – Mathematical and Statistical Issues, procedures and Models." Paper presented at the Worldwide readership Symposium, Vienna, Austria: 2007.

Frankel, Martin, Julian Baim, Michal Galin and Michelle Leonard. "Measurement of Magazine Readership Via the Internet." Paper presented at the Worldwide Readership Research Symposium, Cambridge, MA: 2003

Frankel, Martin, Risa Becker, Julian Baim, Michal Galin and Scott McDonald. "Dazed and Confused: The Characteristics and Behavior of Title Confused Readers." Paper presented at the Worldwide Readership Research Symposium, Prague, Czech Republic: 2005.

Langer, Gary. "Study Finds Trouble for Opt-In Internet Surveys." Sept. 1, 2009
<http://blogs.abcnews.com/thenumbers/2009/09/study-finds-trouble-for-internet-surveys.html>

Schonlau, Matthias, Arthur van Soest, ASrie Kapteyn and Mick Couper. 2009. "Selection Bias in Web Surveys and the Use of propensity Scores." *Sociological Methods & Research* 37: 291-318.

Smith, Philip J., J.N.K. Rao, Michael P. Battaglia, Dani Daniels, and Trena Ezzati-Rice.. 2001. "Compensating for Provider Nonresponse Using Response propensities to Form Adjustment Cells: The National Immunization Survey." *Vital and Health Statistics* 2 (133): 1-17.

Taylor, Humphrey. 2000. "Does Internet research 'Work'? Comparing On-Line Survey results With telephone Surveys." *Journal of Marketing Research* 42:51-64.