

The 2019 Kentucky Governor's Race: A Case Study of State Polling

NBC News and The Marist Poll, in collaboration with Edison Research, conducted several polling experiments in the months leading up to the 2019 Kentucky Governor's contest. The goal was to examine the perceived pitfalls of state polls from the 2016 presidential race and improve poll estimates and election analysis for 2020.

Nate Silver of *FiveThirtyEight*, and others, analyzed the state-by-state vote in 2016 and observed that support for Hillary Clinton and Donald Trump was highly correlated with education. They argued that state polls undercounted white voters without a college degree -- Donald Trump's key group. Since then, many pollsters using a variety of modes of data collection have adjusted samples by education to make up for the difference.

The Kentucky project examines this relationship between education and candidate choice. It asks, *What if there is more going on than just an "undercount" of voters without a college degree?*

In fact, one of the more important takeaways from the Kentucky project is that the need to balance education is actually a clue and *not* a solution. The Kentucky project found the most accurate results of the Kentucky Governor's race occurred when samples of the state were drawn to 1 - account for geography, specifically metropolitan versus non-metropolitan regions, and 2 - were *not* enhanced by listed sample to improve data collection productivity. All models improved when an RDD landline sample (without listed sample) was included. In each case, counties outside of city centers had better coverage and more respondents.

So, a few things we have learned from the Kentucky project:

- The road to improved accuracy in statewide polling lies not just in a statistical adjustment of results (for instance, by education), but in the need to draw **geographically** representative samples in each state keeping urban, suburban, exurban, and rural geography in mind
- Survey samples are not created equally even in scientifically designed RDD (Random Digit Dial) live interviewer methods. Pollsters make many judgements about how to draw a sample (such as including listed sample, deciding the proportion of listed sample to use, matching numbers with other databases to improve contact rates, etc.) which are still consistent with RDD
- Sampling techniques used to improve interviewer productivity have an impact on geography. It appears, the more sampling frames (lists from which telephone samples are drawn) are adjusted to increase productivity in the data collection process (by using listed sample for landlines and billing address based sample for cell phones), **the less representative of exurban and rural communities the sample becomes** even within non-metropolitan counties
- Using billing address based sample for the cell phone frame resulted in demographic disparities (respondents with higher education and income) even when taking into account geographic coverage
- Weighting by demographic characteristics (education, income) will not always account for the diversity of attitudes that exist among individuals who, though similar demographically, have different views based on where they live geographically in a state.

Looking “under the hood,” the RDD-only landline sample (without any listed sample) drawn by geography garnered more respondents in rural areas and less dense clusters in city centers. In a 2016 context, this simple change in sampling found the so-called elusive “shy” Trump voter -- white men without a college education.

Sampling products to improve productivity for cell phone numbers (billing address matches) ironically have the best coverage of the state, that is, the most counties covered. But rather than improve representativeness, they result in samples with a greater proportion of respondents with higher education and income than the RDD-only cell phone sampling frames.

The weight by education “fix” post-2016, if applied in 2020, may result in pollsters “just fighting the last war,” and missing the new realities of 2020. Instead, more consistently representative samples, although certainly a burden in cost and time, may actually provide insight into what becomes the next change in election turnout. With all that has gone on in American politics since 2016, it may be partisan changes in turnout that are the answer to this year’s election puzzle.

Going forward, the NBC News/Marist Poll will:

- incorporate geographic checks in drawing samples for representative statewide surveys
- limit the use of samples enhanced to improve productivity for both landline and cell phone sampling frames
- avoid a one-size-fits-all polling solution, based on weighting samples for what appears to be an undercounted demographic factor without further analysis. In 2016 it was, education; in 2020 the “quick fix” may be party identification.
- carefully track the geographic differences observed from sample to sample and from state to state.

What We Did

This NBC News/Marist Poll research analyzed the results of the Kentucky gubernatorial race in comparison with a series of pre-election polls conducted in the state during the summer and fall of 2019. The experiments tested multiple types of sampling frames for landline and cell phones. Polls conducted by live interviewers during the summer and early fall of 2019 yielded preliminary evidence that the most representative results occurred when the samples accounted for geography, specifically metropolitan versus non-metropolitan regions, and were not enhanced by listed sample to improve productivity.

The final phase of the experiment fielded each model from October 30, 2019 through November 3, 2019 just prior to Election Day. Six different models constructed with various combinations of five different sampling frames revealed very little variation across most topline results. There were, however, several important exceptions.

Sampling	
Treatment A: Landline	RDD plus listed
Treatment A: Cell	RDD with activity code; inactive numbers removed
Treatment B: Landline	RDD selected for metro/non-metro counties within each region of the state
Treatment B: Cell	RDD selected for metro/non-metro counties within each region of the state with activity code; inactive numbers removed
Treatment C: Landline	n/a
Treatment C: Cell	Sample selected based on billing address; includes in-state residents with out-of-area phone numbers

Data							
Treatment		Adults		Registered Voters		Likely Voters	
		Sample Size	MoE	Sample Size	MoE	Sample Size	MoE
Model 1	Treatment A LL + Treatment A Cell	n=1,111	+/- 3.7%	n=947	+/- 4.0%	n=565	+/- 5.2%
Model 2	Treatment B LL + Treatment B Cell	n=1,113	+/- 3.6%	n=942	+/- 3.9%	n=554	+/- 5.1%
Model 3	Treatment A LL + Treatment C Cell	n=1,108	+/- 3.7%	n=950	+/- 4.0%	n=566	+/- 5.2%
Model 4	Treatment B LL + Treatment C Cell	n=1,109	+/- 3.6%	n=953	+/- 3.9%	n=567	+/- 5.1%
Model 5	Treatment A LL + Treatment B Cell	n=1,112	+/- 3.7%	n=937	+/- 4.0%	n=553	+/- 5.1%
Model 6	Treatment B LL + Treatment A Cell	n=1,112	+/-3.7%	n=947	+/-4.0%	n=564	+/- 5.2%

The survey model resulting in the most representative sample (Model 6 RDD landline and RDD Cell) *precisely* mirrored the final Election Day (November 5, 2019) vote totals provided by the Kentucky Secretary of State. Still, models which included samples from other RDD cell frames were also statistically similar.

		Weighted Results						
		Kentucky Likely Voters						
		ACTUAL	Model 1 (A)	Model 2 (B)	Model 3 (C)	Model 4 (D)	Model 5 (E)	Model 6 (F)
		Column %	Column %	Column %	Column %	Column %	Column %	
Kentucky tossup for Governor including those who are undecided yet leaning toward a candidate or have already voted	Andy Beshear, the Democrat	49.19	46%	49%	49%	50%	47%	47%
	Matt Bevin, the Republican	48.83	47%	44%	44%	43%	45%	47%
	John Hicks, the Libertarian	1.97	3%	4%	4%	3%	5%	2%
	Other	n/a	<1%	<1%	<1%	<1%	<1%	<1%
	Undecided	n/a	3%	3%	3%	3%	2%	4%

Looking “under the hood,” though, the analysis of the NBC News/Marist Poll study highlights the distinction between landline samples. These landline frames consisted of RDD-only samples selected in proportion to the population of Kentucky’s metropolitan and non-metropolitan counties, and samples randomly selected in proportion to the population and supplemented with listed phone lines (“RDD+listed”). Here, the RDD-only landline sample drawn by geography provided an informative difference in rural areas. A striking example of the impact was the difference between white men without a college education in each landline frame. In a 2016 context, this could, account for what appeared to be an elusive “shy” Trump voter in state polls but was inexplicably not apparent in national polls.

		Unweighted White men, no college	
		RDD + Listed	RDD only
		Kentucky Adults (A)	Kentucky Adults (B)
		Column %	Column %
Region	Eastern Kentucky	28%	35%
	Bluegrass Country	11%	15%
	Northern KY-Louisville Suburbs	17%	12%
	Jefferson County (Louisville)	18%	15%
	Western Kentucky	26%	24%
Household Income	Less than \$50,000	59%	66%
	\$50,000 or more	41%	34%
Area Description	Big city	17%	13%
	Small city	17%	17%
	Suburban	9%	6%
	Small town	25%	26%
	Rural	32%	39%
Party Registration	Democrat	53%	47%
	Republican	40%	50%
	Other	7%	2%

The problem is exacerbated when sampling products (Models 3 and 4) which use cell phone numbers based upon listed household information (billing address matches) are included. Interestingly, these samples had the best coverage of the state. As a result, it was anticipated that this sampling frame, which oversampled cell phone numbers that matched a billing address, would be a method to control for undercounting exurban and rural geography. Instead, despite the extensive coverage of the state, this sampling frame was more likely to result in a greater proportion of respondents with higher education and income than the RDD-only cell phone sampling frames. This demographic difference neutralized the advantage of geographic coverage of cell phone samples using billing address.

			Unweighted		
			RDD Cell A	RDD Cell B	Billing Address Cell
			Kentucky Adults (A)	Kentucky Adults (B)	Kentucky Adults (C)
			Column %	Column %	Column %
Race/Ethnicity / Household Income	White	Less than \$50,000	35%	37%	31%
		\$50,000 or more	50%	48%	56%
		Difference	-15%	-11%	-25%
Race/Ethnicity / Education	White	Not college graduate	56%	56%	53%
		College graduate	30%	28%	33%
		Difference	26%	28%	20%

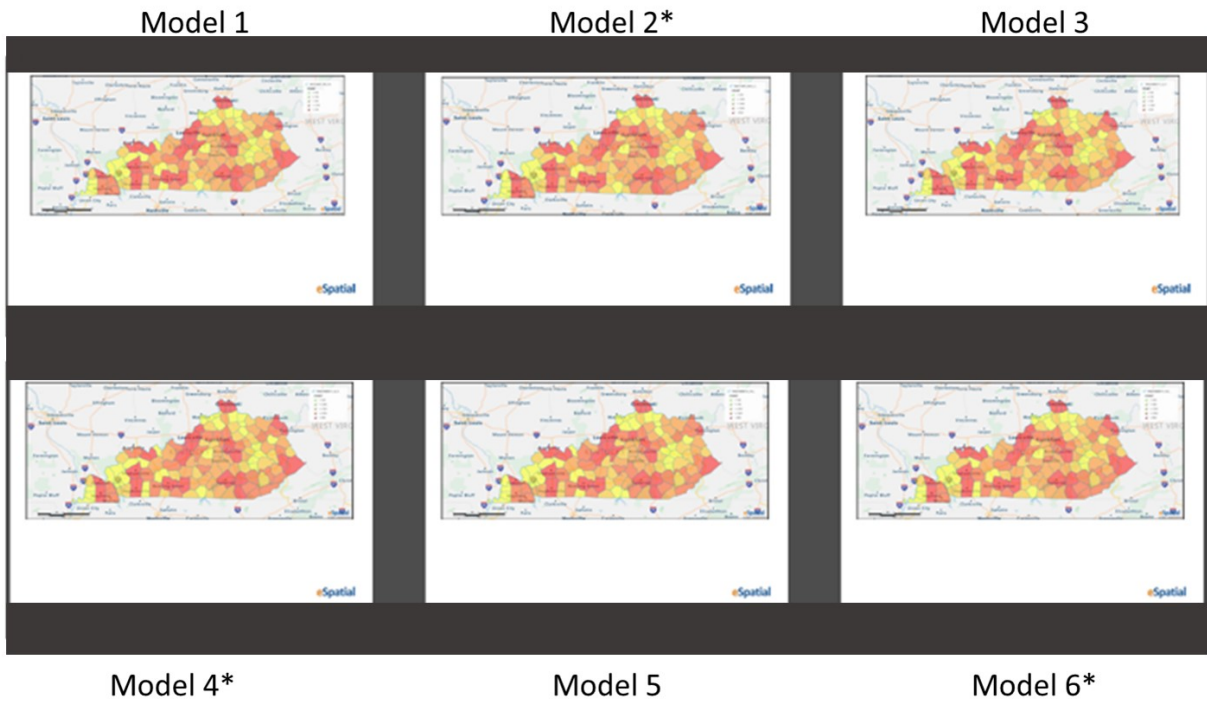
The NBC News/Marist Poll recognizes this experiment was conducted in just one state during an off-year election that featured an extremely unpopular incumbent. Although the results may not be applicable to other states in other elections, they do provide a guide for pollsters when conducting statewide election polls.

The results of this NBC News/Marist Poll study demonstrate that geography is an important factor in creating representative surveys. It further suggests that caution should be used when enhancing samples to improve productivity by matching to listed households whether in landline or cell phone frames. A one-size-fits-all polling solution, based on weighting samples for an undercounted demographic factor, in this case, education, can obscure the underlying problem.

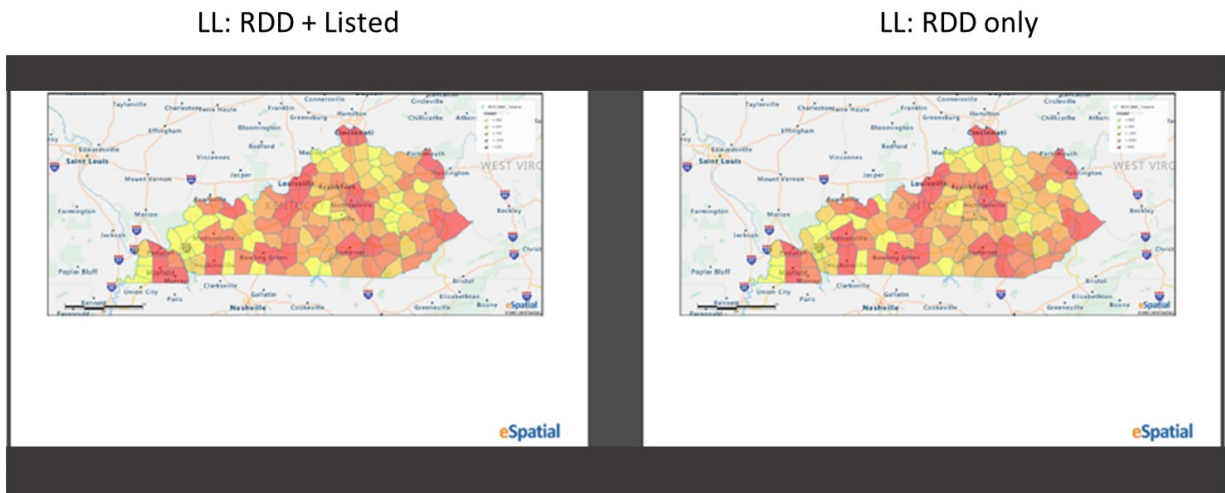
Methodology

This sample analysis is based on 2,775 interviews with Kentucky adults conducted October 30th through November 3rd, 2019 by The Marist Poll. Adults 18 years of age and older residing in the state of Kentucky were interviewed in English by telephone using live interviewers. Mobile phones are treated as individual devices. After validation of age, personal ownership, and non-business-use of the mobile phone, interviews are typically conducted with the person answering the phone. Within each landline household, a single respondent is selected through a random selection process to increase the representativeness of traditionally under-covered survey populations.

The analysis assessed each treatment geographically by county, as well as individual landline and mobile sampling frames. As expected, sample coverage is very similar across models. However, in any combined model that had just RDD landline sample (models 2, 4 and 6), there were counties outside of city centers that had better coverage.



While the combined models showed similar coverage, there are some nuanced differences when looking at the landline and cell sampling frames individually. The RDD-only landline sample provided better coverage of the rural areas of the state.



Interestingly, the mobile phone sample based on billing address has the most complete coverage of the state. But, as noted above, this frame results in survey respondents with higher income and education.

CELL: RDD A

CELL: RDD B

CELL: Billing Address

