



Report for the Associated Press
November 2015 Election Studies in Kentucky and Mississippi

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

GfK Custom Research

December 2015



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 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.

Our first examination of using an online poll to predict election outcomes was in November 2014 for elections that took place in Georgia and Illinois; the results were reported separately. KnowledgePanel (KP) is the largest online panel in the United States with over 55,000 members for which panelists are selected with known probabilities from an ABS frame that represents U.S. households. Due to its size, this sample can be useful for state-specific studies. Even with the largest sized probability panel, some states, especially when filtered for likely voters may have smaller samples than desired. One technique GfK has been developing is to use non-probability samples to supplement the KnowledgePanel. By understanding and adjusting for the biases present in the non-probability samples we can blend the samples to enable larger sample sizes. Our general findings were that using our probability-based sample (KnowledgePanel®) alone with the online completion mode was superior to outcomes obtained by the non-probability sample (NPS) alone. However, we used KnowledgePanel as the basis to adjust for NPS biases which enabled us to blend the samples together as a calibrated larger



sample. We found the calibrated and blended solution yielded reasonably close approximations to actual vote proportions, outperforming the accuracy of exit polls (before they were weighted to final election outcomes). In addition, the results were quite comparable when looking at demographics and attitudes related to vote choice when compared with exit polls. Another purpose of the first study was to examine the influence of likely voter models; we found that the simpler model gave similar, and sometimes closer, approximations to the vote outcomes.

To follow-up on the first comparison study, we conducted a second study using a web-based survey in Kentucky and Mississippi in November 2015. We screened for self-identified registered likely voters drawn from two different sample types: 1) a probability-based panel (GfK's KnowledgePanel®) and 2) non-probability sample sources. We compared the actual election outcomes for the Governor, Secretary of State, and Attorney General in both states as well as Lieutenant Governor in Mississippi among registered likely voters from the probability and a combination of the probability and non-probability online samples using our calibration methodology.

In this paper we summarize the results of this second pilot study. Some goals for this study were:

1. Replication of likely voter model results from Study 1;
2. Comparing online sample estimates with other election poll outcomes;
3. Examining the optimal combination of KP and NPS samples; and
4. Comparing pre-election weighting with election outcome weighting in demographic and attitudinal variables.

The vote outcome estimates from the online study were close to the actual election outcomes, accurately predicting the winner of each race using both the KP-only and the



calibrated data. We also found similar attitudes among key demographic groups within the online likely voters for both sample types. We compared the new, simpler likely voter model developed in the IL-GA study against a slightly abbreviated version of the traditional model and found again that the new model was superior in election outcomes to the traditional model. Overall, we again found that results from online surveys with alternative sample types can be a viable alternative to traditional exit poll methodology. We provide some lessons learned that will be used to help inform the next round of pilot testing.

Method

The study fielded in both Kentucky and Mississippi for 8 days, launching October 27 at 5 pm Eastern and closing at 12:22 pm Eastern on Election Day, November 3, 2015 (the earliest poll closing was 6 pm Eastern in Kentucky). Similar to the fielding that occurred in the first study for Georgia and Illinois, we selected all available sample members from the KnowledgePanel. For the non-probability sample (NPS) we specified detailed quota cells in both states using sex, age group, race-ethnicity, and education nested quotas. Once a quota cell was filled, any additional completes entering the survey who fell into a particular cell were deemed as ‘not qualified’. Note that the NPS sample obtained for Mississippi was less than specified by quotas, primarily due to the short field period.

The total number of qualified completes is shown in Table 1 by sample type and state. We had 937 completed interviews for Kentucky and 720 completed interviews for Mississippi.



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

State	Sample Type	Likely Voter	Not Likely Voter	Total
Kentucky	KP Sample	243	87	330
	NPS	393	214	607
Mississippi	KP Sample	131	44	175
	NPS	355	190	545

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, education, and income level. KP and NPS 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 where we calibrate using additional attitudinal and behavioral dimensions that have been found to differentiate between probability-based and NPS respondents (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.

Results

Likely Voter Model Analyses. We first compared results using the new likely voter model with a modified version of the traditional voter model. Unlike the Georgia-Illinois study where respondents were randomly assigned to the traditional likely voter model or a shorter, stated intention to vote measure, for the Kentucky-Mississippi study everyone was asked all items, with the traditional model using the shorter vote likelihood measure from the new model and a similar



vote location item (rather than a separate item as used in the Georgia-Illinois study). The traditional model led to identification of a smaller subset of likely voters, however, all traditional model likely voters were classified as likely voters in the new model, with additional respondents also classified as likely voters in the new model.

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 required 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 also limited to respondents who were registered to vote and based on responses to two additional questions; it 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.

The results for each model are shown in Table 2 by sample source with the overall average candidate error presented at the bottom of the table. We compared the use of all sample without regard to a likely voter model and then the two likely voter models. We found that across all seven races the average absolute deviation between election outcomes and survey results (average absolute candidate error) was largest when no likely voter model was used and smallest with the new vote likelihood model for both KP sample only (average candidate error was 3.0%) and for the combined calibrated sample (KP + NPS) (average candidate error was 3.2%).



Table 2. Vote Outcomes by Sample Source and Likely Voter Model

	Actual Vote %	KP-only All	KP-only Traditional Likely Voter	KP-only New Likely Voter	Calibrated All	Calibrated Traditional Likely Voter	Calibrated New Likely Voter
Kentucky Governor							
Conway (D)	43.8%	45.9%	48.3%	44.9%	46.2%	48.0%	45.8%
Bevin (R)	52.5%	48.6%	45.6%	49.7%	49.3%	47.2%	49.3%
Kentucky Secretary of State							
Grimes (D)	51.2%	58.4%	56.7%	54.9%	54.9%	54.5%	54.0%
Knipper (R)	48.8%	41.6%	43.3%	45.1%	45.1%	45.5%	46.0%
Kentucky Attorney General							
Beshear (D)	50.1%	54.8%	53.2%	52.3%	54.8%	54.8%	54.0%
Westerfield (R)	49.9%	45.2%	46.8%	47.7%	45.2%	45.2%	46.0%
Mississippi Governor							
Gray (D)	32.3%	35.6%	30.2%	32.3%	36.1%	36.3%	36.4%
Bryant (R)	66.4%	56.7%	65.5%	63.7%	60.2%	61.3%	61.0%
Mississippi Lt. Governor							
Johnson (D)	35.9%	51.5%	40.5%	42.7%	43.3%	39.6%	40.1%
Reeves (R)	60.5%	45.2%	55.1%	53.3%	53.8%	57.9%	57.3%
Mississippi Secretary of State							
Graham (D)	35.6%	40.7%	27.9%	30.2%	41.1%	37.7%	39.4%
Hosemann (R)	61.4%	55.9%	67.8%	65.8%	56.4%	59.9%	58.2%
Mississippi Attorney General							
Hood (D)	55.3%	56.0%	53.9%	55.3%	57.0%	54.0%	53.9%
Hurst (R)	44.7%	44.0%	46.1%	44.7%	43.0%	46.0%	46.1%
Avg. Absolute Candidate Error		6.1%	4.2%	3.0%	4.3%	3.4%	3.2%
Average Spread Bias		-12.2%	-2.9%	-3.2%	-8.6%	-6.0%	-5.7%

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 estimates based on all sample with no likely voter filter for both KP and Calibrated sample with had the highest average spread bias, while the KP-only sample had the lowest average spread bias. Though the new likely voter model had the lowest average absolute candidate error, it had slightly higher average spread bias.

Comparison with Pre-election Polls. We located two polls conducted in the month prior to the election for the offices in Kentucky. The first poll was conducted by Mason-Dixon Polling & Research from Oct. 6 to 8, 2015. The second poll was conducted by SurveyUSA from Oct. 23 to 26, 2015. For Mississippi, we located one poll conducted by Mason-Dixon Polling & Research from Oct. 21 to 23, 2015. All polls were telephone interviews using both landlines and mobile phones. We summarize the results in Table 3, with the proportions reported for candidate selection as a proportion of total candidate selections.



Table 3. Pre-Election Polls Compared with KP and Calibrated Sample Election Results¹

	Actual Vote %	Mason- Dixon	SurveyUSA	KP-only Likely Voter	Calibrated Likely Voter
Kentucky Governor					
Conway (D)	43.8%	47.8%	49.5%	44.9%	45.8%
Bevin (R)	52.5%	45.6%	44.0%	49.7%	49.3%
Kentucky Secretary of State					
Grimes (D)	51.2%	51.1%	57.5%	54.9%	54.0%
Knipper (R)	48.8%	48.9%	42.5%	45.1%	46.0%
Kentucky Attorney General					
Beshear (D)	50.1%	54.9%	57.3%	52.3%	54.0%
Westerfield (R)	49.9%	47.6%	42.7%	47.7%	46.0%
Mississippi Governor					
Gray (D)	32.3%	29.5%		32.3%	36.4%
Bryant (R)	66.4%	69.5%		63.7%	61.0%
Mississippi Lt. Governor					
Johnson (D)	35.9%	35.9%		42.7%	40.1%
Reeves (R)	60.5%	62.0%		53.3%	57.3%
Mississippi Secretary of State					
Graham (D)	35.6%	29.3%		30.2%	39.4%
Hosemann (R)	61.4%	68.5%		65.8%	58.2%
Mississippi Attorney General					
Hood (D)	55.3%	53.2%		55.3%	53.9%
Hurst (R)	44.7%	46.8%		44.7%	46.1%
Avg. Absolute Candidate Error		3.1%	6.9%	3.0%	3.2%

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 Mason-Dixon poll results from:

<http://mason-dixon.com/wp-content/uploads/2015/10/KY-1015-Poll-Part1.pdf>

<http://mason-dixon.com/wp-content/uploads/2015/10/KY-10-15-Poll-Part-21.pdf>

<http://yallpolitics.com/index.php/yp/post/42506/>

<http://mason-dixon.com/wp-content/uploads/2015/10/MS-10-15-Poll-2.pdf>

³ SurveyUSA poll results from: <http://www.surveyusa.com/client/PollReport.aspx?g=24d84559-f5f3-4cb8-b234-49672e61463f>

Generally, the KP-only and Calibrated KP+NPS sample performed better than the pre-election poll conducted by SurveyUSA in terms of average candidate error. Although the average



candidate error is similar to Mason-Dixon, both Mason-Dixon and SurveyUSA incorrectly identified Conway as the winner in the Governor's race in Kentucky whereas the KP only and Calibrated samples correctly identified Bevin as the winner. In nearly all cases, the pre-election polling and the GfK online polling overstated the vote for the Democratic candidate. We believe this specific area warrants further study to determine possible reasons for the bias and then possible ways to adjust for such bias without overcorrection.

Comparing Pre-election Weighted with Election Weighted Results. We next compared demographics and attitudes of respondents based on vote choice for each state and for each office election. Appendix B reflects normal demographic weighting for KP sample and calibration for the blended calibrated sample (KP + NPS). Appendix C shows results for KP sample and calibrated sample after each had been post-stratified to the election outcome. We found that, in general, KP-only sample and calibrated sample had comparable demographic and attitudinal profiles for each candidate choice, and that these results were not substantially affected by post-stratifying by election outcomes, similar to what we obtained in the GA-IL study. Because there were no exit polls conducted in the Kentucky and Mississippi elections, we did not have the exit polls for comparison.



Conclusions

Generally, we replicated results for the new likely voter model found in the prior GA-IL study, finding that in these two states (KY-MS), the new voter model was superior to the traditional model (and better than using no likely voter model at all) for both KP-only sample and Calibrated KP+NPS sample.

Results for the full KP-sample only showed less overall bias than the full KP+NPS calibrated sample.

Lessons Learned

1. A likely voter model is essential to improve prediction, and a simpler model is better.
2. Results for demographics and attitudes show that similar results are obtained for both KP-only and blended, calibrated samples.