Ground Truth Validation of Survey Estimates of Split-Ticket Voting with Cast Vote Records Data*

Alexander Agadjanian† Jonathan Robinson‡

August 26, 2019

Abstract

From signaling trends in nationalization and partisanship to clarifying preferences for divided government, split-ticket voting has received copious attention in political science. Important insights often rely on survey data, as they do among practitioners searching for persuadable voters. Yet it is unknown whether surveys accurately capture this behavior. We take advantage of a novel source of data to validate survey-based estimates of split-ticket voting. Cast vote records in South Carolina (2010-18) and Maryland (2016-18) provide anonymized individual level choices in all races on the ballot for every voter in each election, serving as the ground truth. We collect an array of public and private survey data to execute the comparison and calculate survey error. Despite expectations about partisan consistency pressures leading to survey underestimates, we find that surveys generally come close to the true split-ticket voting rates in our set of races. Accuracy varies, but notably is more consistent for split-ticket voting in a given dyad of national races (e.g., President vs. U.S. House) than in one with state races, as the former is often of greater interest in research and practice.

*Prepared for presentation at the Annual Meeting of the American Political Science Association, Washington, DC, August 29 – September 1, 2019. An earlier version of this paper was presented at the 2019 American Association for Public Opinion Research Conference (Toronto, Canada) on May 17, 2019. Thanks to Shiro Kuriwaki for making cleaned South Carolina cast vote records available to us, as well as for helpful guidance and discussion.

†Research Associate, MIT Election and Data Science Lab. Email: aagadjanian96@gmail.com

‡Lead Research Scientist, Catalist, LLC. Email: jrobinson@catalist.us

1 Context and Motivations

Split-ticket voting has long been an object of fascination for political observers. The topic has enjoyed a rich tradition in political science, having been called "one of the most often analyzed topics in the study of American political behavior" (Mulligan, 2011). Although approaches with aggregate election data do receive use (e.g., Burden and Kimball, 2002), individual level analysis often opens more avenues for research questions and avoids ecological fallacies, leading many researchers to turn to surveys for data. Academics have used survey-based split-ticket measurements to study preferences for divided government and policy balancing (Alvarez and Schousen, 1993; Beck et al., 1992; Born, 1994; Garand and Lichtl, 2000; Lacy et al., 2019; Lewis, 2001; Sigelman, Wahlbeck and Buell Jr., 1997), partisanship and polarization (Campbell and Miller, 1957; Jacobson, 2013), the effects of news on voting (Moskowitz, 2018), and the psychological and contextual determinants of split-ticket voting itself (Davis, 2015; Davis and Mason, 2016; McAllister and Darcy, 1992; Mulligan, 2011).

Split-ticket voting has also drawn considerable public interest, and especially so in recent years. Many accounts have highlighted an over time decline in split-ticket voting (Abramowitz and Webster, 2016), with news outlets frequently devoting attention to the trend (Brownstein, 2016; Kilgore, 2018; Phillips, 2016; Skelley, 2018). People often interpret this growing relationship between election outcomes across offices as a sign of increasing nationalization and widening scope of partisanship's influence. Lastly, political practitioners take great interest in split-ticket voting as well, as identifying individual level cracks in partisanship can aid voter targeting and persuasion efforts.

With all the surrounding interest and stake in split-ticket voting, it's important to not only study it, but also scrutinize the tools used to do so. We focus on surveys in this paper. Scholars have long wondered about the validity of surveys for the study of split-ticket voting (Burden and Kimball, 2002; Moskowitz, 2018). The common way to assess error in vote choice survey data is to compare a survey vote distribution for an office to the actual distribution for the office from an election. Because election results only record vote totals for a single office, the actual distribution of vote totals from different offices in combination of course cannot be deduced. This puts actual split-ticket rates and survey error calculation out of reach. The accuracy of survey split-ticket measurement remains unclear.

2 New Data and Directions

We remedy this uncertainty by using a relatively novel data source that provides true levels of split-ticket voting: cast vote records (CVRs). Referred to as ballot image logs in other contexts, CVRs have been employed by researchers before. Most recently, Kuriwaki (2019*a*) standardizes individual CVRs from South Carolina for all general elections from 2010 to 2018 to comprehensively analyze split-ticket voting, uncovering defection from party choices for national offices further down the ballot. Other scholars estimate voter ideologies from CVRs to examine the role of spoiler candidates in elections (Herron and Lewis, 2007) and understand the policy basis of candidate support (Lewis, 2001).

Equipped with more comprehensive data than in the past—complete records spanning two states and seven unique elections—we take a different direction, using CVRs as ground truth benchmarks of split-ticket voting, against which we compare survey estimates. Several reasons motivate this effort. First, we take advantage of a unique opportunity to validate a ubiquitous survey measure that has not previously been inspected. Second, we seek to attain a finer understanding of survey error. Fine-grained validations of survey data are sparse, but are useful for clarifying the reliability of survey data (Ansolabehere and Hersh, 2012; Rogers and Aida, 2014) or how to improve it (Kennedy et al., 2016; Meyer and Mittag, 2019). Third, we address potential for partisan consistency bias in surveys, which could lead to underestimates of split-ticket voting and thus an exaggerated sense of partisanship.

3 Expectations about Survey Error

Beyond simply validating a survey measurement for its own sake, we have theoretical reasons to suspect survey error. These apply to both vote choice reports from pre-election surveys (mainly from media and campaign polls) and post-election reports (academic surveys like the CCES). First, typical reasons for survey error for any one office vote distribution might compound when looking at two offices in combination. For pre-election polls, preference and turnout decisions can change between the time of a survey and a future election. For post-election polls, overstated support for winning candidates (Wright, 1993), question wording effects (Box-Steffensmeier, Jacobson and Grant, 2000), and forgetfulness (Wright, 1990)—especially as time after the relevant election day increases—can impair recall accuracy. Voters appear to remember past vote choice for a high profile race like the presidential one well (Rivers and Lauderdale, 2016), but this might not apply to lower level races.

We also theorize error could specifically run in the direction of understating split-ticket voting. Pressures for partisan consistency and cognitive dissonance avoidance could affect vote reports across different offices. In pre-election surveys, respondents may not yet be knowledgeable about candidates, assume they will toe the party line, and respond to vote intent questions accordingly. Post-election, survey-takers may similarly assume partisan consistency when asked for vote choice on a range of offices. The need to recall vote choice may exacerbate a partisanship influence here through forgetfulness: as voters struggle to recall their vote across several offices further out from the election date, they might assume they voted their party consistently, particularly for less salient races. This could also reduce dissonance among partisans unwilling to reveal defection. Some surveys, like the CCES, only provide party labels (not candidate names) for vote choice questions in state office elections—this might further obscure vote recall and even encourage partisan consistency.

4 Data and Methods

We now discuss the two data sources we focus on—CVRs and surveys—and the approach to analysis we take to investigate accuracy of survey split-ticket rates.

4.1 Cast Vote Records

We calculate ground truth estimates of ticket-splitting from CVRs, anonymized individual level ballot data files that record votes but are not linkable to other data sources. When a voter casts their ballot in an election, two things happen. Their poll book, with their voter ID number and other personal identifiable information, is sent to their state's voter registration database in order to generate vote history data, but is separated from the voter's actual ballot. This ensures the secret ballot—the central pillar of our electoral system in the US. The votes voters cast in their ballots get sent to the cast vote record (or ballot image log) database where individual level votes are stored.

Our CVR data only comes from certain elections in two states: South Carolina and Maryland. For South Carolina, we collect these records for all general elections from 2010 to 2018 (five in total), taken from the South Carolina State Election Commission's website (Kuriwaki, 2019*a*). For Maryland, we obtain CVRs for 2016 and 2018, drawn from the Maryland State Board of Elections through an open/public records request made of files generated by ClearBallot, a firm that provides software for auditing election tabulation systems using CVR data. CVR data is growing in popularity in academic research but has also seen a rise in interest among election administrators as an essential part of risk-limiting audits that states conduct after elections to verify results and examine for foul play (Kuriwaki, 2019*b*).[1]

CVR data come in multiple forms depending on the vendor the state uses to tabulate

[1] In this study, we focus only on CVRs in elections for which we also have survey data to compare against, but the records allow for examination of races all the way down to County Sheriff in the case of South Carolina and City Council in the case of Maryland. See Kuriwaki (2019*a*), who takes full advantage of this wealth of individual vote choice data.

its cast vote records and how centralized the system is for vending them to the public, but always at the level of an individual voter. These data have minimal identifiable information. Information on geographies allows researchers to define the contest being voted on; these include very granular districts (e.g., city council), but mostly involve precinct as the most granular level of geography and county as the least granular. CVRs do not come in a readily usable format. Candidate names or ballot measure results populate the fields, which must be converted into candidate party through auxiliary information (e.g., aggregation of public election results). For South Carolina, we worked with Shiro Kuriwaki's Ballot Image Log Project to use a common set of standardized data to ensure consistency in coding and findings across research projects using this new data source.

4.2 Surveys

We aimed to find surveys that asked about vote choice in as many races on the ballot as possible for the seven unique state-year elections covered by the CVRs (South Carolina for 2010-2018 and Maryland for 2016-2018). Sources for our data range from public academic datasets, public data accessed through the Roper Center for Public Opinion Research, and personal contacts that led to disclosure of data from public pollsters (e.g., Winthrop, Washington Post-UMD) and private ones (Civis Analytics). For the latter portion, some data sources retain their identity, while others ask that we redact their names; they appear as "Anonymous" or "Pollster A/B" in our presentation of data and results. For one anonymous pollster, we have data for separate surveys from both July and October of that election year. These are distinguished in the table and in later figures.

Table 1 presents our amassed survey data, with abbreviations for the office that each data source asks vote choice for. The paucity of state polling in recent elections made compiling a large dataset difficult, but we still gather at least some data in every unique election. Survey timing varies, as survey field dates occur in the months preceding the election or in the month following it. In several cases, the Cooperative Congressional Election Study (CCES)

State	Sources	Year
South Carolina	CCES: US Sen, Gov, US House, AG, SoS, State House Winthrop: US Sen, Gov Anonymous: Gov, US House – CD5	2010
South Carolina	CCES: Pres, US House, State Sen, State House Winthrop: Pres, US House – CD7	2012
South Carolina	CCES: US Sen, US Sen Special, Gov, US House, SoS, State House Winthrop: US Sen, US Sen Special, Gov	2014
South Carolina	CCES: Pres, US Sen, US House, State Sen, State House Civis: Pres, US Sen	2016
South Carolina	CCES: Gov, US House, AG, SoS, State Sen, State House Civis: Gov, US House, State Sen, State House Anonymous: Gov, US House – CD1 July & Oct. Surveys	2018
Maryland	CCES: Pres, US Sen, US House Civis; Washington Post; Anonymous: Pres, US Sen	2016
Maryland	CCES: US Sen, Gov, US House, AG, State Sen, State House Civis: US Sen, Gov, US House, State Sen, State House Anonymous: US Sen, Gov, AG	2018

Table 1: Summary of Collected Survey Data

was large enough to include sizable samples for South Carolina and Maryland, and expands our coverage significantly. The CCES was the most comprehensive source in terms of races asked about, covering not only national races but also state offices such as Attorney General, Secretary of State, State Senate, and State House.[2] As the table also shows, we also include survey data that only covers a certain congressional district. In later analysis, we calculate the CVR split-ticket rate for these singular congressional districts in order to make for a direct comparison.

[2]We initially considered including CCES split-ticket estimates for only validated voters in addition to overall estimates. These vote validated estimates proved almost identical to the overall ones. Given this closeness and concern about double-counting CCES estimates when distributions of survey error are calculated, we drop the CCES vote validated split-ticket rates from our analysis. Nevertheless, it is worth noting this similarity, which itself is interesting and perhaps a testament to the accuracy of survey-based vote choice—even with the inclusion of non-voters in those estimates.

4.3 Standardizing and Measurement

We apply a few modifications to both the CVR and survey data before carrying out split-ticket calculations. First, we only consider three possible vote choice options for any office: Democrat, Republican, or a non-major party. This approach excludes abstention, a decision prompted by uncertainty over how to best capture abstention in surveys.[3] Second, we do not consider data for 2018 Maryland State House races, which include multi-member house districts. Our survey data asks vote choice as if all voters encounter just one decision, making it unclear how to bridge the survey and CVR data for comparison.

Lastly, we only consider contested races—those that involve both a Democratic and Republican candidate. This is in line with other similar work (Kuriwaki, 2019*a*), and recognizes that split-ticket voting that involves uncontested races might not reflect true partisan defection so much as choosing the only option on the ballot that happens to be the out-party in some cases. Ignoring vote choice for uncontested races has sizable reductions in the sample size we attain in some cases, presaging a few underpowered calculations in later analysis. Rates of uncontested races are as follows: SC US House - 28% (2012, 2014); SC State Senate - 76% (2012), 85% (2016); MD State Senate - 32% (2018); SC State House - 70% (2010), 84% (2012), 77% (2014), 74% (2016), 64% (2018); MD State House - 25% (2018). This amounts to a 2% decrease in valid US House vote choice reports, a 41% decrease in valid State Senate vote choice reports, and a 37% decrease in valid State House vote choice reports (note that respondent home districts are not necessarily distributed in proportion to actual districts).

To measure split-ticket voting in both the CVR and survey data, we focus on a "top-of-ticket" approach: comparing vote choice for an office at the top of the ticket (e.g., the Presidential race in 2012 and 2016) to each lower ballot office. This method produces office

[3] This problem is most relevant to pre-election polls, where we do not know whether a lack of reported vote choice indicates abstention or indecision. If we treat non-voting as abstention, then this could artificially inflate split-ticket rates, as we end up treating undecided voters and those unsure about their turnout status as voters making a choice. These exist in our data, especially in polls further out from election.

dyads.[4] For every office dyad that we consider in our set of elections, voters are coded as having a split-ticket for any of the following combinations: Democrat × Republican, Democrat × Other, Republican × Democrat, Republican × Other, Other × Democrat, and Other × Republican. The denominator includes all of these combinations in addition to Democrat × Democrat and Republican × Republican, the straight-ticket options. We use survey weights wherever available in the public and private data, or calculate our own weights in cases that they are not included.

In total, we derive ground truth CVR estimates for 28 unique office dyads across our different state-years. As mentioned before, because a few surveys only cover single congressional districts, we also calculate CD-specific CVR rates, increasing our total to 31 unique dyads. After applying the previous modifications that removed parts of the data, the resulting dataset is 45 survey-based split-ticket estimates that are unique for a given state, year, and office dyad. The combined respondent count (represented as the weighted total) for all the samples from which we calculate split-ticket rates is 29,340. After we link the survey estimates to the CVR estimates, we compute "net error" by subtracting the CVR split-ticket rate from the corresponding survey split-ticket rate (Survey - CVR). Negative values signify survey underestimation of the true split-ticket rate for this measure. We also consider absolute error (taking the absolute value of our net error measure) in other parts of the analysis for the sake of easier interpretation.

[4]This contrasts from a other approaches such as calculating the split-ticket rate for the entire ballot or a mixed ballot rate. Based on the number of offices asked about, our survey data is too limited to accomplish these alternatives. Nevertheless, our measurement choice is acceptable given that past researchers, using survey data with similar limitations, largely used office dyads as well—our survey validation seeks to speak to this past work.

5 Results

5.1 Visualizing Survey and CVR Rates

Figure 1 displays all 45 split-ticket surveys rates along with the CVR estimates for the same races. These are organized by seven subplots indicating each unique state-year election in our data. The combination of subplot title (the text following "Top of the ticket:") and an x-axis value name make up the office dyads of interest. CVR estimates always appear in black-colored points, except in cases where they are specific to a congressional district, corresponding to a survey with the same coverage. The CD-specific CVR rates are colored with different shades of grey, as indicated on the graph legend. For reference, at the top of each subplot, we display the true split-ticket rate (the CVR rate). Different colors also denote different survey sources from which the split-ticket rates come. As noted before, the CCES is the most common source here, appearing in purple. 95 percent confidence intervals are shown for all the survey split-ticket estimates.

In a handful of cases, such as President × State Senate in SC 2016 and US Senate × State House in SC 2014, the confidence intervals are very large, making for uncertain estimates. After removing uncontested races, this significantly reduced the sample size for some races and thus create a few highly uncertain estimates. Interpreting these numbers should be approached with caution. We leave these estimates in the data as a matter of completeness. In later analysis, we try to account for these more uncertain estimates.

The CVR split-ticket rates range from 5 percent to 27, with a median of 12 and mean of 13. The survey rates vary a bit more, going from a minimum of 4 to as high as 42, with a median of 13 and mean of 15. As one example, consider the Maryland 2018 subplot with the US Senate race at the top of the ticket. In this election, a Republican candidate won the Governor's race amid a strong Democratic stronghold otherwise, heightening expectations of split-ticket voting. As the CVR estimate shows, 27 percent of voters split their tickets, straying from consistent partisan voting between their US Senate and Governor's race selections.

This represents the highest true split-ticket voting rate in our data. The CCES comes closest to the true rate, as the split-ticket rate there is 24 percent. Civis Analytics (turquoise green-colored point) at 34 percent and our anonymous "Pollster B" (lime green-colored point) at 42 percent miss the mark by a greater amount.

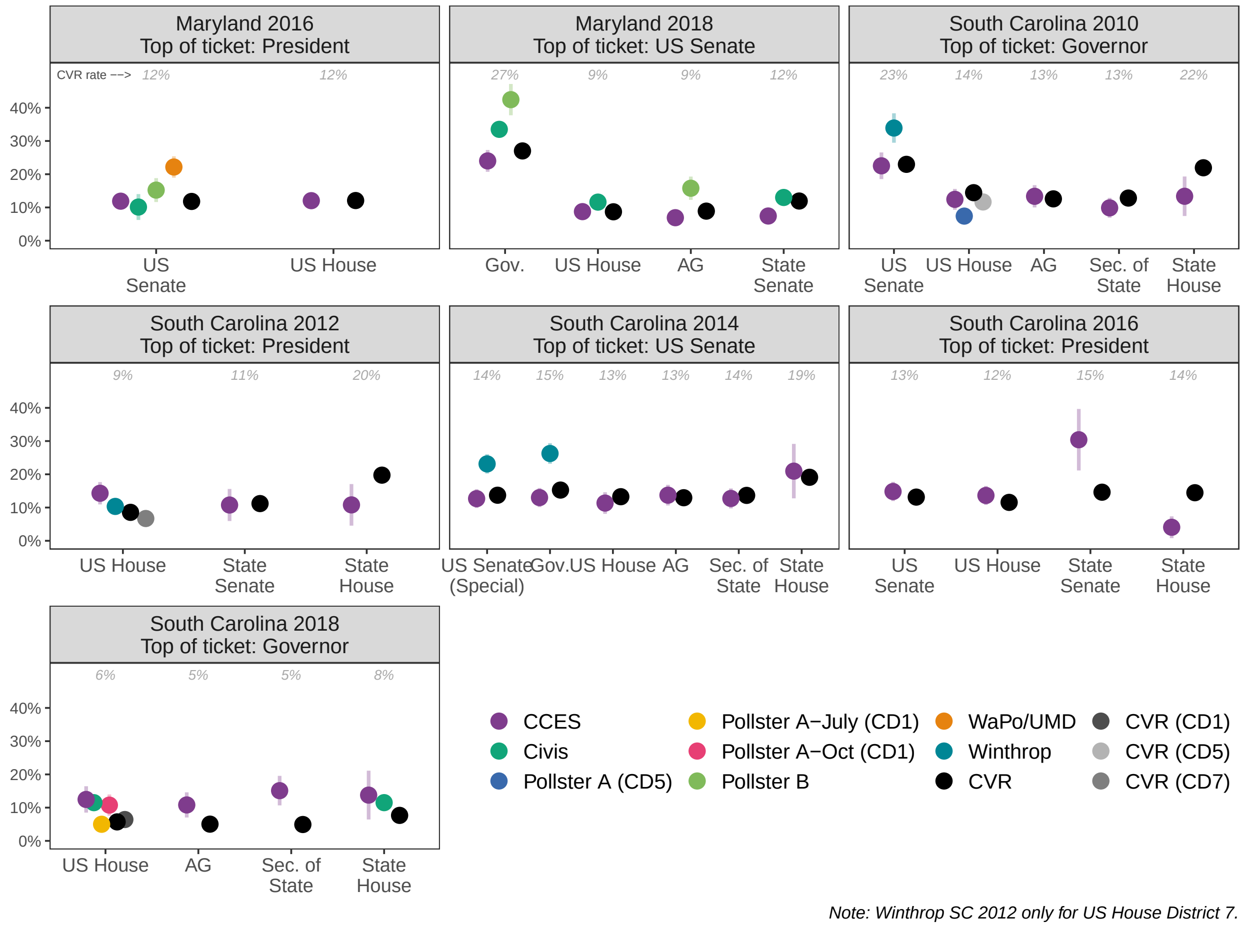

Figure 1: Top of Ticket vs. Lower Level Split-Ticket Rates (Surveys and CVRs)

This initial look at the data contains a lot of information, but from a qualitative glance, it appears that surveys mostly do well in approximating the true split-ticket voting rate. The degree of error varies, but does not appear to be consistently in the direction of over- or under-estimating the true rate. Figure A.1 in the Appendix plots the net error rate for each estimate to pinpoint the exact deviations. Notably, cases that reveal sizable survey error, such as Governor × State House for SC 2010, President × State House for SC 2012, and President × State Senate for SC 2016, tend to have much larger confidence intervals in Figure 1 and thus less certainty. Extreme cases like these might arise more because of sample size issues than with the quality of the survey estimates themselves. In most other contexts, survey rates appear fairly accurate—in other words, the colored dots come close to the black/grey colored dots, the ground truth.

5.2 Distribution of Survey Error

We now move into summarizing the rates of survey error in a more systematic way. Figure 2 presents the distribution of net error for our 45 split-ticket estimates (the values that appear in Figure A.1). Again, error is calculated as survey rate minus CVR rate, such that negative values connote survey underestimation of the true rate and positive values show overestimation.

Survey error is roughly normally distributed. We also display a few key statistics on the graph. The mean net error is +2.14, which indicates that, on average, surveys overestimate the true split-ticket rate by a few points. This is generally small, and notably the error runs in the opposite direction of what we expected based on concerns about partisan consistency pressures. Overall, 60 percent of our calculated survey rates *overestimate* the true rate to some degree. The first quartile is -1.69 and the third quartile is 5.78, reinforcing the greater overestimation pattern. The standard deviation is 5.87, signifying that a good amount of variability is present in the level of survey error. The root mean square error is 6.18. Lastly, in the Appendix, we include a plot of the distribution of absolute survey error (Figure A.2); the mean error here is 4.66 points and the standard deviation is 4.11.

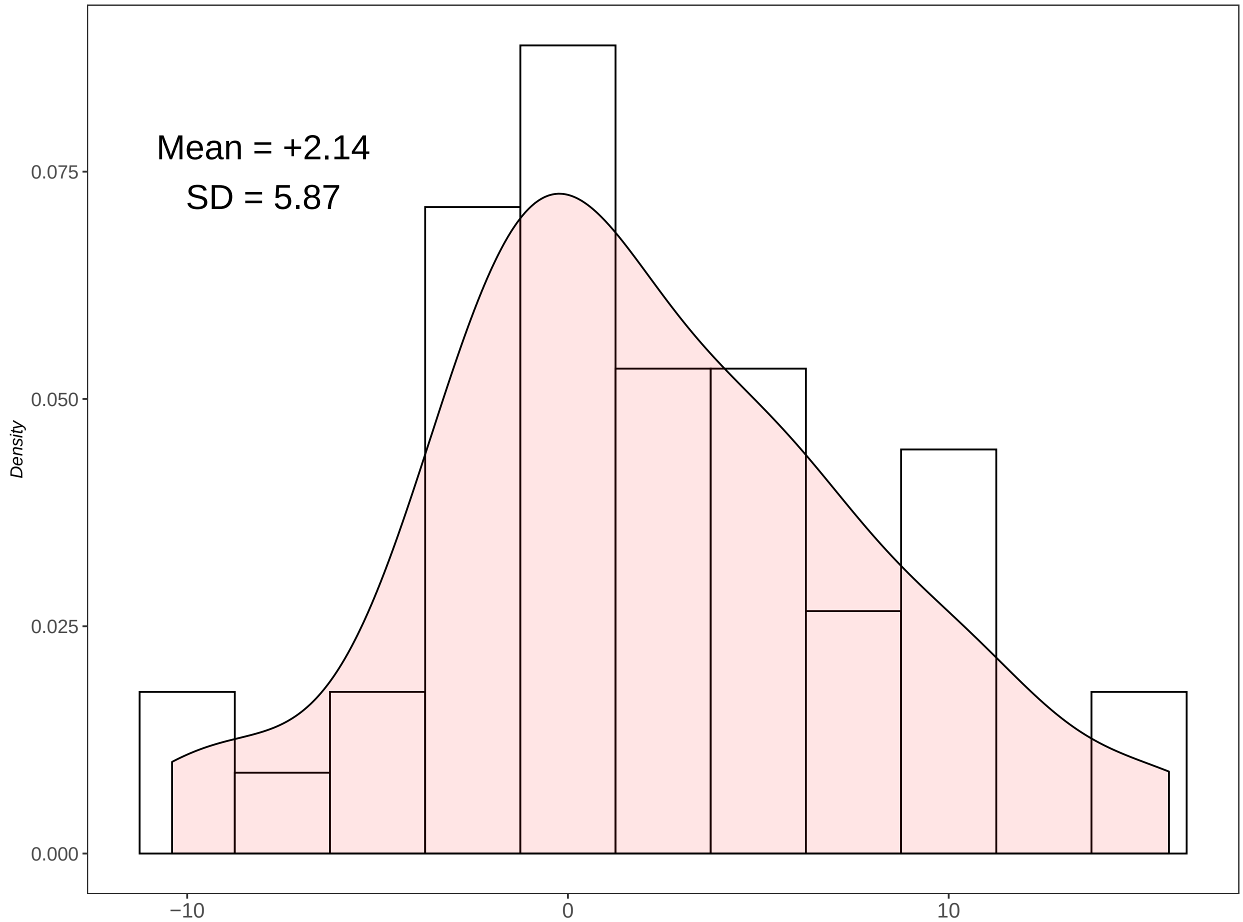

Figure 2: Distribution of Net Error (Survey - CVR)

5.2.1 Office Level Breakdown

We also choose to highlight one other interesting breakdown of survey error—by the levels of offices (national or state) involved in the office dyads. Split-ticket voting across office levels has long been a point of interest (Kuriwaki, 2019*a*; Campbell and Miller, 1957). Variation in split-ticket rates here could clarify the extent of nationalization of elections. It is therefore worth checking if survey error varies by the composition of an office dyad—whether it involves two national races, one national race and one state level race, or two state races. Figure 3 demonstrates the distributions for each of these three categories.

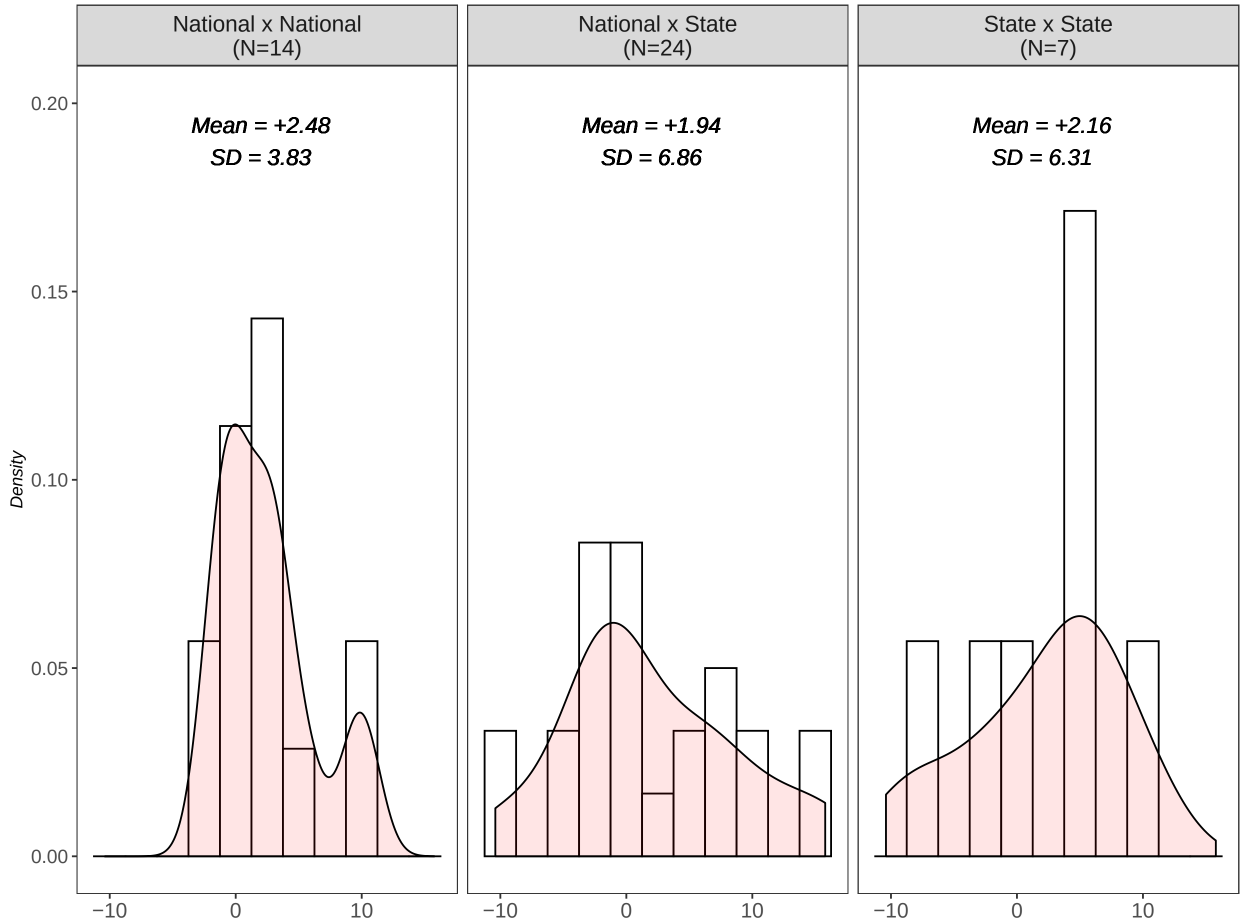

Figure 3: Distribution of Net Error (Survey - CVR) by Office Level Combination

On average, all office level combinations overestimate the true split-ticket by roughly similar extents—National × National by 2.48 points, National × State by 1.94 points, and State × State by 2.16 points. The only notable difference occurs for how tightly error is distributed: National × National has a smaller standard deviation (3.83) than the other two office combinations. Office dyads that involve two national-level races are thus more consistently accurate, avoiding the extreme cases of over- and under-estimation that plague dyads that include at least one state office.

5.3 Simulation Approach

As noted before, some of the survey rates included in error calculations come with relatively large amounts of uncertainty. To better account for uncertainty in these estimates, we apply robustness checks on our initial analysis to make sure that no outlier estimates are driving our results. We generated simulated data from the underlying distribution using the mean and standard deviation of each split-ticket voting estimate. This gives less weight to the noisier survey estimates (those with higher SDs) and more weight to the more precise estimates. The simulated data allows us to plot a better qualitative sense of the distribution of split-ticket voting in the data. For mean inferences, the results will be the same, but we get much more precise estimates of the quantities of interest this way. Our estimates should also produce more accurate estimates of the standard deviation, while in other approaches, singular high-influence points could bias the results. These simulated distributions are shown in the Appendix, in Figures A.3 (net error), A.4 (absolute error), and A.5 (net error by office level combination). Comparing results from Figure 2 (net error, unadjusted) and Figure A.3 (net error, simulated) yield similar takeaways, but the latter conveys a better sense of the underlying distribution of results.

5.4 Explaining Survey Error

Finally, to shed light on what drives survey error, we test for associations between survey estimate characteristics and error in our pool of 45 observations using OLS regression. We set absolute error rather than net error as the outcome for the sake of easier interpretation, and regress it on several variables: office level combination (National $\times$ State and State $\times$ State, with National $\times$ National as the reference group), state, year, a dummy for the survey taking place after the election (which is equivalent to a dummy for CCES surveys), the absolute distance in days between the mean field date in a survey and the relevant election date, and the sample size for the survey from which an estimate is drawn (the weighted total, which we then take the log of because of its right-skewness).

Table 2 shows the results of this regression. The underpowered analysis makes it hard to draw any definitive conclusions, but there are interesting patterns worth noting. Split-ticket rates that involve two national level races (the reference group) have slightly smaller error than the other combination types, but these differences are not statistically significant. The significant negative coefficient on the post (CCES) dummy indicates that CCES/post-election surveys has smaller error. It is unclear if this has more to do CCES or with voters more accurately expressing their choices across different offices after they actually make those decisions. Integrating pre-election CCES data will make for an informative next step, though the pre-election wave asks fewer office votes than the post-election wave does (which originally motivated our use of post-election data in this paper).

Distance from an election correlates positively with absolute survey error, as might be expected and in line with past literature (Belli et al., 1999). Every additional day further away from when the election takes place increases the absolute error by 0.05 points ($p = 0.02$). The estimate is small but may accumulate in cases when surveys are fielded months in advance of an election. Post-election surveys in our data (the CCES) collect responses as late as about one month after an election; days between survey date and election date will not be large enough to contribute sizable error.

Lastly, sample size—represented by the logged sum of the weights from which a split-ticket estimate is drawn—has a negative relationship with absolute error. The coefficient on the log-transformed predictor indicates that for every 10% increase in sample size that a split-ticket estimate is based on, absolute error decreases by 0.25 points. As the sample grows, error decreases. This is intuitive and expected, but is worth pointing out given previous results, such as from Figures 1 and A.1. As noted earlier, many of the largest errors came from polls that relied on small sample sizes. These estimates should be given less weight when evaluating results, and this specific regression estimate confirms that noisy, small sample estimates are to some degree deflating the accuracy of survey split-ticket rates.

Table 2: Regression of Absolute Error on Survey Characteristics

	Absolute Error
National x State	1.718
	(1.695)
State x State	2.213
	(2.723)
South Carolina	0.139
	(1.636)
2012	0.888
	(2.645)
2014	0.861
	(2.126)
2016	3.310
	(2.377)
2018	1.443
	(1.775)
Post-election (CCES)	-7.943^{***}
	(2.062)
Days until/from election	0.050^{**}
	(0.021)
log(Sample size)	-2.618^{***}
	(0.952)
Constant	23.558^{***}
	(7.992)
N	45
R-squared	0.416

$^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

6 Conclusion

A popular topic both in and outside of political science, split-ticket voting has long been studied with survey data, but with little to no knowledge about the accuracy of this measurement approach. We fill this gap in understanding by taking advantage of cast vote records—anonymized individual level ballots—to validate survey split-ticket estimates in certain state-years, and more specifically to address potential for survey underestimation of the truth. Evaluating split-ticket rates from various public and private surveys, we find that, in general, surveys accurately capture the true rates of split-ticket voting. Fine-grained survey measurement of vote choice such as for two offices in combination is reliable and does not result in greatly compounded error.

To summarize, average error is far from substantial (+2.14 net error). Most of the estimates run in the opposite direction of our expectations, over- rather than under-estimating. Split-ticket comparisons that involve two national races—which are more commonly used in past work—better avoid extreme cases of error. Variability in accuracy exists, but as earlier discussion of extreme estimates and the modeling suggests, this might have more to do with (unexpected) small sample size problems than the quality of surveys themselves. The more reliable estimates from our data suggest surveys do well in capturing true split-ticket rates.

We conclude by discussing next steps and areas of improvement for our efforts to validate survey split-ticket measurement. First, while disregarding abstentions from CVRs and nonresponse/uncertainty from surveys was justifiable, integrating these other voting options in some way will be informative and perhaps better reflect the data conditions that researchers face. Second, a concern we do not address here is distinguishing between overall and split-ticket error rates—does the split-ticket error we find between offices A and B arise only because overall vote distributions for offices A and B by themselves were off to begin with? Finding a way to separate out error with marginal rates from error with combination rates will be a key next step. Lastly, obtaining surveys with comprehensive offices asked about (beyond just the CCES) will open new avenues, allowing for a more complete office combination approach (instead of only office dyads) and for straight/mixed ballot rates.

References

Abramowitz, Alan I. and Steven Webster. 2016. "The Rise of Negative Partisanship and the Nationalization of U.S. Elections in the 21st Century." *Electoral Studies* 41:12–22.

Alvarez, R. Michael and Matthew M. Schousen. 1993. "Policy Moderation or Conflicting Expectations?: Testing the Intentional Models of Split-Ticket Voting." *American Politics Quarterly* 21(4):410–438.

Ansolabehere, Stephen and Eitan Hersh. 2012. "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate." *Political Analysis* 20(4):437–459.

Beck, Paul Allen, Lawrence Baum, Aage R. Clausen and Charles E. Smith. 1992. "Partisan Ambivalence, Split-Ticket Voting, and Divided Government." *American Political Science Review* 86(4):916–928.

Belli, Robert F., Michael W. Traugott, Margaret Young and Katherine A. McGonagle. 1999. "Reducing Vote Overreporting in Surveys: Social Desirability, Memory Failure, and Source Monitoring." *Public Opinion Quarterly* 63(1):90–108.

Born, Richard. 1994. "Split-Ticket Voters, Divided Government, and Fiorina's Policy-Balancing Model." *Legislative Studies Quarterly* 19(1):95–115.

Box-Steffensmeier, Janet M., Gary C. Jacobson and J. Tobin Grant. 2000. "Question Wording and the House Vote Choice: Some Experimental Evidence." *Public Opinion Quarterly* 64(3):257–270.

Brownstein, Ronald. 2016. "Voters Don't Split Tickets Anymore.".
URL: *https://www.theatlantic.com/politics/archive/2016/09/senate-trump/498287/*

Burden, Barry C. and David C. Kimball. 2002. *Why Americans split their tickets: Campaigns, competition, and divided government.* Ann Arbor: University of Michigan Press.

Campbell, Angus and Warren E. Miller. 1957. "The Motivational Basis of Straight and Split Ticket Voting." *American Political Science Review* 51(2):293–312.

Davis, Nicholas T. 2015. "The Role of Indifference in Split-Ticket Voting." *Political Behavior* 37(1):67–86.

Davis, Nicholas T. and Lilliana Mason. 2016. "Sorting and the Split-Ticket: Evidence from Presidential and Subpresidential Elections." *Political Behavior* 38(2):337–354.

Garand, James C. and Marci Glascock Lichtl. 2000. "Explaining Divided Government in the United States: Testing an Intentional Model of Split-Ticket Voting." *British Journal of Political Science* 30(1):173–191.

Herron, Michael C. and Jeffrey B. Lewis. 2007. "Did Ralph Nader Spoil a Gore presidency? A Ballot-level Study of Green and Reform Party Voters in the 2000 Presidential Election." *Quarterly Journal of Political Science* 2(3):205–226.

Jacobson, Gary C. 2013. "Partisan Polarization in American Politics: A Background Paper." *Presidential Studies Quarterly* 43(4):688–708.

Kennedy, Courtney, Andrew Mercer, Scott Keeter, Nick Hatley, Kyley McGeeney and Alejandra Gimenez. 2016. Evaluating Online Nonprobability Surveys. Technical report Pew Research Center.

Kilgore, Ed. 2018. "2018 Midterms Offered More Proof That Split-Ticket Voting Is a Thing of the Past.".
URL: *http://nymag.com/intelligencer/2018/11/2018-midterms-split-ticket-voting.html*

Kuriwaki, Shiro. 2019*a*. Party Loyalty on the Long Ballot: Is Ticket Splitting More Prevalent in State and Local Elections? Technical report Working Paper.

Kuriwaki, Shiro. 2019*b*. The Administration of Cast Vote Records in U.S. States. Technical report Working Paper.

Lacy, Dean, Emerson M. S. Niou, Philip Paolino and Robert A. Rein. 2019. "Measuring Preferences for Divided Government: Some Americans Want Divided Government and Vote to Create It." *Political Behavior* 41(1):79–103.

Lewis, Jeffrey B. 2001. "Estimating Voter Preference Distributions from Individual-Level Voting Data." *Political Analysis* 9(3):275–297.

McAllister, Ian and Robert Darcy. 1992. "Sources of Split-ticket Voting in the 1988 American Elections." *Political Studies* 40(4):695–712.

Meyer, Bruce D. and Nikolas Mittag. 2019. "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety Net." *American Economic Journal: Applied Economics* 11(2):176–204.

Moskowitz, Daniel J. 2018. Local News, Information, and the Nationalization of U.S. Elections. Technical report Working Paper.

Mulligan, Kenneth. 2011. "Partisan Ambivalence, Split-Ticket Voting, and Divided Government." *Political Psychology* 32(3):505–530.

Phillips, Amber. 2016. "Is split-ticket voting officially dead?".
URL: *https://www.washingtonpost.com/news/the-fix/wp/2016/11/17/is-split-ticket-voting-officially-dead/?utm*$_t$*erm* $= .43073df3cef8$

Rivers, Douglas and Benjamin Lauderdale. 2016. "Beware the phantom swings: why dramatic bounces in the polls aren't always what they seem.".
URL: *https://today.yougov.com/topics/politics/articles-reports/2016/11/01/beware-phantom-swings-why-dramatic-swings-in-the-p*

Rogers, Todd and Masahiko Aida. 2014. "Vote Self-Prediction Hardly Predicts Who Will Vote, and Is (Misleadingly) Unbiased." *American Politics Research* 42(3):503–528.

Sigelman, Lee, Paul J. Wahlbeck and Emmett H. Buell Jr. 1997. "Vote Choice and the Preference for Divided Government: Lessons of 1992." *American Journal of Political Science* 41(3):879–894.

Skelley, Geoffrey. 2018. "Split-Ticket Voting Hit A New Low In 2018 Senate And Governor Races.".

URL: *https://fivethirtyeight.com/features/split-ticket-voting-hit-a-new-low-in-2018-senate-and-governor-races/*

Wright, Gerald C. 1990. "Misreports of Vote Choice in the 1988 NES Senate Election Study." *Legislative Studies Quarterly* 15(4):543–563.

Wright, Gerald C. 1993. "Errors in Measuring Vote Choice in the National Election Studies, 1952-88." *American Journal of Political Science* 37(1):291–316.

Appendix

A Additional Figures

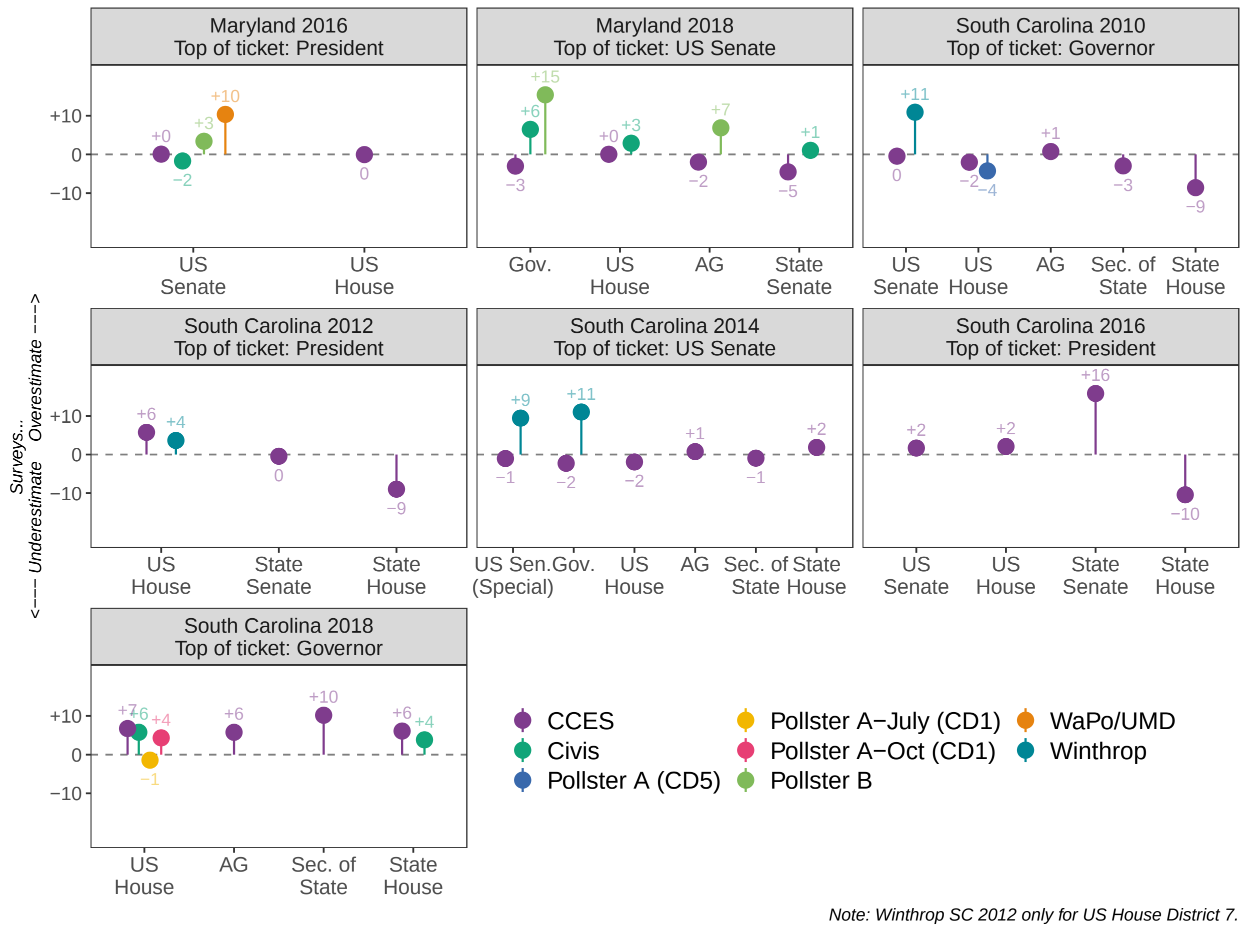

Figure A.1: Net Error (Survey - CVR) by Race

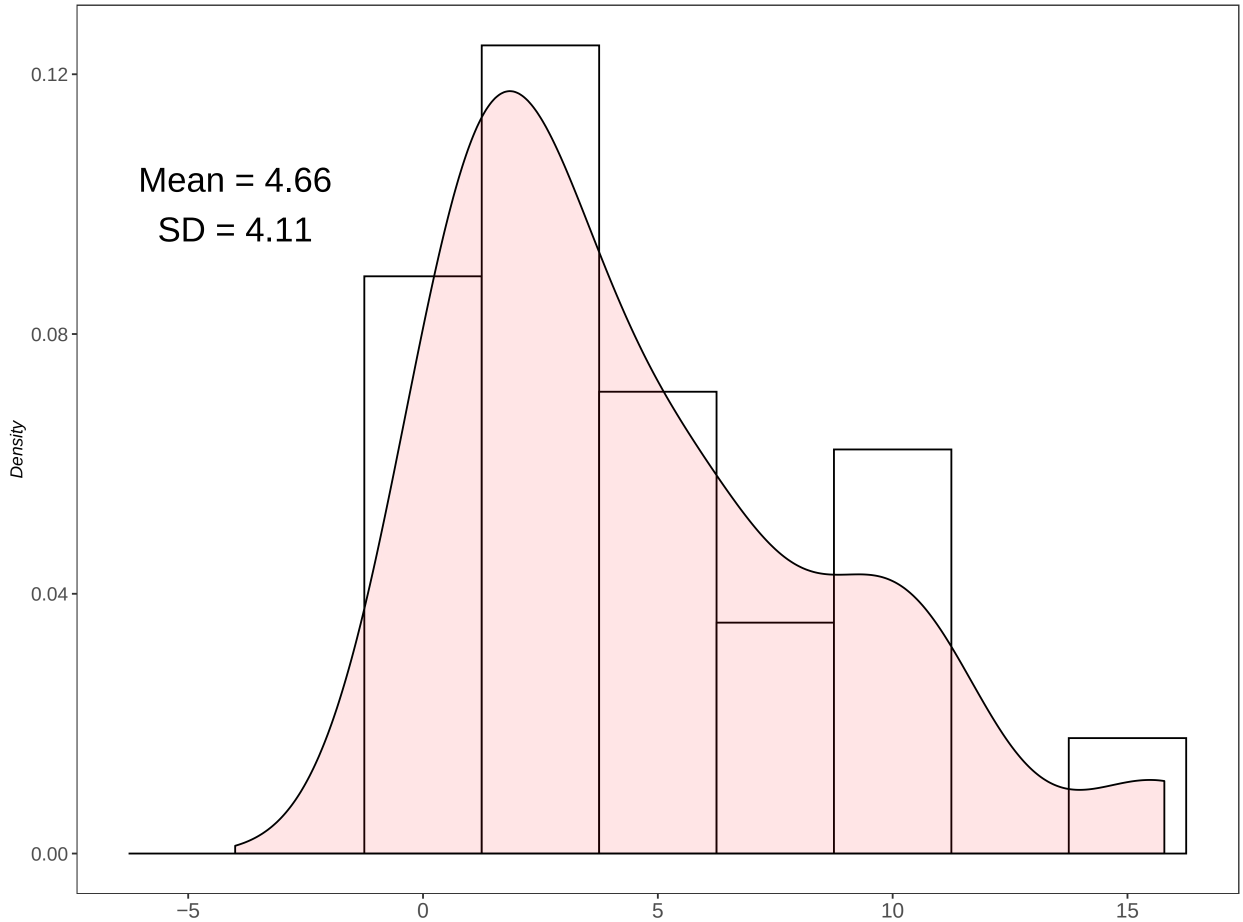

Figure A.2: Distribution of Absolute Error (Survey - CVR)

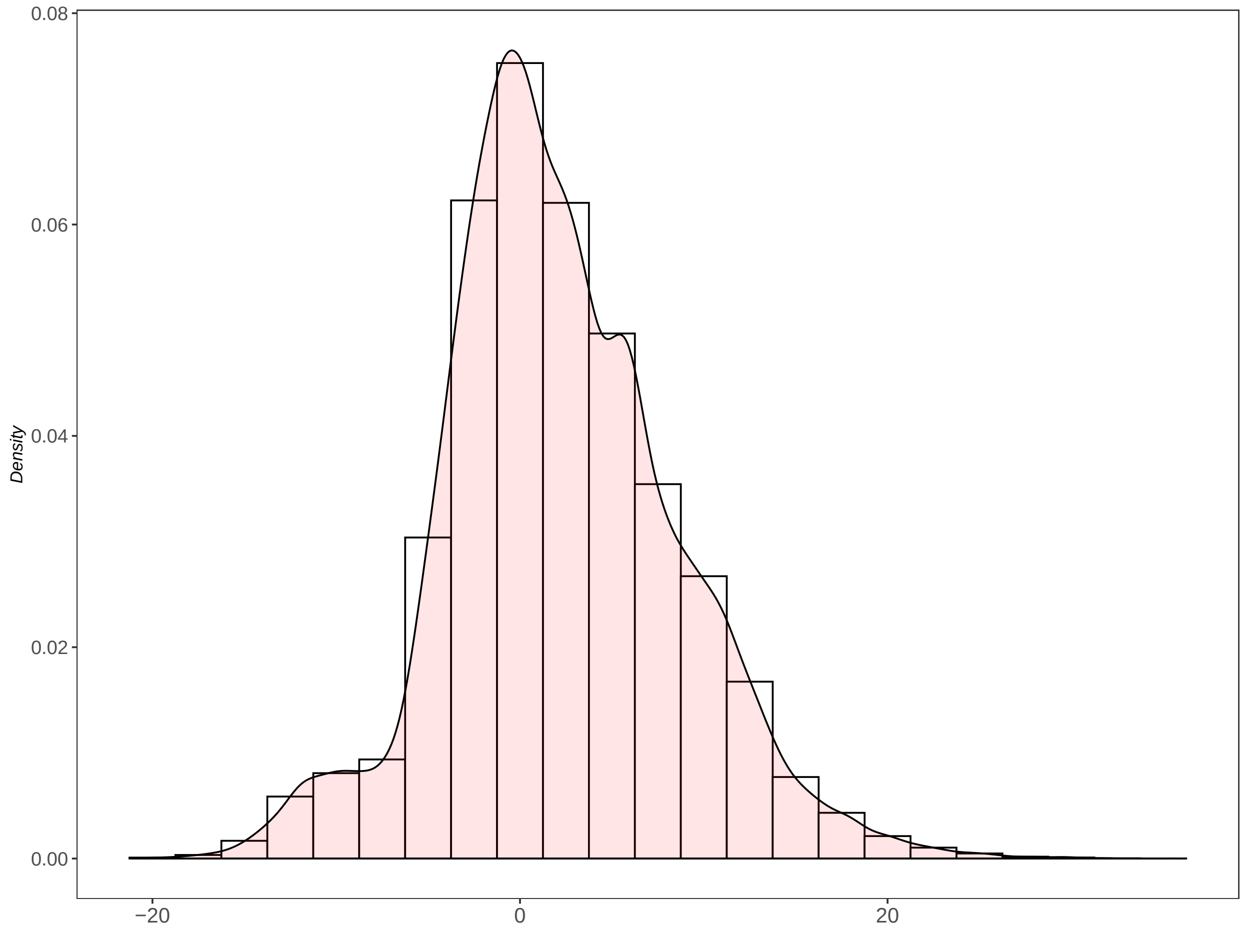

Figure A.3: Simulated Distribution of Net Error (Survey - CVR)

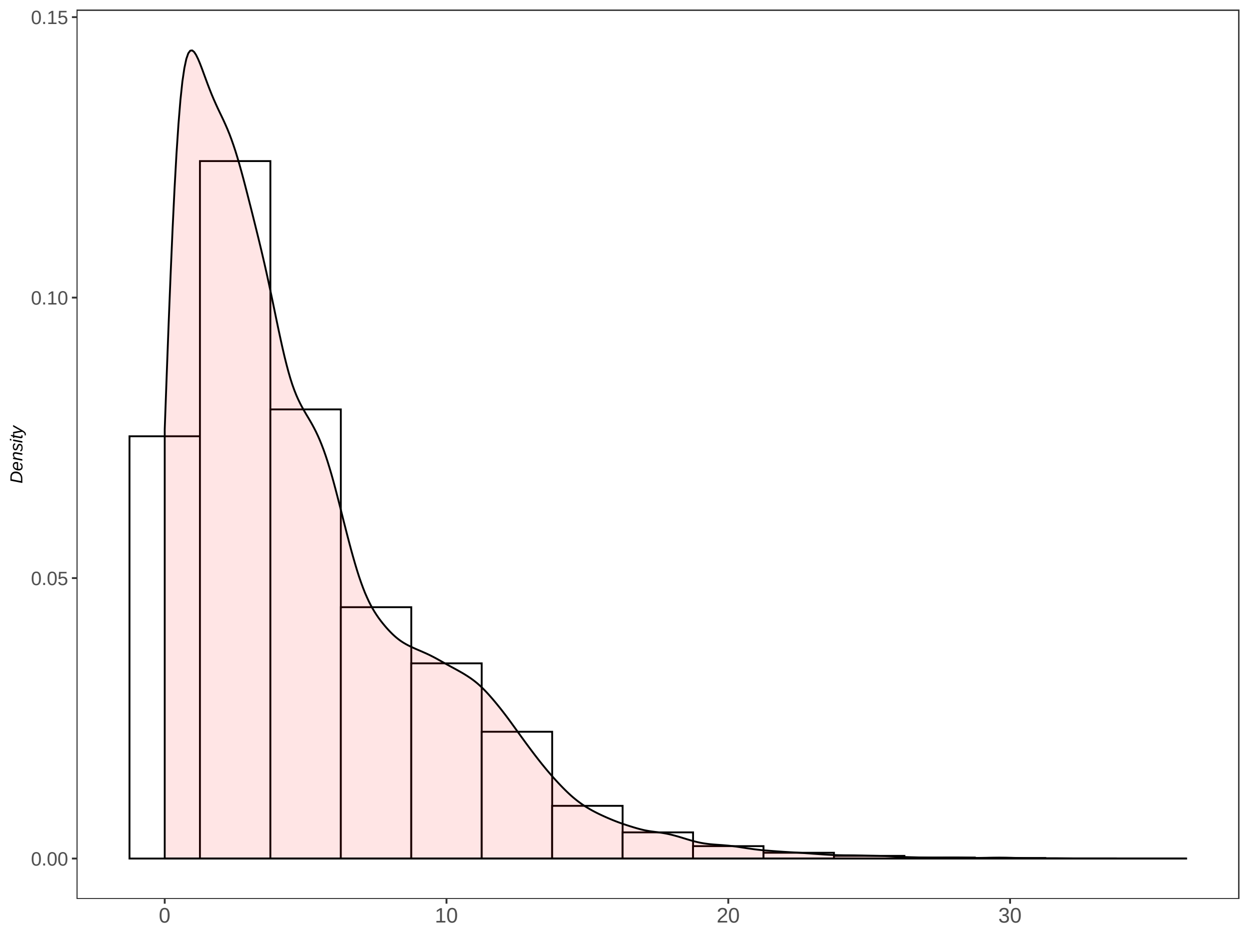

Figure A.4: Simulated Distribution of Absolute Error ($| \, Survey - CVR \, |$)

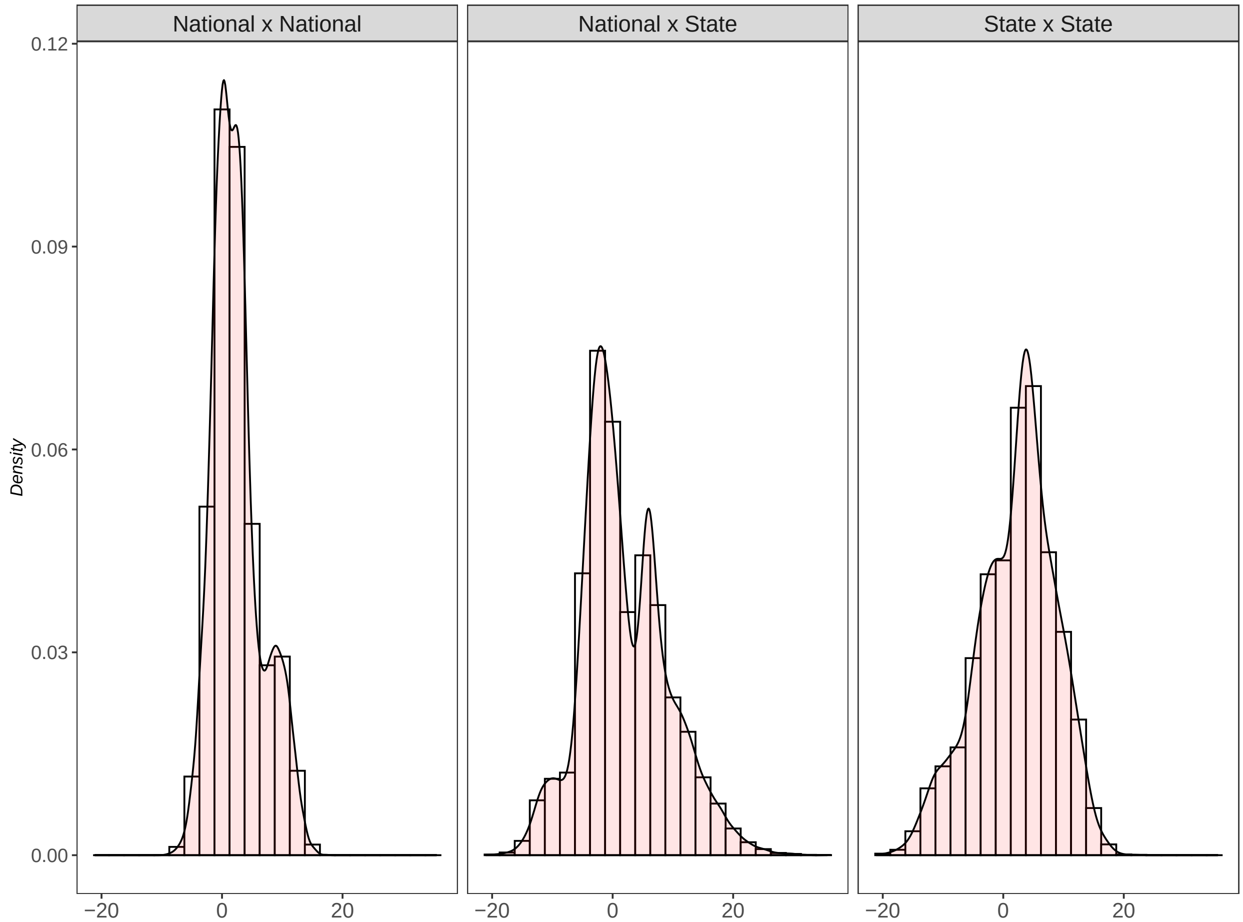

Figure A.5: Simulated Distribution of Net Error (Survey - CVR) by Office Level Combination