

The CPS Voting and Registration Supplement Overstates Minority Turnout*

Stephen Ansolabehere
Harvard University
sda@gov.harvard.edu

Bernard L. Fraga
Emory University
bernardfraga@emory.edu

Brian F. Schaffner
Tufts University
brian.schaffner@tufts.edu

Conditionally Accepted at the *Journal of Politics*

January 1, 2021

Abstract

The Current Population Survey (CPS) is a key source of information on who votes. Turnout estimates derived from the CPS are often cited in academic research on participation, widely used in the calibration of surveys, and central to ongoing legal and policy debates over the protection of voting rights in the United States. We compare CPS estimates to official voter turnout records from 2008-2018, and document consistent, significant discrepancies that call into question the reliability of CPS turnout statistics. Specifically, the CPS overestimates Black and Hispanic turnout relative to non-Hispanic Whites, whether relying on turnout rates as a share of eligible citizens or the racial/ethnic composition of the voting population. Sampling error and commonly used adjustments to CPS estimates do not account for or correct this bias, and thus academics and policy makers should use discretion when judging recent shifts in voter turnout with survey data.

Keywords: voting; survey measurement; Black turnout; Latino turnout; voting rights

*Supplementary material for this article is available in the online edition. Replication files for the analyses presented in this article will be available in the *Journal of Politics* Data Archive on Dataverse (<http://thedata.harvard.edu/dvn/jop>).

Researchers have long relied on the CPS’s Voting and Registration Supplement as an authoritative source of data on turnout in the U.S., including in studies examining turnout among racial and ethnic minority groups (Jackson 2003; Wolfinger, Highton and Mullin 2005; Sides, Schickler and Citrin 2008; Francia and Orr 2014; Rocha et al. 2010). For example, Leighley and Nagler (2013) use the CPS to conclude that “where thirty years ago black turnout lagged substantially behind white turnout, in 2008 it was the same as white turnout” (p. 31). Think tanks and policy makers rely heavily on the CPS to craft public policies governing elections. And in 2013, Chief Justice John Roberts pointed to CPS data as central to the Court’s decision to strike down Section 4(b) of the Voting Rights Act in *Shelby County v. Holder*. Based on CPS estimates, the Court concluded “that African-American voter turnout exceeded white voter turnout in five of the six States originally covered.” The Court’s reading of the CPS data was correct, but there are growing concerns among researchers that the CPS may not accurately measure turnout, particularly for minority citizens (McDonald 2007; McKee, Hood and Hill 2012; Bauman 2018).

The CPS data are also used for calibrating weights for surveys that are widely used to understand American politics. The Cooperative Congressional Election Study (CCES), Exit Polls, Associated Press’s VoteCast, and many other surveys all use the CPS to calibrate weights for voters or registered voters. The CPS is also used in recent applications of multilevel regression and post-stratification (MRP) both in political science as well as election polling and modeling (Ghitza and Gelman 2013).

While the government-produced CPS performs better than many other surveys in collecting data on political participation, it is still susceptible to problems such as misreporting and non-response (Traugott and Katosh 1979; Cassel 2004; Bernstein, Chadha and Montjoy 2001; Ansolabehere and Hersh 2012; Cuevas-Molina 2017; Enamorado and Imai 2018; Jackman and Spahn 2019). The CPS relies on respondents to self-report whether they are registered to vote and whether they voted in the most recent election. If some groups have more errors in self-reported turnout than others, it may lead to inaccurate conclusions when comparing participation rates (Cuevas-Molina 2017; Enamorado and Imai 2018). For example, McKee, Hood and Hill (2012) find that black turnout rates in Georgia in 2008 were significantly lower than what the CPS reported.

Does the CPS provide an accurate picture of turnout patterns across racial and ethnic groups? We test the validity of the CPS’s turnout estimates by comparing those rates to what voter files indicate. We show that the CPS overstates turnout rates more for minority voters than for whites

and that this bias cannot be easily corrected. Our findings thus raise concerns about the use of the CPS for statistical analysis of participation.

Comparing Voter File and CPS Turnout Estimates

We compare turnout for different racial groups as estimated by the CPS with turnout for those same groups based on voter files. A growing body of academic literature relies on voter files to study electoral behavior, and voter records are now a key source of information used to validate survey self-reports of turnout (McDonald 2007; Ansolabehere and Hersh 2012). We rely on data from six southern states where self-reported race/ethnicity is provided during the voter registration process: Alabama, Florida, Georgia, Louisiana, North Carolina, and South Carolina.¹ For these states we can determine how accurately racial/ethnic differences in voter turnout are reflected in the CPS.

The voter file data were provided by the voter file firm Catalist. Catalist has an established track record of producing estimates of the total number of voters that closely align with semi-official statistics as computed by McDonald (2017) (Fraga 2018). The voter file data are the baseline comparison, but they themselves may have flaws (Nyhan, Skovron and Titiunik 2017). If vote history was incomplete in some counties or states due to faulty record keeping, then the voter file data would not be a valid point of comparison. However, we find that the voter file data provided by Catalist does not suffer from any widespread lack of coverage. In fact, as we show in the online appendix, the turnout estimates using vote history (from Catalist) and official turnout figures differ by no more than 1.3 percent from 2008 to 2018 and these deviations do not appear to be more common in precincts with higher numbers of minority voters.

Turnout rate calculations consist of a numerator, the number of people of a given race or ethnicity who voted, and a denominator, the number of people of a given race or ethnicity who were eligible to vote. We gathered the number of individuals who are listed as having voted in general elections from 2008-2018 by self-reported race/ethnicity, establishing these figures to be the most valid available count of the number of persons voting by race in each of the six states. We estimate the eligible electorate for each racial group based on U.S. Census Bureau Population Estimates Program and American Community Survey data. As discussed in Appendix A.1, we construct a citizen voting-age

¹Mississippi and Tennessee also ask registrants to volunteer their race/ethnicity, but in the above mentioned states well over 90% of registrants provide this information.

non-incarcerated population estimate that approximates the denominator used in CPS calculations.

We compare voter file and Census derived statistics to the CPS estimates in the same six states. Drawing on the publicly available raw, individual-level November CPS Voting and Registration Supplement data, we are able to precisely replicate the number of voters and rates of voter turnout reported by the Census Bureau in their official publications. The CPS also calculates turnout rates by dividing their tally of the number of persons who voted by the sampled population that is both over 18 and is a naturalized or native-born U.S. citizen.²

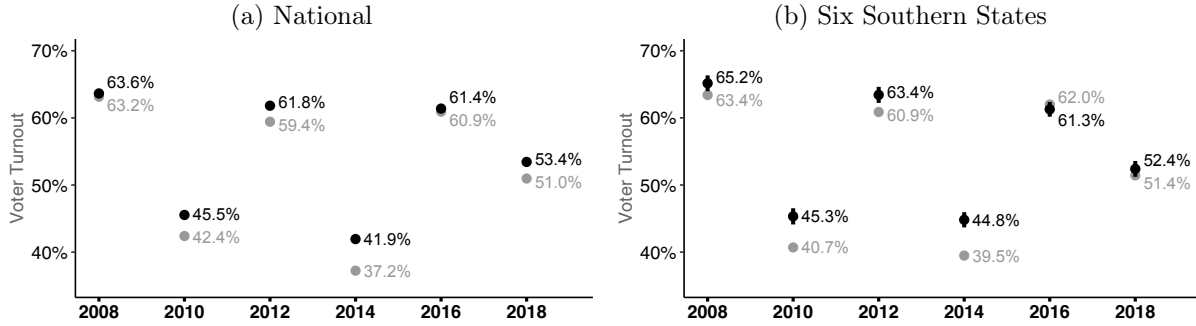
The CPS estimate of turnout is based on a sample, and thus random sampling error may account for deviations we see from the true population parameter. The CPS uses a multi-wave household-based stratified sampling scheme in order to obtain data on over 140,000 Americans every month. They do not provide data on the sampling units they use, but do provide margins of error based on generalized variance parameters for some top line statistics. Previous research establishes that inferences relying on CPS-provided measures of uncertainty are unreliable (Davern et al. 2006). Following recommendations from Davern et al. (2007), we instead use the available geographic and household indicators to approximate this complex sampling scheme and compute 95% confidence intervals for CPS rates of voter turnout by state and race/ethnicity.

Results

Figure 1 shows the overall turnout rate in elections held from 2008 through 2018 as well as the CPS estimates for those rates. Figure 1a presents these comparisons nationally. The CPS estimates are presented with 95% confidence intervals, though these are smaller than the dots in the plot. In presidential election years, the CPS over-estimates the turnout rate by between .4 and 2.4 percentage points. In midterm election years, when turnout is lower, the CPS over-estimates turnout by 2.7 to 4.7 percentage points. Only in 2008 does the actual turnout rate fall within the 95% confidence interval for the CPS estimate. Generally speaking, the discrepancies in Figure 1 grow larger when turnout drops, suggesting that turnout self-reports captured by the CPS may fluctuate less than turnout itself.

²Those not responding to CPS voting questions are considered non-voters, a practice questioned by many researchers. We use CPS reported figures for consistency with official estimates, and discuss the impact of non-response on CPS estimation error later in the paper.

Figure 1: Overall Voter Turnout, CPS vs. State Estimates



Note: Black points and confidence intervals indicate estimates from the Current Population Survey. Gray points are estimates of the number of voters in each election divided by the nonimprisoned citizen voting-age population, as estimated by McDonald (2017).

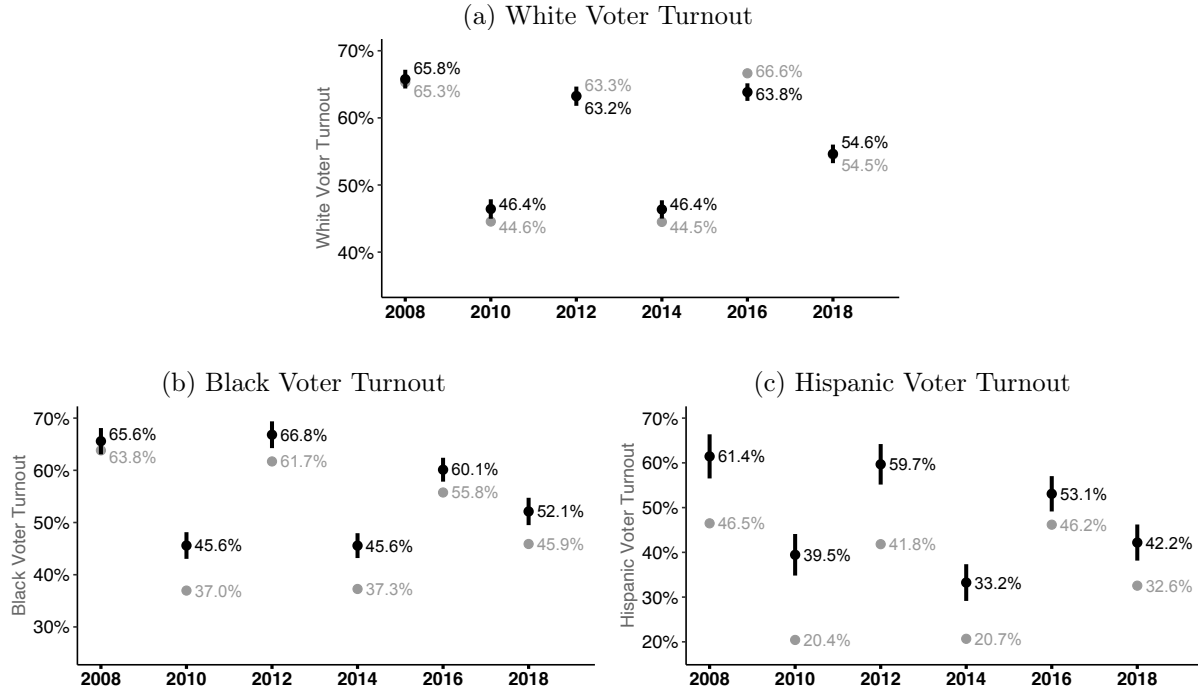
Figure 1b shows the same overall turnout comparisons, but this time limiting the comparison to the six southern states that we focus on for our comparison of turnout rates by race/ethnicity. The overall patterns are similar; however, in the most recent two elections, the actual turnout rate falls within the 95% confidence interval of the CPS estimate.

Now we turn to our comparison of actual turnout rates and CPS estimates for the non-Hispanic White, African-American, and Hispanic populations in the six southern states. These are presented in Figure 2. The first important point from this figure is that Black turnout did not, in fact, exceed white turnout in 2012. In every election plotted in the figure, the turnout rate among African Americans based on the voter file data lags behind the turnout rate among whites. In the most recent two elections, this gap has been quite large. In 2016, the turnout rate among whites was 10.8 percentage points higher than the turnout rate among blacks. In 2018, the gap in turnout rates was 8.6 percentage points. Yet, researchers relying on the CPS to examine turnout rates in these six states would find a gap in participation rates of just 3.7 points in 2016 and 2.5 points in 2018.

The plots in Figure 2 show clearly why this happens – the CPS consistently overstates the participation rate among Blacks and Hispanics while it sometimes underestimates participation among whites. Figure 2a shows that in two of the three presidential elections, the CPS accurately estimates the turnout rate among whites in these six states, though in 2016 the CPS data actually underestimated white turnout by 2.8 points. In midterm elections held in 2010 and 2014, the CPS data overstated the white turnout rate by about 2 points.

Figure 2b shows that the CPS over-stated the turnout rate among Blacks by more than 4

Figure 2: Voter Turnout by Race, CPS vs. Voter File Estimates



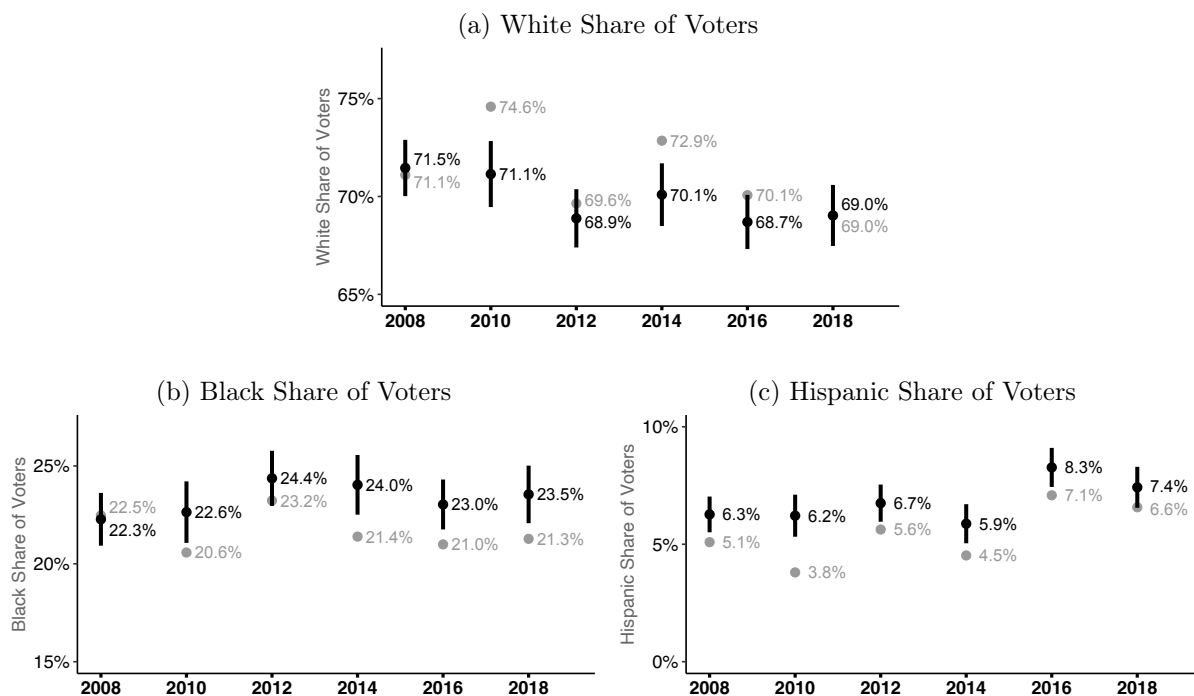
Note: Includes voters in six southern states with comprehensive data on race/ethnicity of registrants (AL, FL, GA, LA, NC, SC). Black points and 95% confidence intervals indicate estimates from the Current Population Survey. Gray points are estimates of the number of voters by race/ethnicity in each election divided by the nonimprisoned citizen voting-age population by race, using data from Catalist and the U.S. Census Bureau.

percentage points in five of the six elections. Even in 2008, the peak year for African American turnout, the CPS still overestimated Black turnout by nearly 2 percentage points (though it was within the 95% confidence interval). Figure 2c shows a similar pattern for Hispanics. The CPS estimates for Hispanic turnout are always much higher than the figure suggested by the voter files.

While turnout rates are often the key statistic for those studying turnout, there is sometimes more interest in the racial composition of those who voted, particularly among analysts who use the CPS to calibrate survey weights or MRP estimates (e.g., Ghitza and Gelman 2013). On this metric, the story is much the same. Figure 3a shows that in the six southern states, the CPS almost always produces an estimate of the white share of the electorate that is lower than what it actually was, though on several occasions the actual composition was within the 95% confidence interval. By contrast, Figure 3b shows that in four of the six elections, the CPS estimate of the black share of the electorate is statistically significantly higher than the actual percentage of voters who were

Black. Finally, Figure 3c shows that the CPS over-states the Hispanic share of the electorate by a statistically significant amount in all but one election cycle. The CPS appears to present a picture of the electorate that is more racially diverse than what the voter files indicate.

Figure 3: Share of Voters by Race, CPS vs. Voter File Estimates



Note: Includes voters in six southern states with comprehensive data on race/ethnicity of registrants (AL, FL, GA, LA, NC, SC). Black points and 95% confidence intervals indicate estimates from the Current Population Survey. Gray points are estimates of the share of voters by race/ethnicity in each election, using data from Catalyst.

Why do the Voter File and CPS estimates differ?

There are at least four reasons why the CPS and vote history data may differ. First, non-response to questions about registration and turnout in the CPS may cause over-estimates of turnout. Non-response to the turnout section of the CPS survey varies by the race/ethnicity of the respondent (Fraga 2018), which could cause the bias in these estimates if all non-respondents are assumed to be non-voters. Hur and Achen (2013) suggest a series of adjustments to the CPS data that instead drop those who do not answer the registration and turnout questions, then reweight state estimates to McDonald (2017)'s turnout data. This technique assumes that, conditional on responding to the CPS voting questions, subgroup turnout rates are accurately measured. However, as we show in the Online Appendix, applying these corrections produces estimates of white voter turnout that are

biased even further downward in the six states we examine, and continues to produce estimates of Black and Hispanic turnout that are as much as 13 percentage points greater than what the voter file indicates. Shares of the electorate (shown in Figure 3), which are not denominator-dependent, are also biased in favor of whites and, in the case of Black and Hispanic shares, well outside of the margin of error in recent elections.

A second possibility would be incomplete record keeping on the state voter files, particularly if there was more missing data among minorities than for whites (Nyhan, Skovron and Titiunik 2017). If voter files were significantly under-counting minority turnout, such a pattern would be evident from the fact that the turnout records in the vote history files would not match the official certified vote counts. In Table A.1 of the online appendix, we show that the voter file counts are very close to the official vote tallies provided by the states. Additionally, when we look at calculations related to each group’s share of the electorate (Figure 3) the discrepancies we document persist. Therefore, it is unlikely that missing data on the voter files can account for these turnout differences.

A third possibility is that minority voters might mis-report turnout at higher rates than whites. Voter validation studies generally find higher rates of over-reporting turnout among minorities (McKee, Hood and Hill 2012; Enamorado and Imai 2018).³ In Table A.3 of the online appendix, we show how over-reporting is more common among Black and Hispanic respondents in the CCES surveys. This provides some support for the mis-reporting explanation, though we do not know whether these patterns would also exist in the CPS survey. Mis-reporting may be less pronounced in the CPS survey since it is not inherently a political survey. To fully test this hypothesis, it would be necessary to match CPS respondents to voter file data to determine the extent to which over-reporting might be causing errors.

A fourth possible explanation for the patterns we observe may relate to non-response bias or differential panel attrition. Individuals recruited for the CPS take surveys in four consecutive months, then are re-contacted again eight months later to take four more surveys in consecutive months. Participation in the CPS is voluntary and panel attrition is significant; one estimate is that panel attrition is 20% over the year between the first and the fifth wave (Rivera Drew, Flood and

³Some studies find that racial differences in over-reporting disappear after controlling for other factors, but what matters here is whether such patterns manifest in the aggregate.

Warren 2014). Studies of panel attrition in other surveys have found that non-voters are more likely to attrit than voters and non-whites also tend to have higher attrition rates (Smith and Son 2010). If Black and Hispanic non-voters are especially likely to attrit from the CPS, then it may lead the sample to have a higher share of Black and Hispanic voters relative to the population. Since no CPS panel spans two election cycles, it is not possible to test this possibility without matching the CPS sample to voter file data.

Conclusion

The analysis presented here points to a serious bias in the CPS turnout estimates, one that has significant consequences for the inferences we make about racial/ethnic disparities in turnout rates. The CPS relies on a standard approach to studying the turnout – simply ask survey respondents whether they voted. But biases in the composition of the CPS samples and/or in responses to the turnout questions are likely responsible for the errors we document here (Smith and Son 2010; Dahlgard et al. 2019; Lahtinen et al. 2019). To better diagnose the problem, the CPS should conduct a voter validation study akin to those undertaken by other surveys. A validation of the CPS Voting and Registration Supplement would help to document the biases in the estimates produced by the Census bureau, and may open the door to methodological innovations that ensure policy debates are informed by data accurately reflecting the composition of the American electorate. In the meantime, we suggest that analysts use caution when making inferences about variation in turnout rates by racial and ethnic groups.

References

- Ansolabehere, Stephen and Eitan Hersh. 2012. “Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate.” *Political Analysis* 20(4):437–459.
- Bauman, Kurt. 2018. “How Well Does the Current Population Survey Measure the Composition of the U.S. Voting Population?” <https://www.census.gov/library/working-papers/2018/demo/SEHSD-WP2018-25.html>.
- Bernstein, Robert, Anita Chadha and Robert Montjoy. 2001. “Overreporting Voting: Why It Happens and Why It Matters.” *Public Opinion Quarterly* 65(1):22–44.
- Cassel, Carol A. 2004. “Voting Records and Validated Voting Studies.” *Public Opinion Quarterly* 68:102–108.

- Cuevas-Molina, Ivelisse. 2017. Voter Turnout Overreports: Measurement, Modeling and Deception
PhD thesis University of Massachusetts Amherst.
- Dahlgard, Jens Olav, Jonas Hedegaard Hansen, Kasper M Hansen and Yosef Bhatti. 2019. “Bias in Self-reported Voting and How it Distorts Turnout Models: Disentangling Nonresponse Bias and Overreporting Among Danish Voters.” *Political Analysis* 27(4):590–598.
- Davern, Michael, Arthur Jones Jr., James Lepkowski, Gestur Davidson and Lynn A. Blewett. 2006. “Unstable Inferences? An Examination of Complex Survey Sample Design Adjustments Using the Current Population Survey for Health Services Research.” *Inquiry* 43(3):283–297.
- Davern, Michael, Arthur Jones Jr., James Lepkowski, Gestur Davidson and Lynn A. Blewett. 2007. “Estimating Regression Standard Errors with Data from the Current Population Survey’s Public Use File.” *Inquiry* 44(2):211–224.
- Enamorado, Ted and Kosuke Imai. 2018. “Validating Self-reported Turnout by Linking Public Opinion Surveys with Administrative Records.” <https://dx.doi.org/10.2139/ssrn.3217884>.
- Fraga, Bernard L. 2018. *The Turnout Gap: Race, Ethnicity, and Political Inequality in a Diversifying America*. New York: Cambridge University Press.
- Francia, Peter L. and Susan Orr. 2014. “Labor Unions and the Mobilization of Latino Voters Can the Dinosaur Awaken the Sleeping Giant.” *Political Research Quarterly* 67(4):943–956.
- Ghitza, Yair and Andrew Gelman. 2013. “Deep Interactions with MRP: Election Turnout and Voting Patterns Among Small Electoral Subgroups.” *American Journal of Political Science* 57(3):762–776.
- Hur, Aram and Christopher H. Achen. 2013. “Coding Voter Turnout Responses in the Current Population Survey.” *Public Opinion Quarterly* 77(4):985–993.
- Jackman, Simon and Bradley Spahn. 2019. “Why Does the American National Election Study Overestimate Voter Turnout.” *Political Analysis* 27(2):193–207.
- Jackson, Robert A. 2003. “DIFFERENTIAL INFLUENCES ON LATINO ELECTORAL PARTICIPATION.” *Political Behavior* 25(4):339–366.
- Lahtinen, Hannu, Pekka Martikainen, Mikko Mattila, Hanna Wass and Lauri Rapeli. 2019. “Do surveys overestimate or underestimate socioeconomic differences in voter turnout? Evidence from administrative registers.” *Public Opinion Quarterly* 83(2):363–385.
- Leighley, Jan E. and Jonathan Nagler. 2013. *Who Votes Now?: Demographics, Issues, Inequality,*

and Turnout in the United States.

- McDonald, Michael P. 2007. "The True Electorate A Cross-Validation of Voter Registration Files and Election Survey Demographics." *Public Opinion Quarterly* 71(4):588–602.
- McDonald, Michael P. 2017. "Voter Turnout." *United States Elections Project*. <http://www.electproject.org/home/voter-turnout/voter-turnout-data>. Last Accessed March 7.
- McKee, Seth C., M.V. III Hood and David Hill. 2012. "Achieving Validation: Barack Obama and Black Turnout in 2008." *State Politics and Policy Quarterly* 12(1):3–22.
- Nyhan, Brendan, Christopher Skovron and Rocío Tituniuk. 2017. "Differential Registration Bias in Voter File Data: A Sensitivity Analysis Approach." *American Journal of Political Science* 61(3):744–760.
- Rivera Drew, Julia A, Sarah Flood and John Robert Warren. 2014. "Making full use of the longitudinal design of the Current Population Survey: Methods for linking records across 16 months\m {1}." *Journal of economic and social measurement* 39(3):121–144.
- Rocha, Rene R., Caroline J. Tolbert, Daniel C. Bowen and Christopher J. Clark. 2010. "Race and Turnout: Does Descriptive Representation in State Legislatures Increase Minority Voting?." *Political Research Quarterly* 63(4):890–907.
- Sides, John, Eric Schickler and Jack Citrin. 2008. "If Everyone Had Voted, Would Bubba and Dubya Have Won?" *Presidential Studies Quarterly* 38(3):521–539.
- Smith, Tom William and Jaesok Son. 2010. *An analysis of panel attrition and panel change on the 2006-2008 General Social Survey Panel*. NORC/University of Chicago.
- Traugott, Michael W. and John P. Katosh. 1979. "Response Validity in Surveys of Voting Behavior." *Public Opinion Quarterly* 43(359-377).
- Wolfinger, Raymond E., Benjamin Highton and Megan Mullin. 2005. "How Postregistration Laws Affect the Turnout of Citizens Registered to Vote." *State Politics & Policy Quarterly* 5(1):1–23.

Biographical Statements

Stephen Ansolabehere is the Frank G. Thompson Professor of Government at Harvard University, Cambridge, MA 02138.

Bernard L. Fraga is an Associate Professor in the Department of Political Science at Emory University, Atlanta, GA 30322.

Brian F. Schaffner is the Newhouse Professor of Civic Studies in the Department of Political Science and Tisch College at Tufts University, Medford, MA 02155.

Online Appendix: Supporting Information for
The CPS Voting and Registration Supplement Overstates Minority Turnout

Contents

A.1	Constructing the Denominator for Turnout Rate Calculations	1
A.2	Results After Hur and Achen (2013) Correction	2
A.3	Differences between voter file counts and official vote tallies	5
A.4	Turnout Over-reporting by race and ethnicity in the CCES surveys	8
A.5	State-by-state Differences in Overall Turnout	9
A.6	State-by-state Differences in White Turnout and Vote Share	10
A.7	State-by-state Differences in Black Turnout and Vote Share	12
A.8	State-by-state Differences in Hispanic Turnout and Vote Share	14
A.9	Surveys using CPS Data	16
A.10	Percent of eligible respondents over-reporting turnout by race	16

A.1 Constructing the Denominator for Turnout Rate Calculations

To calculate rates of voter turnout when using Catalist’s estimate of the number of voters, we construct a denominator that is derived from U.S. Census Bureau estimates of the population. As a base, our denominator uses data from the Census Population Estimates Program (PEP) and the related Census Intercensal Estimates.⁴ The PEP uses data on births, deaths, and migration to estimate the yearly resident population of the United States after each decennial Census, while the Intercensal Estimates are an interpolation of this information between the 2000 and 2010 Census. Each dataset provides estimates of the population by age and state for non-Hispanic White, Black, American Indian/Alaska Native, Asian, Native Hawaiian and Pacific Islander, and Hispanic racial/ethnic groups. For the analyses in this paper, we use estimates of the 18+ population total, non-Hispanic White, Black, and Hispanic populations.

We supplement the PEP and Intercensal Estimates with citizenship estimates from the American Community Survey (ACS).⁵ The ACS is a large-scale household survey, and uses the PEP/Intercensal estimates of the population by age, sex, and race/ethnicity as the principal source of information when weighting person-level estimates.⁶ The ACS asks a series of demographic, employment, and household characteristic questions, including asking respondents and their families about their nativity and citizenship. The ACS provides annual 1-year estimates of these characteristics for states, age groups, and racial/ethnic categories. For the voting-age population in each state, each year, and for each racial/ethnic group, we extract the smoothed ratio of the voting age population to the citizen voting-age population, then multiply this ratio by the PEP/Intercensal Estimates for the corresponding year.

As demographers have documented, the PEP, Intercensal, and ACS estimates are not free of errors. Indeed, the 2010 Census itself (which is the base for the PEP and Intercensal estimates) undercounted Black and Hispanic residents, and overcounted Whites likely in an attempt to com-

⁴<https://www.census.gov/programs-surveys/popest.html>. The PEP estimates we use in this paper are the “Vintage 2018” data (released in 2019).

⁵<https://www.census.gov/programs-surveys/acs>.

⁶<https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html>.

pensate for said undercounting.⁷ These errors make it difficult to know the precise composition or size of the voting-eligible population in the United States, or to develop reliable estimates of uncertainty for these quantities. However, for the purposes of this paper, we only seek to provide a comparable denominator to that used by the Current Population Survey (CPS) when contrasting rates of voter turnout. Importantly, the CPS sampling design and weighting strategy rely on the same Census data (currently ACS data, and prior to 2014 decennial Census data) we use for our denominator;⁸ at least some systematic errors in the quantities listed above would likely impact the CPS estimates as well. That said, the rotating panel design and non-mandatory nature of the CPS means that there are other administration-driven deviations that we cannot account for.

In a final effort to make our denominator as comparable to the CPS as possible, we adjust our estimate of the citizen voting-age population in two ways. First, since the CPS is based on a survey of respondents in November of the election year, and weighted to be comparable to national PEP estimates for November, we interpolate our citizen voting-age population estimates to November 1 of each election year. Second, the CPS sampling frame excludes individuals in institutional group quarters, who make up approximately 1.4% of the U.S. population. 60% of U.S. residents in institutional group quarters are in prisons or juvenile centers, and in every state aside from Vermont and Maine, imprisoned persons are not eligible to vote. We thus remove adult prisoners from our denominator, using data from the National Prisoner Statistics Program as coordinated by the Department of Justice.⁹ These estimates are provided by state and race/ethnicity.

A.2 Results After Hur and Achen (2013) Correction

Figures A.1 and A.2 reproduce main text Figures 2 and 3, respectively, but apply the turnout overreport correction discussed in [Hur and Achen \(2013\)](#). As the figures show, applying the [Hur and Achen \(2013\)](#) correction does not remove the bias in minority turnout versus white turnout,

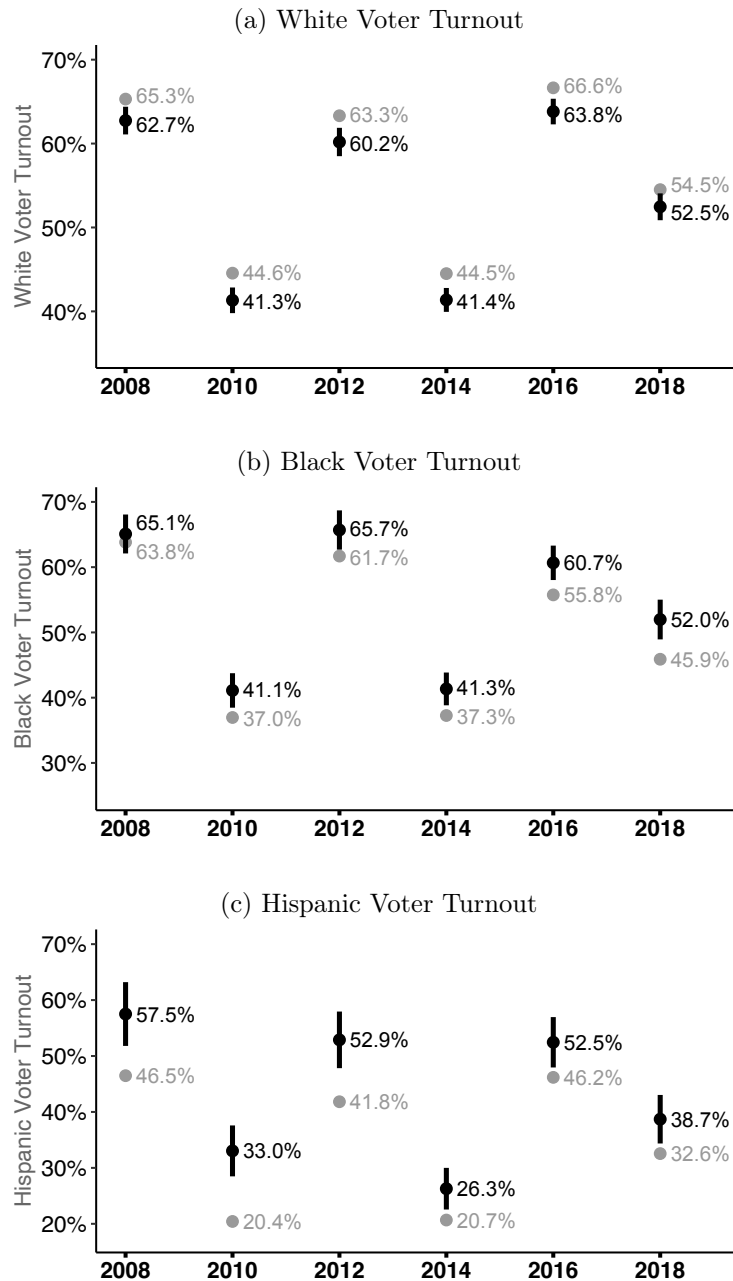
⁷see Table 1 in “Census Coverage Measurement Estimation Report: Net Coverage for the Household Population in the United States.” available at <https://www2.census.gov/programs-surveys/decennial/2010/technical-documentation/methodology/g-series/g03.pdf>.

⁸https://www.bls.gov/cps/sample_redesign_2014.pdf.

⁹<https://www.bjs.gov/index.cfm?ty=dcdetail&iid=269>.

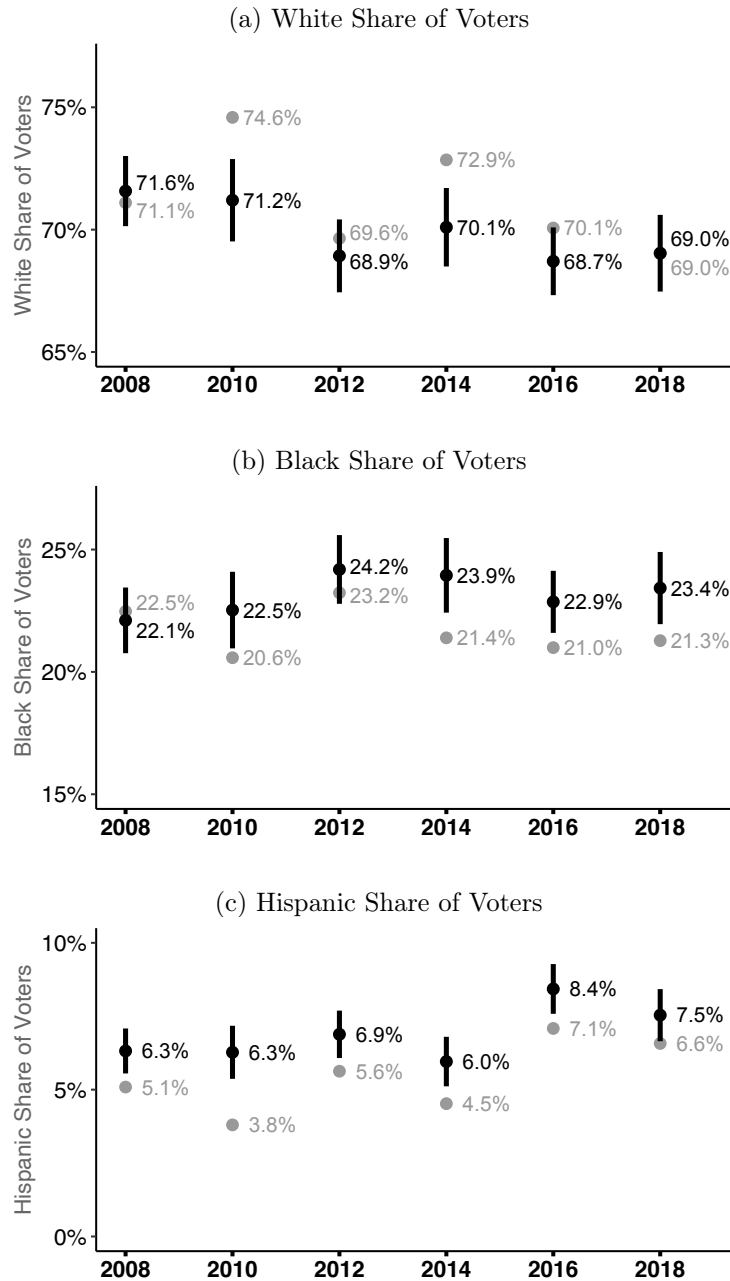
and in many cases, exacerbates the bias that overstates minority voting relative to white voting.

Figure A.1: Voter Turnout by Race, CPS with Hur and Achen (2013) Adjustment vs. Voter File Estimates



Note: Includes voters in six southern states with comprehensive data on race/ethnicity of registrants (AL, FL, GA, LA, NC, SC). Black points and 95% confidence intervals indicate estimates from the Current Population Survey, after applying the adjustment outlined in Hur and Achen (2013). Gray points are estimates of the number of voters by race/ethnicity in each election divided by the nonimprisoned citizen voting-age population by race, using data from Catalist and the U.S. Census Bureau.

Figure A.2: Share of Voters by Race, CPS with Hur and Achen (2013) Adjustment vs. Voter File Estimates



Note: Includes voters in six southern states with comprehensive data on race/ethnicity of registrants (AL, FL, GA, LA, NC, SC). Black points and 95% confidence intervals indicate estimates from the Current Population Survey, after applying the adjustment outlined in Hur and Achen (2013). Gray points are estimates of the share of voters by race/ethnicity in each election, using data from Catalyst.

A.3 Differences between voter file counts and official vote tallies

As we note in the paper, the number of voters on the voter files maintained by Catalist for each state are quite close to the actual number of votes tallied in the official vote counts collected by Michael McDonald. These tallies are shown in Table A.1. One concern raised by the small deviations that do exist is the possibility that the undercounts on the voter file may originate mostly from non-white precincts. Such a pattern would lead us to potentially understate the share of the electorate comprised of non-whites.

Table A.1: Deviation in voter file totals and official vote counts, by year

Year	Official	Catalist	Total Deviation	% Deviation
2006	12,387,448	12,220,415	-167,033	-1.3%
2008	22,761,646	22,612,750	-148,896	-0.7%
2010	14,964,889	14,890,618	-74,271	-0.5%
2012	23,095,990	22,934,619	-161,371	-0.7%
2014	15,520,376	15,598,483	78,107	0.5%
2016	24,822,710	24,680,514	-142,196	-0.6%
2018	20,992,345	20,749,153	-243,192	-1.1%

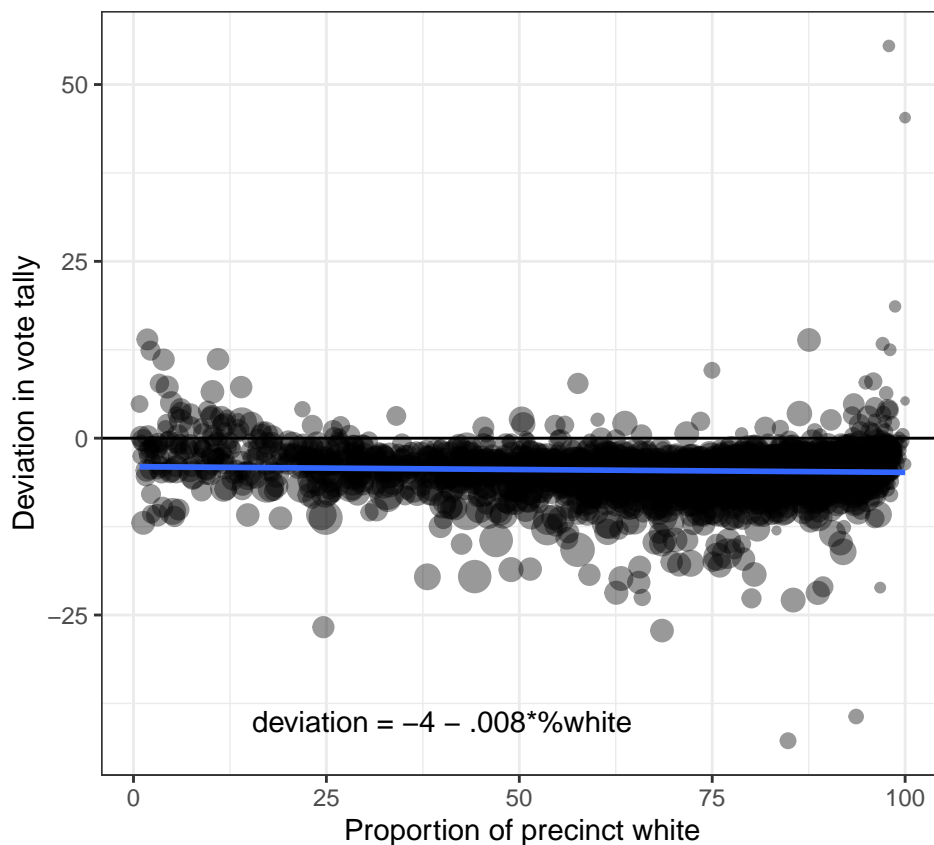
Note: Table shows the number of voters in the six southern states as reported by McDonald (<http://electproject.org>) compared to the number reported by the Catalist voter file data.

While we generally view the deviation totals to be too small to change our overall conclusions, we did investigate the question of whether these deviations may be especially likely to occur in non-white precincts. To do this, we acquired data from the North Carolina State Board of Elections (NCSBE) for the 2016 general election. Specifically, we acquired the official precinct level ballot count from the NCSBE website. We also acquired the North Carolina vote history file as well as a snapshot of the voter file from the day of the November 2016 election. Using these three data sources, we created two measures that we plot against each other in Figure A.3. On the y-axis is a measure identical to what is shown in the last column of Table A.1 – the percentage point difference between the official vote tally in each precinct minus the vote tally calculated from the state’s vote history file. On the x-axis we simply calculate the percentage of registered voters in each precinct who are non-Hispanic whites.

Figure A.3 shows no evidence for the possibility that undercounts on the voter file come mostly from precincts with more racial or ethnic minorities. Instead, there is a slight negative slope indicating that, on average, precincts that are more white tend to undercount votes to a slightly higher

extent than precincts with more voters who are people of color.

Figure A.3: Relationship between racial composition of precinct and deviation of voter file count from official vote tabulation in North Carolina, 2016



Note: Y-axis is the percentage point difference between official vote tally in each precinct and the number of voters on the voter file from that precinct for the November, 2016 general election. X-axis is the percent of registered voters in the precinct who are non-Hispanic white. N = 2,653 precincts. Data points are sized according to the number of registered voters in the precinct. All data come from the North Carolina State Board of Elections (<https://dl.ncsbe.gov/>).

As an additional robustness check, we evaluated whether the Catalist voter file data accurately reflects the number of voters by race. To do so, we acquired the official count of the number of white, Black, and Hispanic voters who voted in the 2012, 2014, and 2016 general elections as published by the Georgia Secretary of State's office. We then compared the share of voters who are white, Black, or Hispanic in the Georgia estimates to the Catalist estimates, removing individuals who are not from these groups.

Table A.2 provides the results of this comparison, where we see that in each election and for each group the Catalist data provides substantively similar estimates of the share of voters by race.

Table A.2: Share of Voters by Race in Georgia, Catalist vs. State Statistics

	<i>State Statistics</i>	<i>Catalist</i>
2014		
White	0.680	0.681
Black	0.309	0.307
Hispanic	0.011	0.012
2016		
White	0.671	0.672
Black	0.306	0.303
Hispanic	0.024	0.025
2018		
White	0.654	0.654
Black	0.321	0.320
Hispanic	0.025	0.026

Note: Only includes White, Black, and Hispanic voters. Proportions do not sum to 1 due to rounding. “State Statistics” based on counts of voters by race from the Georgia Secretary of State’s website.

A.4 Turnout Over-reporting by race and ethnicity in the CCES surveys

Table A.3 uses the CCES surveys from 2008 through 2018. Each of these surveys has been matched to voter files by Catalist and vote history has been appended. The table shows the rate at which no validated vote record was located among white, Black, and Hispanic respondents who claimed to have voted in the post-election wave of the survey.

Table A.3: Percentage of CCES self-reported voters without a validated vote record, by race

Year	Whites	Blacks	Hispanics	Year	Whites	Blacks	Hispanics
2008	21%	26%	30%	2014	29%	44%	48%
2010	19%	40%	38%	2016	30%	41%	47%
2012	23%	29%	35%	2018	25%	37%	40%

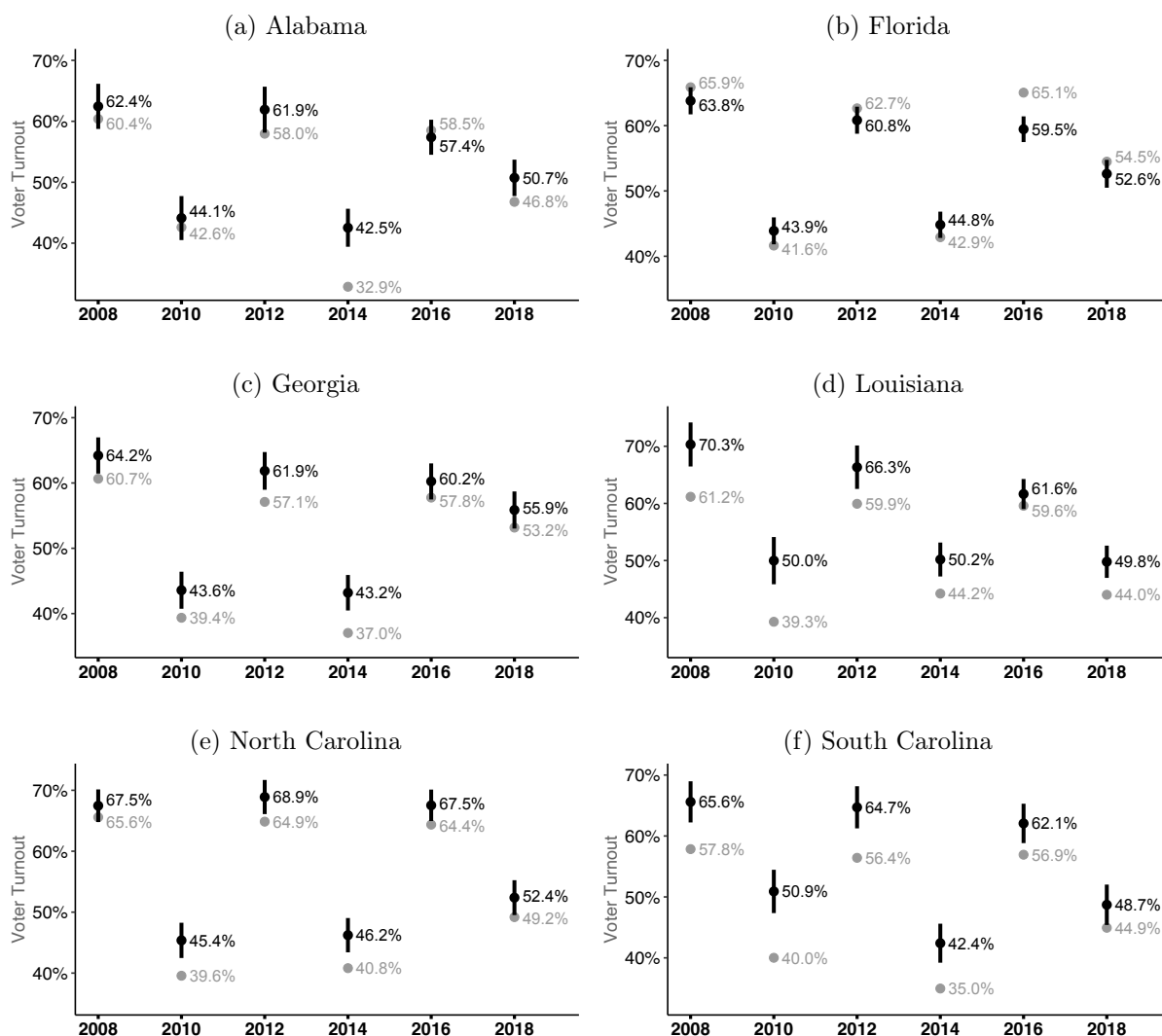
Note: Table shows the percentage of white, Black, and Hispanic respondents who claimed to have voted in the post-election survey for whom no validated vote record was found by Catalist. Sampling weights applied to ensure that sample is representative of American adult population.

Table A.3 shows that whites who claimed to have voted are consistently more likely to actually have a validated vote record when compared to African Americans and Hispanics. These patterns vary across years, but the differences are often quite large. For example, in 2010, Catalist failed to find a voting record for 20% of whites who claimed to have voted, but 41% of Blacks who claimed to have voted were unmatched to a valid vote record. Similarly large gaps in over-reporting were evident in 2014 and 2016, whereas the gaps were smaller (though still significant) in 2008 and 2012.

A.5 State-by-state Differences in Overall Turnout

In Figure A.4, we present a state-by-state version of Figure 1 in the main text. Alabama and Florida often, though not always, have CPS-estimated turnout rates that are closer to McDonald’s estimates than the other states. However, Florida’s 2016 turnout is the only instance where the CPS estimate appears to significantly understate overall rates of voting participation.

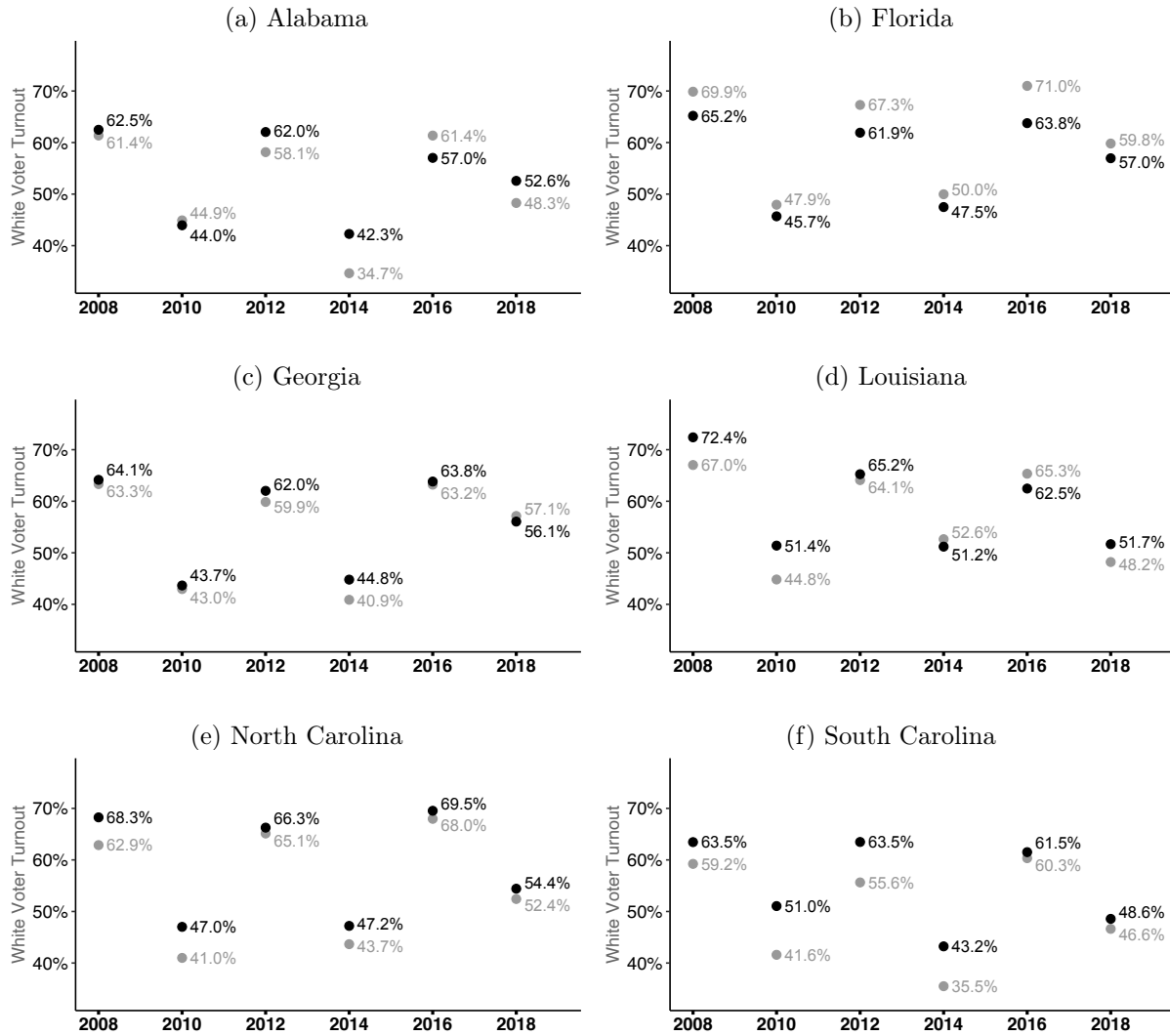
Figure A.4: Overall Voter Turnout, CPS vs. State Estimates



Note: Black points and 95% confidence intervals indicate estimates from the Current Population Survey. Gray points are estimates of the number of voters in each election divided by the nonimprisoned citizen voting-age population, as estimated by McDonald (2017).

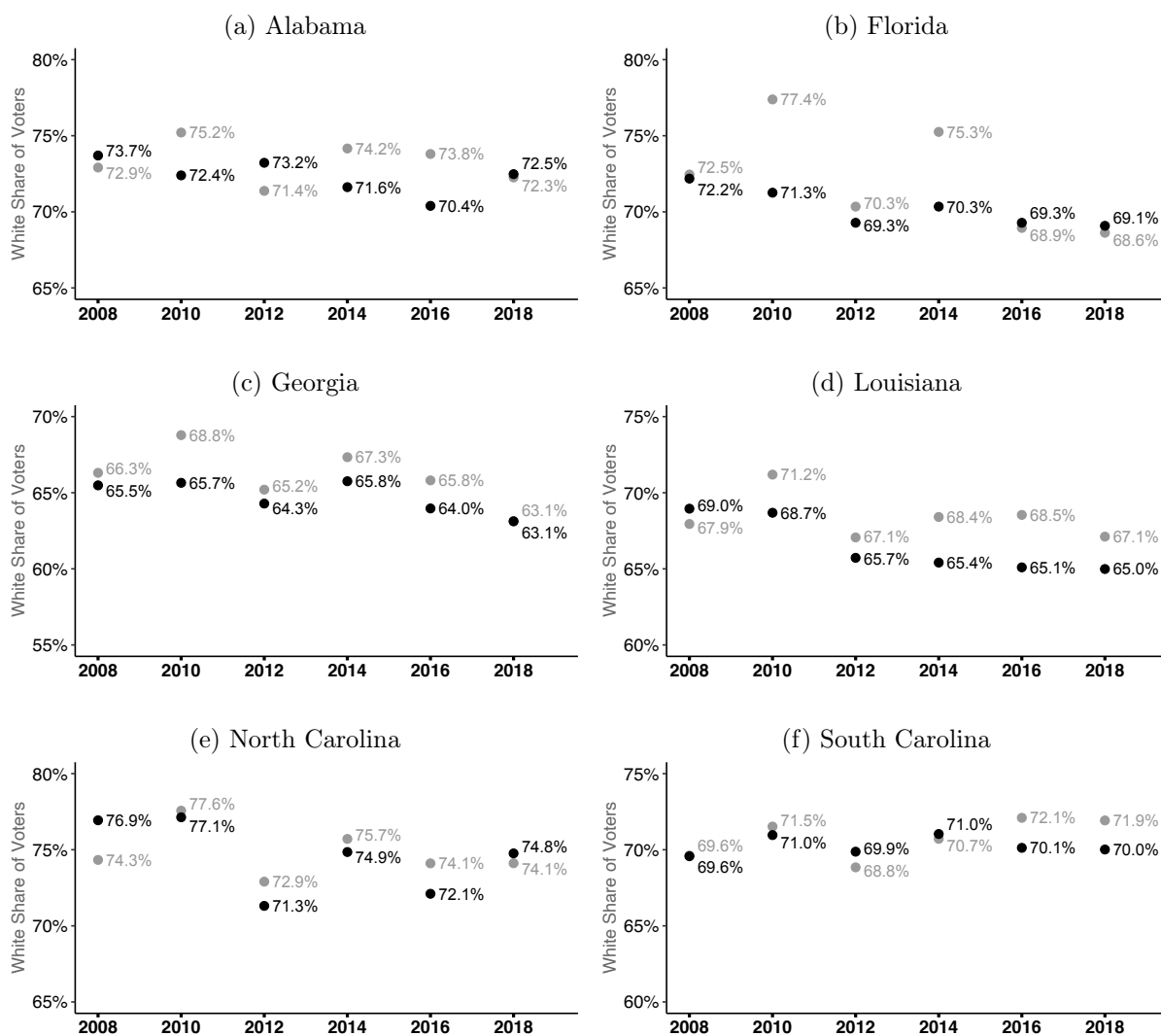
A.6 State-by-state Differences in White Turnout and Vote Share

Figure A.5: White Voter Turnout, CPS vs. Voter File Estimates



Note: Black points indicate estimates from the Current Population Survey. Gray points are estimates of the number of voters by race/ethnicity in each election divided by the nonimprisoned citizen voting-age population by race, using data from Catalyst and the U.S. Census Bureau.

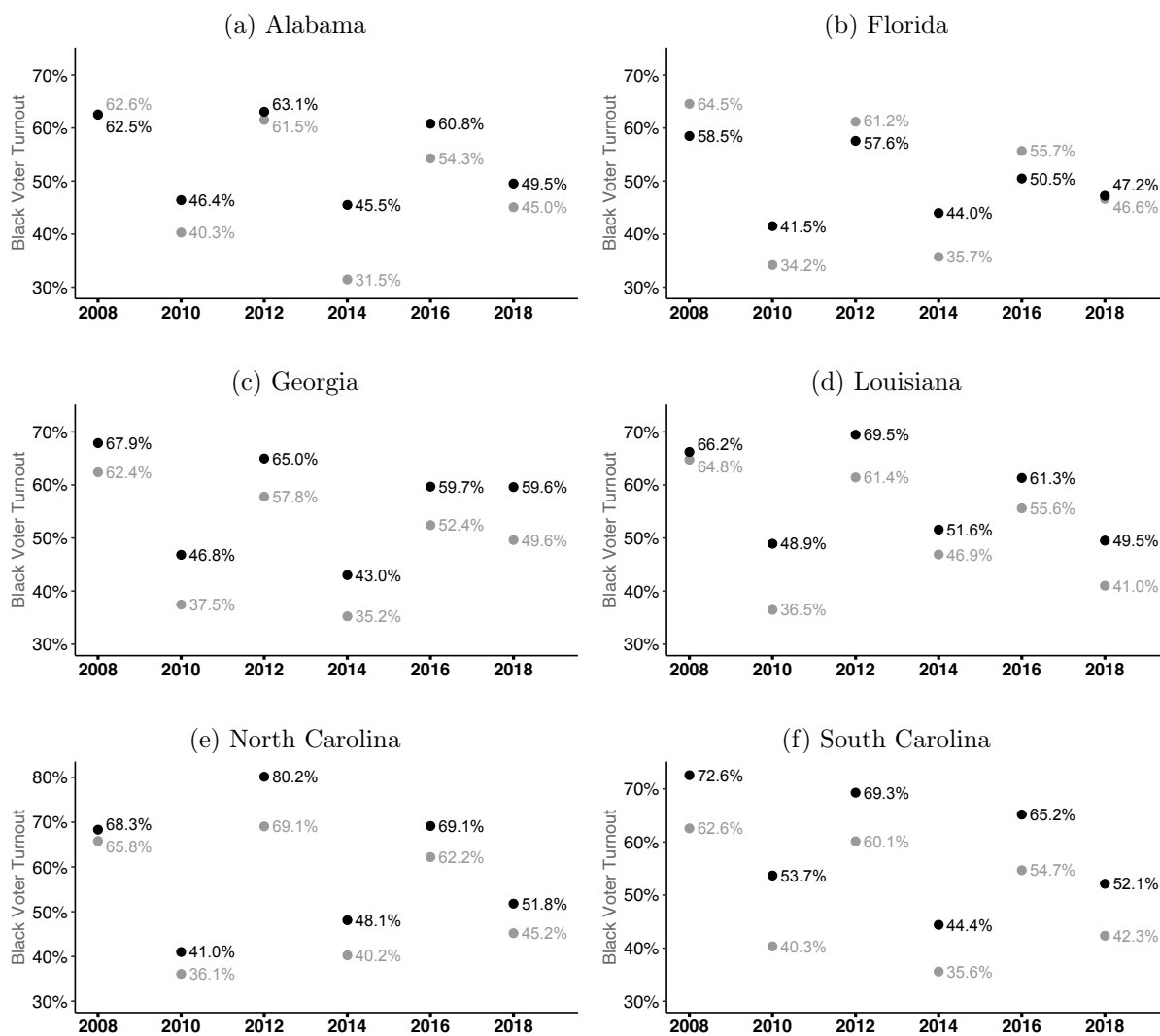
Figure A.6: White Share of Voters, CPS vs. Voter File Estimates



Note: Black points indicate estimates from the Current Population Survey. Gray points are estimates of the share of voters by race/ethnicity in each election, using data from Catalist.

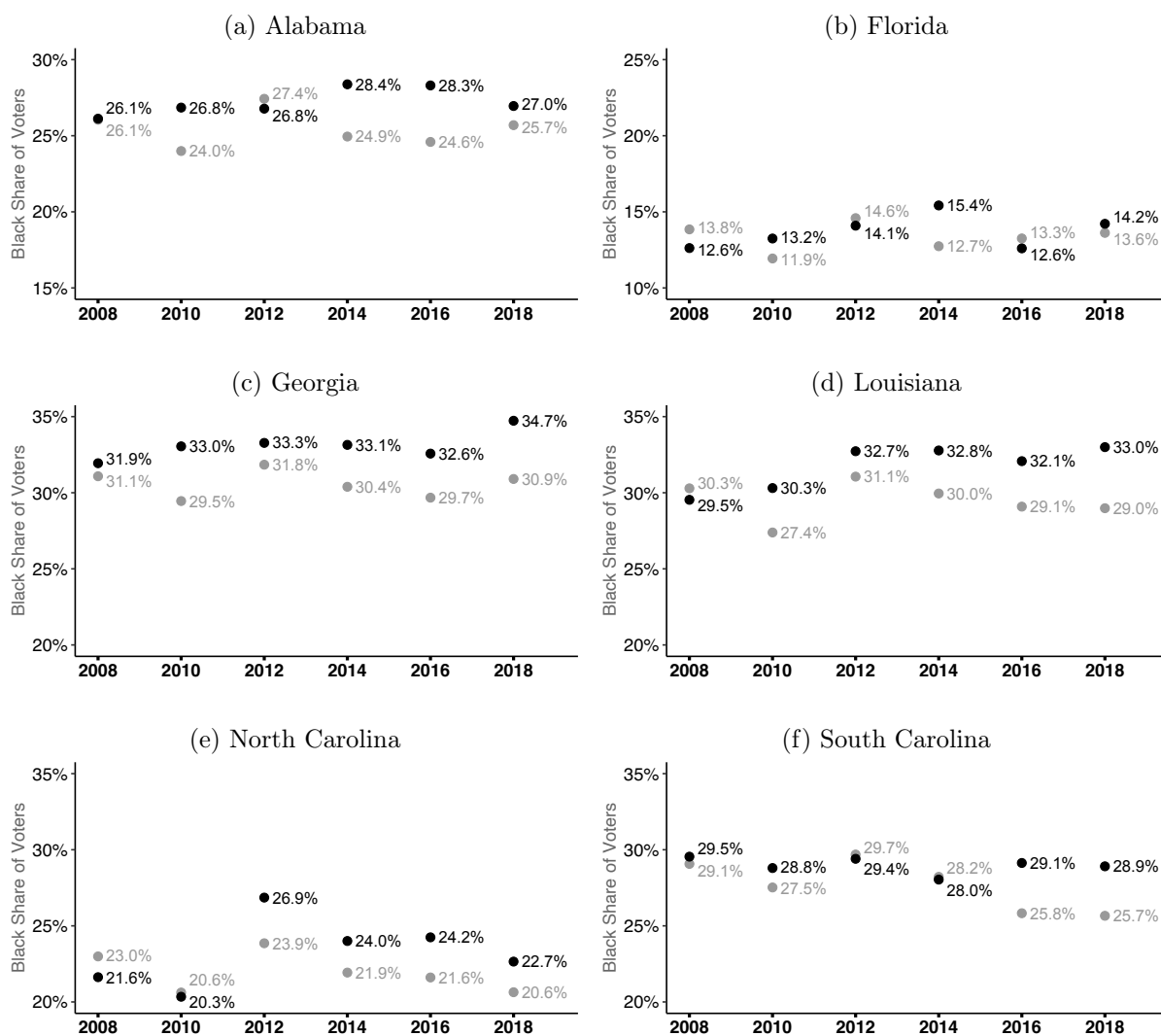
A.7 State-by-state Differences in Black Turnout and Vote Share

Figure A.7: Black Voter Turnout, CPS vs. Voter File Estimates



Note: Black points indicate estimates from the Current Population Survey. Gray points are estimates of the number of voters by race/ethnicity in each election divided by the nonimprisoned citizen voting-age population by race, using data from Catalist and the U.S. Census Bureau.

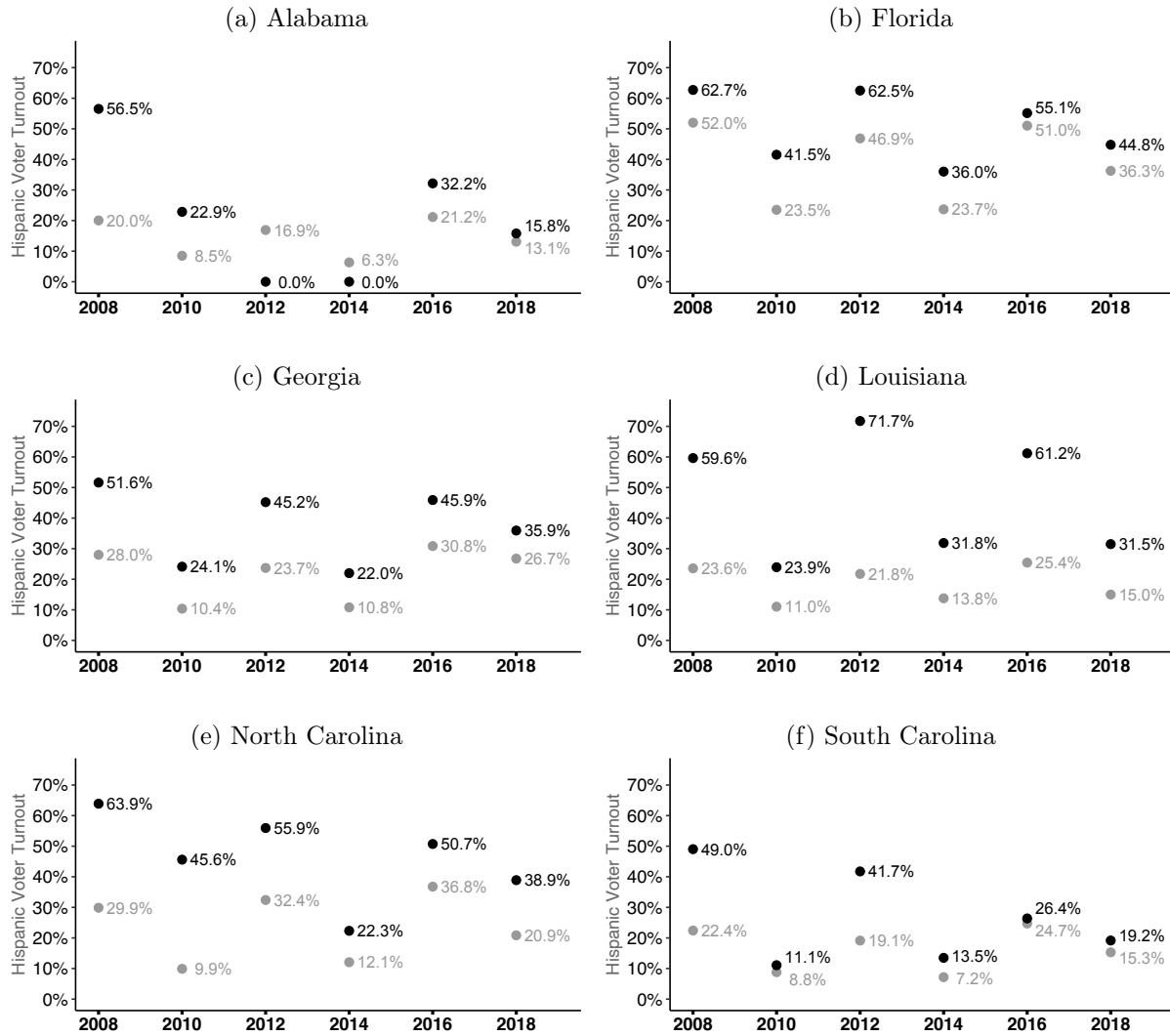
Figure A.8: Black Share of Voters, CPS vs. Voter File Estimates



Note: Black points indicate estimates from the Current Population Survey. Gray points are estimates of the share of voters by race/ethnicity in each election, using data from Catalist.

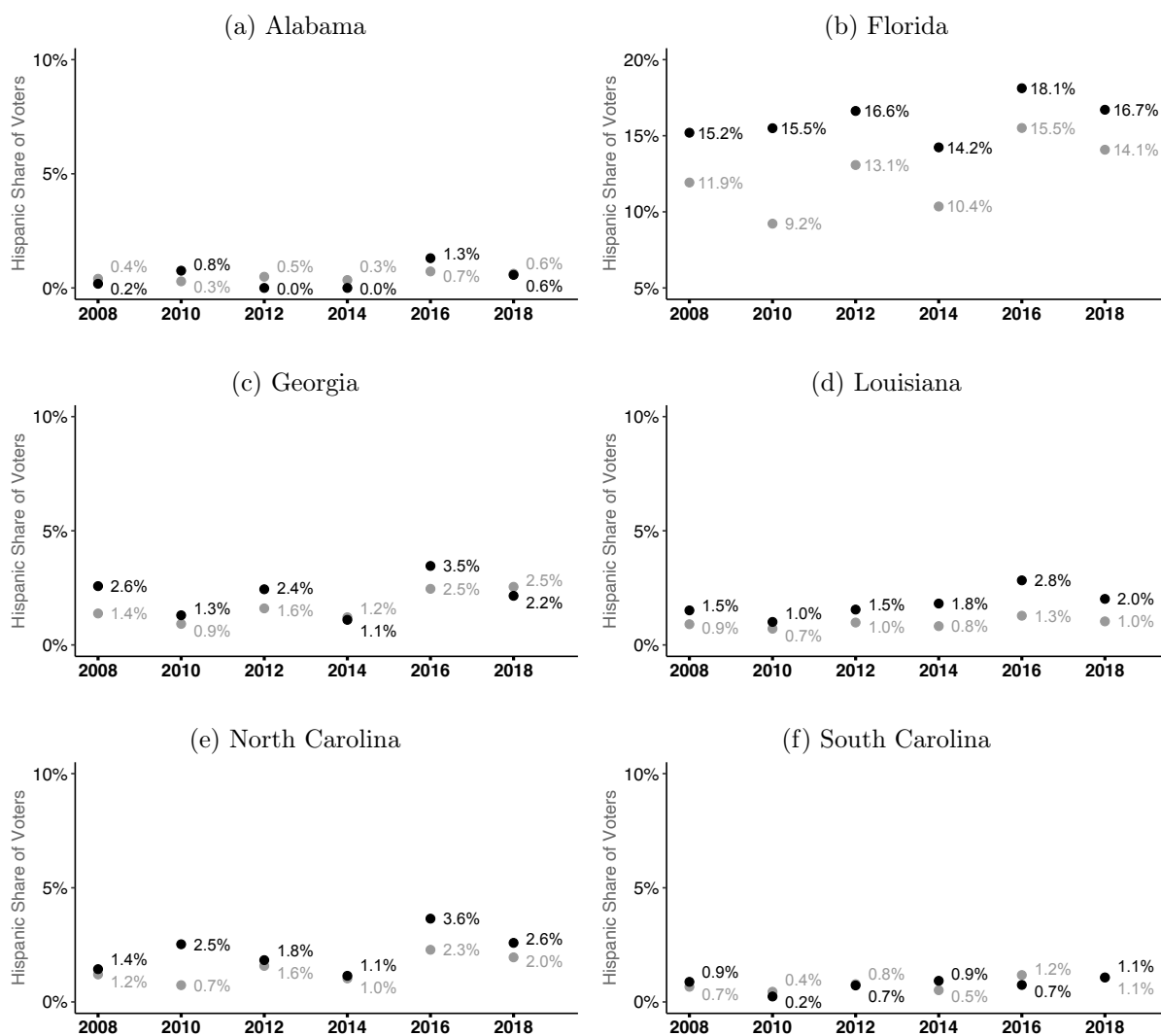
A.8 State-by-state Differences in Hispanic Turnout and Vote Share

Figure A.9: Hispanic Voter Turnout, CPS vs. Voter File Estimates



Note: Black points indicate estimates from the Current Population Survey. Gray points are estimates of the number of voters by race/ethnicity in each election divided by the nonimprisoned citizen voting-age population by race, using data from Catalist and the U.S. Census Bureau.

Figure A.10: Hispanic Share of Voters, CPS vs. Voter File Estimates



Note: Black points indicate estimates from the Current Population Survey. Gray points are estimates of the share of voters by race/ethnicity in each election, using data from Catalist.

A.9 Surveys using CPS Data

Table A.4: Major election surveys that weight or match to CPS, 2016-2020

Survey	Weight to CPS?	Weight to CPS/VRS?	CPS/VRS for target sample?
American National Election Study (2016)	Yes	No	NA
Associated Press VoteCast (2018)	Yes	Yes	NA
Cooperative Congressional Election Study	No	No	Yes
National Election Pool Exit Polls (2018)	Yes	Yes*	NA
New York Times/Siena	Yes	Yes*	NA
Pew America Trends Panel	Yes	No	NA
VOTER Survey	No	No	Yes

Note: * denotes surveys that use the CPS to weight to educational attainment of voters only.

A.10 Percent of eligible respondents over-reporting turnout by race

Table A.5: Percentage of CCES respondents over-reporting turnout, by race

Year	Whites	Blacks	Hispanics	Year	Whites	Blacks	Hispanics
2008	19%	26%	23%	2014	25%	34%	35%
2010	14%	29%	24%	2016	28%	37%	43%
2012	22%	27%	32%	2018	21%	29%	31%

Note: Table shows the percentage of white, Black, and Hispanic respondents who said that they voted but did not according to records from Catalist. Sampling weights applied to ensure that sample is representative of American adult population.