

How I weighted the Centerline National Poll and why I dropped 107 cases

How the sausage was made... and a cautionary tale about what surveys you believe.

I believe in radical transparency – not just the type that you get with super-detailed methodology statements, but also the “how my brain worked through the data” kind. Also because I found some weird patterns in the data that led to discovering 107 cases that appeared to be junk responses, so I want to share the lessons for using online opt-in panel data here, *even with careful scrubbing*. Here’s the relevant part of the methodology statement:

This survey of 1,570 U.S. adults age 18 and over was conducted by Centerline Research & Strategy from September 26 to October 1, 2026. 1145 respondents completed the survey via a nonprobability online panel sample provided by RepData, and 425 were probability-based text-to-web respondents sampled from the voter file. Survey responses were screened for quality on measures of attention, time to completion, and suspected bot or AI behavior, and suspicious cases were discarded during data collection. Sampling and fielding were managed by TrueDot. Post-fielding, 107 cases were removed for straight-lining through the survey (choosing the top choice on every question); 101 of those were from the online panel and 6 from text-to-web.

The important piece here is the sample makeup. It’s mostly an opt-in online panel, paired with about 27 percent text-to-web from the voter file. Any time you are working with opt-in panel data, there are about a million pitfalls to watch out for – and a ton of polls use these sources because of their low cost. It does pay, however, to get the sample from a reputable place that screens respondents, like [RepData](#), and to use a fielding team that scrutinizes every complete for signs of bots, AI, or low quality, like [TrueDot](#). I didn’t initially spend much time digging into data quality given those layers of security. What I eventually found wasn’t anything they would have screened for – and it’s a judgment call on whether it even warrants exclusion from the data.

But first, the process:

I prepared weights for adults, registered voters, and likely voters, and crafted my definition of likely voters ahead of time, so that when I got the data I was ready to run it. Once I had the final data, I created unweighted crosstabs to explore it before applying the demographic and region weights. I dug into the data again before adding partisan weights. I tweaked the weight values so that I wasn’t more than doubling or less than halving any unweighted value. There were quotas on data

collection, so most values did not have this problem, although a few of the crosstab cells (e.g., young women) had to be adjusted. Along the way, though, I noticed that as I tweaked the weights, the data would move in counterintuitive ways. The biggest red flag was that Trump's approval *went up* at the same time the Democratic advantage on the generic ballot increased. Those two things don't go together.

This is where I diverged from a lot of polling practices – I stopped and dug in, rather than assuming it's just weird people. It cost me two days of going back and forth between other tasks that had to get done and the data, but I needed to find out what was happening. I adhere to the adage that if something looks wrong, it probably is. I traced the cross-pressured cases to a group of mostly young respondents of color. Even then, the patterns made little sense – they were mostly Hispanic women and Black men, and *all of them* under age 45. I would have been less suspicious if it had been all young men, since they shifted toward Trump in 2024, but the fact that it was Hispanic women set off alarm bells.

Since that made little sense to me, I asked Claude to search the data for the cases and what commonalities they have, and the response came back with 47 cases that shared these features:

- The cases answered with only top-box or positive responses elsewhere in the survey.
- 44 of the 47 came from the panel sample (a 3.9% rate, against 0.7% for TTW).
- Their median completion time was 13.4 minutes, against 16.2 for everyone else.

I confirmed the patterns in the data, then asked Claude to identify other cases that shared that pattern. I asked for cases that fit both criteria: Straight-lining positive responses and top answer options, plus consistently giving contradictory answers. That added 60 cases to the pile, 57 from the online panel, with a median completion time of 12.7 minutes.

Removing these cases is a judgment call. All of them passed the two attention checks embedded in the survey, took a reasonable amount of time to complete the survey, even if it was shorter than other respondents, and showed no signs of being bots or AI. I have little doubt they were real people taking the survey, and I don't know for sure whether their responses are genuine or junk. It is possible for people to have all sorts of conflicting attitudes. However, most of the time those conflicting attitudes don't involve *both* response patterns that were present, so I strongly believe removing these cases improves data quality and accuracy.

Removing these cases did change the results, especially since they were disproportionately Democrats who approved of Trump because of how the bad response patterns interacted with the survey design – selecting top boxes meant approving of Trump and picking Democrats for the 2026 vote. Removing the cases took 3 percentage points off Trump's approval rating across all adults, registered voters, and likely voters. The change narrowed the generic congressional ballot margin by 1 to 2 points across the three populations.

How removing low-quality cases changes the results

Key measures for U.S. adults, registered voters and likely voters in the Centerline national poll, weighted with and without the flagged cases

	All adults		Registered voters		Likely voters	
	All cases	FINAL Flagged cases removed	All cases	FINAL Flagged cases removed	All cases	FINAL Flagged cases removed
<i>Unweighted n</i>	1,570	1,463	1,339	1,239	1,007	930
<i>Design effect</i>	1.55	1.55	1.30	1.34	1.37	1.40
Trump job approval %						
Approve	34	31	38	35	41	38
Disapprove	61	64	61	64	59	62
Approve minus disapprove	-27	-33	-23	-29	-18	-24
2026 congressional ballot %						
Democratic candidate	45	43	50	49	51	50
Republican candidate	33	33	38	39	41	42
Democratic minus Republican	+12	+10	+12	+10	+10	+8
Party identification with leaners %						
Democrat or lean Democratic	43	42	47	46	47	46
Independent, no lean	22	23	12	13	9	10
Republican or lean Republican	35	35	41	41	44	44
Democratic minus Republican	+8	+7	+6	+5	+3	+2
2024 presidential vote (among those who voted) %						
Donald Trump	44	44	45	45	46	46
Kamala Harris	50	50	50	50	51	50
Harris minus Trump	+6	+6	+5	+5	+5	+4

The FINAL columns are the published estimates. Flagged cases are 107 respondents removed after fielding for low-quality response patterns. Both columns in each group are weighted to the same demographic, regional and religious affiliation targets.

Margins are calculated from the rounded percentages shown. The design effect is the Kish approximation.

Source: Centerline Research & Strategy national poll, September 26–October 1, 2026.

The bigger picture issue is how we work with opt-in online panel data. Many political surveys rely on this data, and many pollsters use a set weighting structure without looking at the data and how weighting changes it. There is a good theoretical reason for this: If you look at the data along the way, your biases can slip into how you're weighting it. Obviously, I was looking at Trump approval and congressional ballot as I was going – I was also looking at the party identification balance and 2024 recall vote. The exact values of these variables were irrelevant to my process; I was looking for patterns in how weighting moved the numbers.

Because of that iterative process, I found 107 problematic cases. I'm reporting 35 percent Trump approval among registered voters instead of 38 percent, and a 10-point generic congressional ballot spread instead of 12. The answer options were rotated on party identification and 2024 vote

recall, so the effect was more even: there was a one-point difference on party identification with leaners, and no difference on 2024 vote recall.

How many fast-operating pollsters who use opt-in panel data do you think are spending two days digging into data inconsistencies before finalizing the results, *then* a few hours writing it up to explain the whole thing? This is transparency, folks. Expect it of pollsters.