



Report for the Associated Press:
Illinois and Georgia Election Studies in November 2014

**Randall K. Thomas, Frances M. Barlas, Linda McPetrie,
Annie Weber, Mansour Fahimi, & Robert Benford**

GfK Custom Research



Overview

Beyond helping news organizations improve their election projections within states, exit polls provide significant insight into voters' political attitudes and reasons for their vote choices. Traditionally, exit polls have selected a representative sample of polling places and conducted in-person interviews among exiting voters, asking for their vote choices, demographics, and attitudes across a number of political and election-related issues. With the rise of early and absentee voting and the decline of participation in polls, exit polls face challenges to both the accuracy and costs of the polls. One response to the rise in early and absentee voting has been to supplement in-person exit polls with data obtained by telephone polling of absentee/early voters, though this increases costs substantially.

As an alternative to these mixed-mode exit polls, we conducted an entirely web-based survey in Georgia and Illinois in November 2014 in the days leading up to the election. We screened for self-identified registered likely voters drawn from two different sample types: 1) a probability-based sample from GfK's KnowledgePanel® (KP) and 2) a non-probability sample (NPS) from various online sources. We compared the actual election outcomes for the Governor and Senate races in those states with the survey results among registered likely voters from both samples. We also compared the survey results to those obtained from National Election Pool (NEP) exit polls, which used both in-person and telephone interviews and were weighted to election outcomes.

In this paper, we summarize the results of this pilot study, which showed that the demographically weighted KP sample showed the least bias of estimates compared to the exit polls and those obtained from a Calibrated combination of the two sample types (KP+NP)



results. We also found similar attitudes among key demographic subgroups within the likely voters for both sample types (KP and Calibrated) to those that were obtained in the non-online exit polls. We conducted an experiment with the likely voter model and considered alternative weighting and data blending strategies to help refine our approach. These results showed that online surveys with alternative sample types can be a viable alternative to traditional exit poll methodology.



Method

This pre-election study was fielded in two states – Georgia and Illinois – with two different sample sources. The first was GfK’s probability-based online panel – the KnowledgePanel® (KP). The second was non-probability online sample (NPS) which blended participants from six different sample sources.

KnowledgePanel is the largest online panel of adults in the US from which probability-based samples can be generated. KP members are primarily recruited using address-based sampling (ABS) methods, including telephone follow-up for refusal conversions. Adults selected to join KP who do not have access to the Internet are provided Internet access and a web-based device at no cost. For this study, the KP sample included all active panel members who were 18 or older and lived in Georgia or Illinois at the time of the study.

The NPS employed a combination of both online panel and river, or intercept, sources of participants who volunteered to participate in surveys in exchange for rewards and sometimes the opportunity to win a prize in a sweepstakes. The sample was selected in accordance with quotas based on sex, age, race, and educational attainment to reflect the population in each state. The benchmarks used for quotas were from the American Community Survey Three Year file 2011-2013.

Table 1 summarizes both the total number of qualified completed interviews by respondents who were 18 years old or older and also the numbers after screening for being registered, likely voters in Georgia and Illinois.



Table 1. Qualified Completed Interviews by State, Voter Status, and Sample Source

State	Sample Source	Likely Voters*	Unlikely Voters	Total Respondents
Georgia	KnowledgePanel	460	194	654
	Non-probability Sample	967	590	1,557
Illinois	KnowledgePanel	751	266	1,017
	Non-probability Sample	1,158	590	1,748

* Likely voters as determined by both Traditional and New Likely Voter models, explained in more detail below.

Survey Fielding. The online pre-election study fielded from October 30 to November 4, 2014 (closing at 8 a.m. Eastern on Election Day). Participants from KP and the non-probability sample sources were invited to the web-based survey with a link containing an embedded unique password to ensure that each survey interview was from a unique participant. After establishing participants' voter registration status, their likelihood to vote, and method of voting, we asked which candidate they would vote for in two separate races – Senator and Governor – based on resident state. For NPS sample, we verified state by asking panelists to provide their ZIP code, and confirmed by asking for state of residence. Following the vote choice section, we asked additional questions about a number of different attitudes that were similarly asked in the NEP exit poll.

Sample Weighting. For pre-election weights, all KP and NPS participants, regardless of voter registration and likelihood to vote, were separately weighted to state-level population benchmarks from the American Community Survey using weighting targets based on age, gender, race, ethnicity, and education. KP and NPS survey data were then combined using an optimal blending process in proportion to their respective effective sample sizes after demographic weighting (Fahimi, 1994) and GfK's calibration methodology that included specific



attitudinal and behavioral questions we have been found to differentiate between probability-based and respondents from NPS (Fahimi et al., 2015). These questions included weekly time spent on the Internet for personal use, number of online surveys completed monthly, average daily duration of television viewing, tendency to be an early adopter of new products, frequency of coupon use when shopping, and number of moves in the past five years.

Election Weighting. In order to enable more appropriate comparisons with the final exit poll results, which are weighted to the final election outcomes, the sample weights computed above were then post-stratified to election outcomes for each sample type (KP or Calibrated KP+NPS) for each election in each state.

Likely Voter Model test. One part of this study was designed to test the traditional AP likely voter model used for polling against a simpler likely voter model, which took less time for respondents to complete. Respondents were randomly assigned to the traditional likely voter model or a shorter, stated intention to vote method.

The traditional model is limited to respondents who are registered to vote. It is based on a complex set of definitions that includes past vote frequency, past voting behavior, whether or not they have already voted, likelihood to vote, interest in news about the election, and knowing where to vote. This model requires eight survey questions based on four different patterns of survey answers to define a likely voter. This model is very similar to what many others in the polling sector use.

The new likely voter model was much simpler. Under this model, likely voters are limited to registered voters and includes those who 1) already voted or say they will definitely vote or 2) say they probably will vote and also indicated that they always or nearly always vote in elections. Estimates of candidate vote percentage show very few significant differences by



sample within the likely voter models. Comparing the models across estimates by sample, 70% of the time the new, stated intention model was closer to the actual results than the traditional model. This suggests that the stated intention model, which utilized fewer questions, may work well as a substitute for the more complex traditional model, which was also more burdensome for respondents to complete. Due to the efficiency of the new likely voter model, future iterations will employ this simpler version.

National Election Pool (NEP) Exit Polls. In-person interviews in both states were conducted on Election Day among randomly selected voters drawn from a stratified probability sample of polling places within each state. The exit poll results were released in stages (calls), varying in both the number of cases and weighting algorithms used (final stage has the results weighted to the actual election outcomes). The exit poll in Georgia included interviews with 3,117 voters, including 559 (17.9%) who voted by absentee ballot or who were early voters and were interviewed by phone prior to Election Day. Phone interviews were conducted between October 24 and November 2 on both landline phones and cellphones drawn from a random-digit dial sample. The exit poll in Illinois included interviews with 1,307 Election Day voters interviewed at one of 25 randomly selected polling places; telephone interviews among early and absentee voters were not conducted in Illinois.

All exit poll results, including candidate estimates, and descriptions of polling methodology were provided to GfK by the AP. We compared candidate estimates obtained in this study with the best ratio estimates from the final exit poll, which were based on both telephone interviews of absentee/early voters, where available, and on all election day voter interviews at sample precincts. Separately, we compared the telephone poll candidate estimates from absentee/early voters with this study's estimates from absentee/early voters. We also



compared crosstabs from this study to the exit poll crosstabs that were weighted to the final election results. Exit poll candidate estimates are derived from a series of statistical models that use current and past results from a random sample of precincts. The models use exit poll and pre-election poll data until the polls close and the poll results can be replaced by actual vote results from the precinct sample. The best ratio estimate is the model that has the smallest standard error on the difference between the top two candidates.

Results

Estimates for Races. We first compared the actual vote outcomes for each race in each state with the NEP exit poll results and with the results from our study among likely voters. Since the results for GfK data were not substantially different between voter models, we pooled likely voters as identified by the traditional likely voter model and the new likely voter model for subsequent analyses. For this analysis we rely on the results from the best ratio estimates from the exit poll.

Table 2 shows the results for both the Georgia and Illinois Senate and Gubernatorial elections. We calculated the average candidate error across races by averaging the absolute difference between actual and study estimate for the primary candidates. Overall, we found that the KP-only results were closest to actual election outcomes. Across all races in both states, the average exit poll candidate error was 4.4%. By comparison, the average candidate error for the KP-only sample was less than the exit poll at 1.8%. The calibrated KP + NPS average error was 2.3%.

Focusing on the state-specific results, for the Georgia exit poll had an average candidate error of 6.0% while the KP-only was lower at 2.3% and the Calibrated KP+NPS was 2.5%. For



the Illinois races, the exit poll had an average candidate error of 2.9%, while KP-only sample had a smaller error of 1.2% and the Calibrated KP+NPS error was 2.1%.

Another way to compare the results is to look at the party spread bias, a calculation that determines average bias toward one party. It is computed using the actual spread (Democrat proportion of the vote minus Republican proportion of the vote) minus the estimated spread (estimated Democrat proportion minus estimated Republican proportion). A negative average value indicates a Democrat bias and a positive average value indicates a Republican bias. The NEP exit poll results had a significantly higher negative average spread bias reflecting a much stronger Democrat bias than either the KP-only or the Calibrated KP+NPS samples, with the KP-only having the least party bias.

Looking at the spread bias by state, the party spread bias was most strong in Georgia for the exit poll results. The Illinois spread bias for the exit poll results, while highest of all Illinois results, was not of the magnitude seen in Georgia. The Calibrated KP+NPS spread bias was higher in Illinois than in Georgia. The spread bias for KP-only appeared to be smaller than both exit poll and Calibrated, and not systematically biased toward one or the other party, similar to typical unbiased sample result fluctuations one would expect based on sample sizes used.



Table 2. Georgia and Illinois Election Results¹

		Actual Vote % ²	Exit Poll - Best Ratio ³	Pre-election KP-only ⁴	Pre-election Calibration (KP+NPS) ⁵
Georgia Senate Candidate					
	Nunn (D)	45.2%	51.4%	44.7%	44.9%
	Perdue (R)	52.9%	46.5%	49.3%	48.3%
	Spread Bias		-12.6%	-3.1%	-4.2%
Georgia Governor Candidate					
	Carter (D)	44.9%	50.5%	42.1%	43.1%
	Deal (R)	52.8%	47.1%	50.4%	49.6%
	Spread Bias		-11.3%	0.5%	-1.4%
Illinois Senate Candidate					
	Durbin (D)	53.5%	57.7%	52.7%	54.8%
	Oberweis (R)	42.7%	42.0%	43.9%	39.7%
	Spread Bias		-4.8%	2.0%	-4.2%
Illinois Governor Candidate					
	Quinn (D)	46.4%	49.8%	47.5%	47.8%
	Rauner (R)	50.3%	47.1%	48.6%	47.3%
	Spread Bias		-6.6%	-2.8%	-4.4%
Average Absolute Candidate Error			4.4%	1.8%	2.3%
Average Spread Bias			-8.8%	-0.8%	-3.5%

¹ Results are calculated for the two major candidates with minor candidate proportions as part of the denominator. However, undecided is not part of the denominator.

² Actual -- Final, official results

³ Exit Poll -- Best ratio estimate using final exit poll, combined with absentee poll where available

⁴ KP Only -- Estimates from GfK KnowledgePanel

⁵ KP + OP -- Estimates from KnowledgePanel and non-probability sample, blended together using GfK calibration methodology

Early/Absentee Voters in Georgia. Approximately 36.8% of all ballots cast in the Georgia election were early/absentee ballots. A portion of the exit poll data for Georgia was collected by way of a telephone poll in order to capture early and absentee voters. Table 3 compares the results obtained in the actual vote for the two races¹ along with the exit poll results and the KP-only and Calibrated KP+NPS samples. Here again, the KP-only results are closer to actual election outcomes than the NEP exit poll and the online Calibrated solution. The average absolute candidate error for the exit poll early/absentee voters and for the Calibrated KP+NPS



early/absentee voters was equal at 3.8%, while the average absolute candidate error for KP-only was lower at 1.8%. Looking at the spread bias, the NEP exit poll results had a higher average bias than KP-only and the Calibrated KP+NPS samples.

Table 3. Early/Absentee Voter Results

	Actual Vote % among Early/Absentee	Exit Poll - Early/Absentee Voters	Early/Absentee Voters - KP- only	Early/Absentee Voters - Calibration (KP+NPS)
Georgia Senate Candidate				
Nunn (D)	48.4%	53.0%	48.3%	45.4%
Perdue (R)	50.3%	46.0%	48.9%	45.1%
Spread Bias		-8.9%	-1.3%	-2.2%
Georgia Governor Candidate				
Carter (D)	48.0%	51.0%	48.0%	43.3%
Deal (R)	50.3%	47.0%	44.5%	48.0%
Spread Bias		-6.2%	-5.7%	2.4%
Average Absolute Candidate Error		3.8%	1.8%	3.8%
Average Spread Bias		-7.6%	-3.5%	0.1%

Electorate Demographics and Attitudes. While the KP probability sample alone showed a somewhat smaller average candidate error from election outcomes than both the early exit poll (Call 2) and the Calibrated KP+NPS sample, it may be that the calibrated combined sample could be useful for determining the demographic and attitudinal bases for vote choice when compared against the final exit poll results (Call 10). The full results are contained in Appendices B and C when using pre-election weights, while Appendices D and E contains the election-weighted results. For most demographic variables and attitudinal variables, the overall



results of both the KP-alone and Calibrated KP+NPS samples showed similar patterns to what was obtained in the exit polls for all races.

Looking at the election-weighted results showed some demographic differences between the exit polls and our online polls with regard to demographics that were consistent in both the pre-election and election-weighted results – there were more people with a lower education (HS or less) and with a lower income (<\$50K) in the online survey than in the exit polls. These discrepancies may have been due to differences in the sample or weighting, making this an area for further investigation. To look at overall effects for the demographics and attitudes of likely voters we examined the difference between the exit poll results (Call 10 weighted to election outcomes) and the online survey outcomes by computing the absolute difference for each proportion of each level of the demographic and attitudinal variables for each candidate (when the exit poll did not suppress presentation of the results). We averaged the deviations by item and then averaged across all items. We did this for each race within each state and for each sample and weighting combination. Table 4 summarizes the average deviations for these combinations.



Table 4. Comparing Exit Poll Results with Online Survey Voter Demographics and Attitudes

	Pre-election Weighted Results		Results Weighted to Election Outcomes	
	KP-only Sample Average Divergence	Calibrated (KP+NPS) Sample Average Deviation	KP-only Sample Average Divergence	Calibrated (KP+NPS) Sample Average Deviation
GA Senate	7.0%	5.1%	5.9%	3.8%
GA Governor	7.0%	5.0%	5.4%	3.2%
IL Senate	3.6%	3.4%	3.5%	3.0%
IL Governor	4.1%	3.7%	4.1%	3.2%
Average	5.4%	4.3%	4.7%	3.3%

Overall, a review of the tables showed that the story told by the analyses with regard to voter demographics and possible attitudinal motivations by vote choice was fairly consistent across each weighting and sample scenario. The story told under each is relatively consistent with the overall story reflected in the exit poll results. Weighting to election outcomes did reduce the average difference from exit polls for 15 of the 16 race and sample-weighting combinations, all but Illinois Governor results from the KP sample alone which showed the same absolute average deviation after weighting to election outcomes. In addition, calibration lowered the average difference from the obtained results for the exit polls. The combination of calibration and weighting to the election outcomes yielded absolute average deviations of less than 4% in all four races.

Discussion

Overall, for prediction of election outcomes, KnowledgePanel® alone had the lowest average error than exit poll results, even lower than the traditional NEP exit polls. Calibrated



KP+NPS results also out-performed the exit polls, though they had slightly higher errors than the probability-based KP only sample. Further, with appropriate weighting, results for most demographics and attitudinal issues by candidate by race for both states yielded similar results to the exit poll; this is true for both the KnowledgePanel and Calibrated samples. In focusing on the election outcomes, the online probability sample looks to be a viable alternative to in-person exit polls, coming closer to the actual election outcomes than the exit poll estimates.

While the Calibrated solution better approximated the exit poll results for attitudes and demographics than KP sample, the exit poll itself is a survey and we were not able to compare the results to the demographics and attitudes of the full voter population. In addition, exit poll results for both demographic and attitudinal variables could be a result of the sample participating, weighting algorithms used, or mode differences – exit polls are conducted with human interviewers, with either phone or in-person interviewing. All three factors are differences between exit polls and online surveys – sample, weighting, or mode – and could have had some influence on the generally small differences we observed. In spite of these differences, the results of the online survey regarding the demographics and opinions affecting the races in the two states presented a parallel narrative to that obtained by the exit polls, regardless of sample and weighting algorithms used.



References

Fahimi, M. (1994). “*Post-stratification of Pooled Survey Data.*” *Proceedings of the American Statistical Association, Survey Research Methods Section*, Toronto, Canada.

Fahimi, M., Barlas, F. M., Thomas, R. K., & Buttermore, N. (2015). Scientific surveys based on incomplete sampling frames and high rates of nonresponse. *Survey Practice*, 8 (5). ISSN: 2168-0094