# Voter Bias in the Associated Press College Football Poll: Reconducting a 2009 study with new data in a \$1 Billion-dollar industry that has seen significant changes in the past decade

Brent Hensley

University of California, Berkeley

Econ H195B

Spring 2020

#### Abstract

This paper investigates bias in Associated Press College Football poll to determine if voters give preferential to treatment to teams in the voter's state or conferences represented in the voter's state. Using censored Tobit modelling, evidence for bias exists in these areas. Compared to a 2007 study using the same model, the bias found at present is lower.

#### Keywords

Football, Voting, Associated Press College Football Poll, Conference, Pollster, censored Tobit

#### Acknowledgement

I thank Professor Handel for the support and feedback as my advisor, as well as Professor Pouzo for his contributions. I also wish to acknowledge the contribution from Collegepolltracker.com for their help in this research.

### **1** Introduction

A paper published in 2009 titled "Voter Bias in the Associated Press College Football Poll" by B. Jay Coleman, Andres Gallo, Paul M. Mason, and Jeffrey W. Steagall "investigate[d] multiple biases in the individual weekly ballots submitted by the 65 voters in the Associated Press college football poll in 2007." In their research they looked at 9 weeks' worth of poll data, and concluded that they found evidence of bias "toward teams (a) from the voter's state, (b) in conferences represented in the voter's state, (c) in selected Bowl Championship Series conferences, and (d) that played in televised games, particularly on relatively prominent networks." This research paper reconducts their analysis with a current and expanded data set, with the aim of looking for bias in areas (a), (b), and (c) as they have described.

#### **1.1 Historical Context**

Since 2009, the structure of the college football industry has undergone some pretty serious changes. A decade ago, the national champion was determined through a format called the Bowl Championship Series (BCS), which used a computer (that took AP poll rankings as one of the inputs) to determine the 2 best teams and have them face each other at the end of the season in a championship game. Since then, a "Playoff Committee" has been formed, which released rankings weekly for the second half of each football season, and at the end selects the four "best" teams to play off in a 4-team playoff bracket system. While it is not required, the de facto state of the College Football Playoff (CFP) is that each of the four teams must come from one of the "Power 5" conferences, which are considered to be the most elite conferences, as well as equal to each other. These conferences are the Pac-12 (PAC), Atlantic Coast (ACC), Big 12 (B12), Big Ten (BIG), and Southeastern (SEC), and are each comprised of individual member universities from around the country. However, in 2009, there were 6 conferences that were considered to be the

most elite (the BCS conferences): the current Power 5 plus what was then the Big East Conference. In this last 10 years the Big East has dissolved, leading to conference realignment. Six teams from the Big East joined conferences including BIG, B12, and ACC. Other re-alignment occurred around the nation as well, changing the input of "teams in conferences represented in the voter's state" which is one of the variables and biases examined here. A full list of teams who underwent realignment can be found in Table B. Additionally some teams such as Utah (PAC) and Texas Christian (B12) were not in a BCS conference at the time of the previous study but have since joined to one. As of 2019, there are a total of 65 Power 5 teams located in 35 states.

In Figure 1, states are colored according to the conference's teams located within them are associated with. Eight states are home to multiple conferences, and have been designated with multiple colors accordingly, such as Pennsylvania which has 2 Power 5 schools: Penn State which competes in the Big Ten, as well as Pittsburgh which competes in the ACC. Due to varying size of conferences as well as greater density of universities in the eastern United States, the conferences are not evenly distributed both in that they have overlapping reaches as well as are can be found in differing numbers of states. This is numerically outlined in Table 1.

Table 2 shows the distribution of teams and pollsters from a state and conference level perspective. It is worth noting that the number of pollsters per area represented by a conference's footprint is not directly proportionate to the number of schools in the conference. For instance, 24 pollsters live in states home to an SEC school, or 1.7 pollsters per school, whereas 8 pollsters live in the six states home to one of the twelve schools in the PAC.

#### 1.2 Hypothesis

In regard to the biases I examined, my preliminary expected findings were:

Bias towards teams in the same state and represented in the voter's state will remain the same. I expect this for the same reason they drew their initial conclusion, which is that since a voter can only watch so many games in a day, they will still prefer to watch games of teams of which they themselves are a fan, so most likely those close to where they live.

#### **1.3 Motivation**

Given the college football industry is valued at over \$1 Billion dollars annually, there is real value in understanding the structure of power that lies within it. Since the AP Poll Rankings have a broad impact on which teams are perceived as good, or worth watching for the general public, the AP voters can influence which teams end up on prominent TV networks, which bring in large revenues for the schools. Additionally, the perceptions of schools tie directly to the worth of the school's brand and recruiting abilities, meaning AP rankings have an influence beyond just TV ratings. Understanding if there is bias in the poll, and its extent, is valuable to know as it is a rare scenario that 65 journalists get to fill out a ballot weekly and hold influence over an industry of this size. While research has been done on this topic before, the landscape has changed so drastically in recent years that it is worthwhile to reconduct this experiment with current data to see if changes such as the ones enumerated above are lessening proven bias. As a result, this research used the same modelling the original researchers did, censored Tobit, to try and recreate results as directly comparable as possible to the original paper.



### Figure 1: Map of the United States by Power 5 Conference Footprint

Table 1: Dispersion of Schools & Conferences

Conference	# of Teams	# of States
ACC	15	10
BIG	14	11
B12	10	5
PAC	12	6
SEC	14	11

Conf./State	# of Teams	# of Pollsters	State	# of Teams	# of Pollsters
ACC	15	19	MI	2	2
BIG	14	17	MN	1	1
<b>B12</b>	10	8	МО	1	1
PAC	12	8	MS	2	1
SEC	14	24	NC	4	3
AL	2	3	NE	1	1
AR	1	1	NJ	1	1
AZ	2	1	NY	1	1
CA	4	3	OH	1	4
СО	1	1	OK	2	1
FL	3	5	OR	2	1
GA	2	2	PA	2	2
IA	2	1	SC	2	2
IL	2	1	TN	2	2
IN	3	2	ТХ	5	4
KS	2	1	UT	1	1
KY	2	1	VA	2	1
LA	1	2	WA	2	1
MA	1	2	WI	1	1
MD	1	1	WV	1	1

 Table 2: Table Showing Geographical Location of 2018 Pollsters

Note: 6 Pollsters live in states without a Power 5 team

### **2** Literature Review

Since this research was directly inspired by an existing piece of literature, this literature review begins by detailing it and discussing how its shortcomings motivate this research paper, and how it relates to the research question. Additionally, below is a contrast other existing literature and how it applies to this research and anticipated findings.

1. Coleman, B. Jay, Andres Gallo, Paul M. Mason, and Jeffrey W. Steagall. "Voter bias in the associated press college football poll." *Journal of Sports Economics* 11, no. 4 (2010): 397-417.

This paper, originally published a decade ago, is the foundation for this research. I am asking many of the same questions they asked, with the primary focus of my paper being to reconduct their study with current data to examine differences in biases as a result of a changed landscape in college football in the 10 years since their paper was originally published. In their research they looked at 9-weeks' worth of poll data, and concluded that they found evidence of bias "toward teams (a) from the voter's state, (b) in conferences represented in the voter's state, (c) in selected Bowl Championship Series conferences, and (d) that played in televised games, particularly on relatively prominent networks." With the substantial changes since their publication (the introduction of a College Football Playoff and the introduction of conference-based networks that allow global streaming, as detailed in the introduction), the glaring hole in their analysis for current relevance is data from the current system. This has led to the development of this research question.

2. Stone, Daniel F., and Basit Zafar. "Do we follow others when we should outside the lab?

Evidence from the AP top 25." Journal of Risk and Uncertainty 49, no. 1 (2014): 73-102. This paper examines the individual ballots of the voters and sees how an individual is influenced over time by the aggregate results from the other experts on a weekly basis. It relates to this research question in that the data they have studied is the same (individual ballots and the rankings of the teams on them), but it differs critically in that it is not focused on bias as it related between the voters and the teams, instead looking to see if individuals are biased by the beliefs of other voters. In this paper they examine the influence of the other voters, who are presumed to be experts, on an individual's choices, whereas at present voters can also be influenced by the CFP rankings, which is compiled by a separate panel of few select experts. There is no mention of the CFP rankings in their paper, as it was published before the current system was implemented. There is a significant and defined gap in existing literature on the influence of the CFP rankings on the AP poll.

 Andrews, Rodney J., Trevon D. Logan, and Michael J. Sinkey. "Identifying confirmatory bias in the field: Evidence from a poll of experts." Journal of Sports Economics 19, no. 1 (2018): 50-81.

The bias examined in this paper is confirmation bias on the part of the AP voters, who they have found reinforce their beliefs about a team when that team performs better than expected, or "covers the spread" in gambling terms. Like my own question, this research tracks the voters on a weekly basis and how they rank the teams, but it does not consider personal bias for/against certain teams, nor does it account for CFP rankings or major TV network appearances.

 Campbell, Noel D., Tammy M. Rogers, and R. Zachary Finney. "Evidence of television exposure effects in AP top 25 college football rankings." Journal of Sports Economics 8, no. 4 (2007): 425-434.

This paper speaks to the same motivation as for my own paper's research, by analyzing the effects of playing on national TV on a team's ranking in the AP Poll. Their finding is in line with what the 2007 research has found, which is that teams that play on national TV networks do receive an additional boost in the polls, and the boost is especially large when teams that rarely play in

televised games get to and win, when compared to a team with a comparable record that routinely gets to play on national TV due to a large following. It can be extrapolated from this paper that bias in the AP poll has tangible, monetary effects on the schools which are a part of the highest level of college football.

### **3 Data & Variables**

#### 3.1 Data

The data used records a "ranking" value for every team from a Power 5 conference (65 teams total), for all 63 pollsters, for all 16 weeks of the poll in the 2016, 2017, and 2018 seasons. This leaves me with a data sets consisting of 65,520 data points for each year. This data was obtained from Collegepolltracker.com, which uses a web scraper to gather and format the data from the full ballots of the voter's published weekly by the Associated Press. This data is organized in panel form, with Pollster, Week, and Team as the independent variables, Ranking as the dependent variable, seven binary variables, and a control variable.

#### **3.2 Variables**

The "Ranking" variable takes a value from 0-25, corresponding with where an individual pollster placed a single team on their ballots. Each pollster can select from all teams to complete a "Top 25" ranking, leaving all other teams unranked. Each team is then assigned points based each voter's ballot, so that the team which is ranked first (best) is assigned 25 points for that voter, the second team 24 points... twenty-fifth 1 point, and all subsequent unranked teams receiving 0 points. The total number of points received from all voters is then aggregated, and the teams are ordered based on number of points, with the first-place team in points receiving the number one ranking, and so forth.

The seven dummy variables allow the regression to look for bias as shown in the coefficients. Table 3 provides the summary statistics for these variables. The first binary variable, "InState", takes value 1 if a team is located in the same state as the pollster and a value 0 otherwise, and the variable "InConf" takes value 1 if a team plays in a conference which also has a team in the same state as the pollster. For example, a pollster who lives in California would have "InState"

= 1 and "InConf" = 1 for the teams which are in California (Cal, Stanford, USC, UCLA), have "InState" = 0 and "InConf" = 1 for all teams which are outside California but are members of the Pac-12 (PAC) conference, and have both binary variables equal to 0 for all other teams. Finally, the control variable is the ranking (in points) of the aggregate of the poll. For example, in Week 0 of the 2018 poll, Wisconsin ranked 4 after the votes were aggregated. As shown in Figure 2 however, individual pollsters had ranked Wisconsin as high as 1 and as low as 13. By using 4 as the control variable in this case, we can test if those who ranked that team either higher or lower did so consistently with their geography. The final five variables are all conference specific, with the variable having a value of 1 if the team is in that conference and 0 otherwise.

#### **3.3** Testing

To test for bias. the same model that the original researchers used, censored Tobit, is used here. The data is censored due to the fact that every unranked team is given an ordinal value of 0, regardless of the pollster's latent ratings of those teams. They found that the independent variables converged in a Tobit model, and thus decided to use it for their data. Due to the fact that in a Tobit model the beta coefficient is not simply the effect of x on y as it is in a linear regression mode, the use of a Tobit model becomes necessitated so that the coefficients derived are able to be effectively compared to the original researchers'. A comprehensive table showing the full findings of Coleman, Gallo, Mason, and Steagall (2009) can be found in Table C in the appendix.

This regression takes the form of the following equation:

 $Y_{i,j} = \alpha_i + \beta_{InState} X_{i,InState} + \beta_{InConf} X_{i,InConf} + \beta_{ACC} X_{i,ACC} + \beta_{BIG} X_{i,BIG} + \beta_{B12} X_{i,B12} + \beta_{PAC} X_{i,PAC} + \beta_{SEC} X_{i,SEC} + u_i ,$ 

Where  $Y_{i,j}$  is the expected ranking of team i from individual j,  $\alpha_i$  is the estimated ranking of team i if the team is not in the same state/conference as pollster, and  $X_i$  takes value 1 if the team is in the same state/conference as the pollster, with the Betas serving as coefficients of X.

	Mean	Median	Minimum	Maximum
InState	0.03039	0	0	1
InConf	0.23659	0	0	1
ACC	0.23077	0	0	1
BIG	0.21538	0	0	1
B12	0.15385	0	0	1
PAC	0.18462	0	0	1
SEC	0.21538	0	0	1
Ranking	4.59167	0	0	25
Control	4.60897	0	0	25

**Table 3: Summary Statistics** 

POLLSTER	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AARON MCMANN	Ø	-	G	Ð	<b>i</b> ®t	w	Ø	۲	U	樹	1	R	8	M	ş	٧T	0	- Aufe	TCU	-	₩	LSU	¥	-	10g
ADAM JUDE The Seattle Times	-	Ø	G	w	Ø	<b>P</b>	Ð	0	<b>B</b>	U	\$	R	M	1	TCU	nite in	₩	۱	동	5	Upp	0	Y	77	*
ANDY GREDER St. Paul Pioneer Press	Ø	G	-	Ø	<b>P</b>	w	Ð	U	0	ß	-	5	M	퉛	N	TCU	0	₩	¥	Upp	LSU	٧7	nişin	-	œ
BEN JONES Tuscaloosa News	Ø		<b>1</b> 01	Ø	G	w	Ð	0	U	M	R	₽	8	1	¥	TCU	동	77	-	ninin	0	Upp	₩	南	-
BILL LANDIS Cleveland Plain Dealer	Ø	-	191	G	w	0	Ð	Ø	ß	U	R	-	M	\$		돈	0	TCU	0	77	5	10-	-	LSU	Y
BLAKE TOPPMEYER Knoxville News Sentinel	-	Ø	Ø	101	G	Ð	Ð	w	6	0	U	<b>U</b> 25	8	X	TCU	0		₩	M	٧7	5	Дм	동		A
BOB ASMUSSEN Champaign News-Gazette	Ø	Ø	G	*	r@t	6	Ð	U	6	8	w	8	M	77	R	동	~		Y	-	U.F.	LSU	0	TCU	0
BRANDON MARCELLO	Ø	*	G	w	Ø	U	<b>P</b>	0	ß	Ð	6	8	TCU	M	₩		동	10g	X	6	٧T	-	-	œ	0
BRENT AXE The Post-Standard	-	Ø	w	Ø	G	<b>i</b>	Ð	0	₿	R	6	₩	U	M	TCU	5	0	A	Y		퉛	-	Up	-	Ο
BRETT MCMURPHY Stadium	Ø	-	Ø	w	i 🖗	G	Ð	8	6	6	₿	M	Ł	See.	₩	U		-	0	*	TCU	-	ø	Y	<b>W</b>
BRIAN HOWELL Daily Camera	Ø	-	G	Ø	Ð	w	<b>i B</b> t	U	M	ß	0	-	X	5	Ł	0	5	77	ain a	₩	Ο	¥	TCU	U.S.	*
CHRIS MURRAY Nevada Sports Net	Ø	-	G	w	<b>i</b>	Ø	U	Q	R	0	M	-	8	Ł	Y	ß	5	TCU	and the	₩	0	0	***	77	U.S.
CHRIS SOLARI Detroit Free Press	-	Ø	G	Ð	U	1	6	₿	ø	۲	w	M	8	0	<b>U</b> .;-	- Andre	R	동	٧7	모	₩	-	Y	tcu	Ο
CHUCK CARLTON The Dallas Morning News	Ø		G	W	ወ	1	Ø	₿	6	~	M	U	6	TCU	5	동	R		<b>U</b> 27	٧7	œ	-	A	0	R
CONOR O'NEILL Winston-Salem Journal	Ø	-	Ø	G	w	₿	<b>P</b>	Ð	U	5	6	~	6	R	-	M	TCU	77	0	U.	0	동	ø		A
DAVE SOUTHORN The Idaho Statesman	-	Ø	G	<b>1</b> 00	w	97	Ð	U	6	Ø	0	M	R	5		TCU	퉛	77	0	-	₩	¥	Up	-	<b>W</b>
DAVID BRIGGS Toledo Blade	Ø	*	101	G	Ø	w	1	U	Ð	M	0	R	₿	8	₩	돈	TCU	77	Up;	-	0		-	노르미	
DON WILLIAMS Lubbock Avalanche-Journal	-	Ø	G	Ø	ወ	1	w	₿	U	1	돈	M	R	₩	5	0	77	ninin	TCU	Q	5	0	ga <b>n</b> a	١	
DYLAN SINN The Journal Gazette	Ø	101		Ø	ወ	U	G	R	w	₿	0	₩	TCU	M	5	동	A	-	Y	노모니	77	Up.		5	0
ERIC HANSEN The South Bend Tribune	-	Ø	G	Ø	w	<b>P</b>	U	Ð	0	R	M	alula.	ş	₿	TCU	0	5	١	¥	1	Upp	0	-		77
FERD LEWIS The Honolulu Star-Advertiser	Ø	-	G	W	<b>P</b>	Ð	Ø	₿	0	U	1	5	A	돈	M	VT	alute.	0	<b>U</b> 27	TCU	₩	-	0	-	¥
GARLAND GILLEN WVUE-TV FOX 8	Ø	G	*	Q	r@t	Ø	0	W	VT	鬯	U	~	-	M		R	돈	TCU		Up;	-	8	LSU	¥	0
GARY HOROWITZ KBZY radio	Ø	*	G	<b>i O</b> t	Ø	Ð	U	₿	0	w	3	-	₩	5	M	R	돈	TCU	Up;	alula.	0	0		-	Y
GRACE RAYNOR The Post and Courier	Ø	-	G	Q	Ø	w	<b>i B</b> t	U	ß	8	0	N	1	돈	TCU	77	M	0	₩	U.		¥	8	LSU	œ.
JASON GALLOWAY Wisconsin State-Journal	Ø	-	G	101	Ð	Ø	w	R	0	U	6	₿	M	8	0		돈	77	-	TCU	-		Up.	-	¥
JERRY DIPAOLA Pittsburgh Tribune-Review	Ø	-	G	Q	101	Ø	w	M	U	6	٢	₿	X	8	TCU	돈		٧7	₩	0	¥	U.	5	Дм	-
JIM ALEXANDER The Press-Enterprise	Ø	-	G	W	Ø	5	U	0	Ð	19	TCU	M	₿	8	R	₩	돈	0	٧7	5		0	Y	LSU	
JOEY KAUFMAN Southern California News	-	Ø	G	W	Ø	₿	-	Ð	U	1	0	TCU	8		~	M	돈	R	0	-	Ο	<b>U</b> 27	٧7	U.	¥
JOHN BEDNAROWSKI Marietta Daily Journal	Ø	191	-	G	Ø	U	ዒ	₿	6	R	W	-	돈	M	5	TCU	0		Up;	₩	5	٧7	-	LSU	
JOHN CLAY Lexington Herald-Leader		Ø	1 <sup>®</sup> t	Ø	w	Ð	G	8	Ø	U		M	$\otimes$	0	R	8	F.	١	٧7	Up;	퇀	-		C	¥
JON WILNER San Jose Mercury News	Ø	-	Ø	U	G	-	w	191	0	6	Ð	<b>U</b> 27	TCU	₿	-	M	-	8	٧7		X	돈	Y	A	0
JONNY MILLER WBZ NewsRadio 1030	Ø		Ø	G	101	w	Ð	R	₿	۲	U	M	8	6	nguga	٧7	¥	노모	0	TCU	퉛	-	Up.	œ	Дм
KEITH SARGEANT NJ Advanced Media	Ø		G	1	<b>67</b>	ዒ	w	0	U	R	₿	1	5	퇀	M	TCU	٧7	₩	노모	<b>U</b> 37	¥	0	aine a	5	-
KELLIS ROBINETT Wichita Eagle	Ø	-	G	Q	w	U	<b>i O</b> t	Ø	0	₩	R	₿	1	8	M	돈	77	TCU	and the	¥	0	Up	5	Ο	<b>New</b>
KIRK BOHLS Austin American-Statesman	-	Ø	G	<b>N</b>	1	Ø	w	U	ዒ	6	5	₿	5	<b>U</b> 27	~	77	R		e	ş	gilles	0	TCU	M	ø
LAUREN BROWNLOW WRAL	Ø	*	G	Ø	W	₿	191	0	X	Ð	U	6	TCU	8	~	M		LSU	Up.	Ч	B	-	돈	-	*
MARC WEISZER Athens Banner-Herald	Ø	*	G	Ø	W	Q	191	₿	0	₩	U	6		X	M	ş	tcu	0	٧7		Up.	-	œ	0	노모
MATT BAKER Tampa Bay Times	-	Ø	Ø	G	194	W	0	₿	Ð	U	M	×			₩	8	tcu	0	¥	-	ş	œ	0		10-
MATT BROWN Sports on Earth	Ø	*	W	Ø	G	U	0	TCU	Q	i 🚯	M	₿		and a	6	١	R	5	0	₩	Upp	돈	Q	A	œ.
MATT MCCOY 610 WTVN-AM	Ø	*	Ø	W	G	U	Q	-	101	R	₿	M	0	ş	8	0	₩	TCU	٧7	IJ	•	¥	à	5	

Figure 2: Visualization of Ballots, Week 0 Poll 2018 (Source: Collegepolltracker.com)

### **4 Summary of Results**

The two primary coefficients examined are those of InState and InConf, which have values that varied in each of the three years of the analysis. Only in 2018 were the instate and inconf variables statistically significant, taking values of 0.2721 and 0.3057 respectively. This indicates that for this season, the pollsters give preference in the rankings to teams which play close to them. In 2016 and 2017 however, the values were non-trivial, but not statistically significant on a 95% confidence interval. Due to this, the rest of the findings are based only on data from the 2018 season.

#### 4.1 2018 Results

The 2018 values can be compared to the values found by Coleman, Gallo, Mason, and Steagall in their paper, to examine the strength of the bias in 2007 compared to 2018. They found the coefficients for the state bias to be 0.645, and conference bias to be 0.357 in 2007. These coefficients show sizable bias towards teams that are from the same state as the pollster, and as well as a bias towards teams which play in conference's represented in the voter's state. As far as comparing the coefficients to those of the original study, using a 95% confidence interval, it can be concluded that the bias shown towards teams located in the same state as the pollster has decreased; this factor is significant at the .001 level. It cannot however be concluded that pollsters show less favoritism now than they did in 2007 to teams that play in conference's represented in the pollster's state at a .05 level.

As for the conference specific variables, which looks at bias towards/against teams in certain conferences relative to all teams, without regard for the pollster's geography, the findings are largely contiguous with the findings found in the 2007 study. These values can be found in Table 4. In this analysis, two of the conferences were shown to experience bias under a 95%

confidence interval. The PAC had a coefficient of 0.3348, showing bias towards teams that play in the Pac-12 Conference. In the 2007 study, it was also found that the PAC received favorable treatment from voters. Additionally, this analysis concludes that the AP voters are biased against the ACC, meaning that teams in the ACC receive statistically lower rankings than comparable counterparts in other conferences. This was also the case in the 2007 study, although the evidence for bias against the conference is larger now than in 2007. None of the rest of the conferences have coefficients that show evidence of bias at the .05 level, although there is evidence for bias towards teams from the SEC at the .1 level. This differs from the original study which found the strongest evidence for bias towards the SEC and B12, with marginal bias against the BIG. The coefficient of the largest magnitude found in this research (ACC: -.4156) is still smaller than the magnitude of bias found for/against four of the six conferences looked at in the previous study. This indicates that across the board the AP poll is less biased for or against specific conferences now than it was in the past. This is likely because of the increased prevalence of cable networks and internet streaming of games. In the past, one had only a handful of games to choose from, and due to regional broadcasting there would be a disproportionate number of games from the conference closest to the pollster. With the current media landscape, one can watch a game regardless of where the teams are from, therefore minimizing the bias on conferences.

#### **4.2 Additional Analysis**

To further explore the state and regional biases, a further series of regressions were conducted, with a slimmed number of teams included in each (Table 5). Using the 2018 data, the same tobit regression was run starting with only teams that had received votes in a given week. This would include the teams for each week that any pollster gave votes to, regardless of whether the team ended up ranked in the aggregate results. This analysis excludes all teams for which not

a single pollster voted for. This method of removing teams then continued by only looking at teams who were ranked in the final top 25 for each week, then top 20, top 15, top 10, top 5, and top 4. The top 4 is a valuable metric to examine as is it the top 4 teams in the final CFP poll that participate in the playoffs for the national championship title. While that is a separate poll, the AP voters use their top 4 to indicate which teams they believe are most deserving of a playoff berth, making the *#*5 ranking relatively invaluable compared to the spots above it. Using these methods, it can be seen that the instate coefficient consistently increases as the number of teams included in the analysis decrease, while the InConf coefficient steadily decreases. These values can be found in Table 5, and are visualized in Figure 3. The value of analysis using a slimmed number of teams is to try and eliminate a skew towards 0 in the coefficients; both instate and InConf can take more representative values when teams which any (or multiple) pollsters gave 0 points to are removed, as many teams are assigned a 0 value and as result a pollster's latent ranking for said team cannot be interpreted.

There are two likely explanations for this. In the 2018 season, only 9 teams were at any point ranked in the top 4; as one moves up the poll towards higher rankings, there is more agreement on which teams belong. While there is often little disagreement on which teams are at the very top, there is often debate about the ordering of said teams given they are all usually undefeated. This often translates to a home-state bias, causing each of the pollsters in a state with an undefeated team to put that team as #1. Additionally, there is less of a regional effect as most of the best teams are closely concentrated. In 2018, 4 of the 9 teams ranked in the top 4 were from the South, with 4 more being in the Great Lakes region. Given the teams are not spread evenly across the country, the regional bias at the pollster level is negligible.

Ranking	Coef.	Std. Error	t	p > abs(t)	[95 % Conf	. Interval]
InState ('16)	.1460503	.1324018	1.10	0.270	1134753	.4055579
InConf ('16)	.0345191	.05406	0.64	0.523	0714387	.1404768
InState ('17)	.005136	.1408732	0.04	0.971	2709755	.2812476
InConf ('17)	.0582472	.0574014	1.01	.310	0542596	.1707539
InState ('18)	.2721288	.0912655	2.98	0.003	.0932484	.4510093
InConf ('18)	.3057007	.0380824	8.03	0.000	.2310592	.3803423
ACC	4156457	.0419797	-9.90	0.000	497926	3333654
BIG	0296347	.0367467	-0.81	0.420	1016584	.042389
B12	.0519029	.0432896	1.20	0.231	0329448	.1367506
PAC	.3448385	.0417614	8.26	0.000	.2629861	.4266908
SEC	.0544844	.0370429	1.47	0.141	0181198	.1270887

### Table 4: Tobit Regression

### Table 5: Slimmed Regression Coefficients

	Instate	InConf
Received Votes	0.10982	0.36033
Тор 25	0.20877	0.37851
Тор 20	0.0857452	0.27387
Тор 15	0.19063	0.13782
Тор 10	0.21975	0.05404
Top 5	0.29707	-0.03506
Top 4	0.23226	-0.03155

### Figure 3



### **5** Limitations

This analysis was first and foremost limited in the number of weeks of poll data it took in. For a more complete analysis of bias in the AP poll, a larger sample size should be utilized. Only three years of data was gathered and analyzed, and only one of those years yielded statistically significant results. Given team performance is relatively consistent from one year to the next, a larger sampling of years would also include greater turnover of teams in the polls. Additionally, the research could be expanded to include other D-1 teams in non-power conferences. Although these teams garner much less attention than the Power 5 teams, they can often appear in the rankings. The inclusion of teams would also bring voters which are not currently factored into the analysis. The current method only examines voters who reside in one of the states home to a Power 5 team, however several of them live in places such as Idaho or Connecticut, which are home to teams only in other conferences. As a result, the rankings from voters in these states could have an impact on the variables, while under the current analysis they do not.

### **6** Conclusion

This research confirms the widely held hypothesis that the AP Poll is a biased measure of college football team strength. This hypothesis has been tested and proven before, such as in the 2007 study by Coleman, Gallo, Mason, and Steagall. Using the same modelling as used in that study, it was found that bias towards teams in the voter's state is prevalent but less in 2018 than in 2007, and bias towards teams in conference's represented in the voter's state is comparable between these two years. This continued evidence of bias in a poll that is intended to be one for qualified experts calls into question the continued use of its results, such as the amount of respect given to the poll's rankings and the impact these rankings have on agreements with television networks as well as the impact that rankings have on recruiting.

Moving forward, analysis of bias in the AP poll can be expanded by taking in additional factors. In the AP basketball poll, a comparable poll of college basketball teams taken weekly during the season, many of the voters are the same as those in the football poll. Since all of the Power 5 teams yield competitive basketball teams, these same schools are often featured in the basketball poll. Because the two polls involve the same voters ranking the same school's teams, it could be found that state and regional bias can be found from the same voters in both polls. Another way to expand this analysis in future studies is to incorporate the CFP rankings into the data set, to see if voters become biased based on how a team is ranked by the CFP committee.

### **Works Cited**

- Andrews, Rodney J., Trevon D. Logan, and Michael J. Sinkey. "Identifying confirmatory bias in the field: Evidence from a poll of experts." Journal of Sports Economics 19, no. 1 (2018): 50-81.
- Campbell, Noel D., Tammy M. Rogers, and R. Zachary Finney. "Evidence of television exposure effects in AP top 25 college football rankings." Journal of Sports Economics 8, no. 4 (2007): 425-434.
- Coleman, B. Jay, Andres Gallo, Paul M. Mason, and Jeffrey W. Steagall. "Voter bias in the associated press college football poll." *Journal of Sports Economics* 11, no. 4 (2010): 397-417.
- "College Poll Tracker: Breaking down College Polls Every Week of the Season." *Collegepolltracker.Com*, 2019, collegepolltracker.com. Accessed 18 Dec. 2019.
- Stone, Daniel F., and Basit Zafar. "Do we follow others when we should outside the lab? Evidence from the AP top 25." Journal of Risk and Uncertainty 49, no. 1 (2014): 73-102.

## Appendix

#### Table A: Acronyms & Abbreviations

TERM	ABBREVIATION
BOWL CHAMPIONSHIP SERIES	BCS
COLLEGE FOOTBALL PLAYOFF	CFP
PAC-12 CONFERENCE	PAC
SOUTHEASTERN CONFERENCE	SEC
ATLANTIC COAST CONFERENCE	ACC
<b>BIG TEN CONFERENCE</b>	BIG
<b>BIG 12 CONFERENCE</b>	B12
ASSOCIATED PRESS COLLEGE	AP Poll
FOOTBALL POLL	

### Table B: Conference Realignment

TEAM	PREVIOUS CONFERENCE	CURRENT CONFERENCE
SYRACUSE	Big East	ACC
PITTSBURGH	Big East	ACC
WEST VIRGINIA	Big East	B12
NOTRE DAME	Big East	ACC
RUTGERS	Big East	BIG
LOUISVILLE	Big East	ACC
NEBRASKA	B12	BIG

COLORADO	B12	PAC
TEXAS A&M	B12	SEC
MISSOURI	B12	SEC
MARYLAND	ACC	BIG
UTAH	Non-Power 5	PAC
TCU	Non-Power 5	B12

### Table C: Findings from "Voter Bias in the Associated Press College Football Poll" by

#### Coleman, Gallo, Mason, and Steagall

		Model I	
Sample size Log likelihood AIC Schwarz criterion	I	n = 71,370 -43760 87580 87856	
Variable	Coefficient	Þ	VIF
Intercept	-23.184	<.0001	
MeanCompRank	-17.171	<.0001	9.76
Statebias	0.645	.0006	1.15
Confbias	0.357	<.0001	1.26
Distance	-0.010	.7504	1.27
Prev_National_air	1.468	<.0001	1.71
Prev_ESPN_ESPN2	1.600	<.0001	3.39
Prev_Other_ESPN	0.946	<.0001	2.64
Prev_Other_major_cable	0.673	<.0001	4.42
Prev_Major_regional	0.776	<.0001	1.87
Prev_other	0.713	<.0001	2.08
National_air	0.391	.0007	1.75
ESPN_ESPN2	0.381	.0014	2.00
Other_ESPN	-0.483	<.0001	1.63
Other_major_cable	0.730	<.0001	1.72
Major_regional	0.105	.5107	1.43
Other	0.252	.3447	1.50

### Table C (Continued)

		Model I	
Played	-0.188	.1578	2.15
Lost_this_week	-0.535	<.0001	1.52
At_least_one_loss	-3.906	<.0001	1.17
Losses_beyond_one	-10. <b>792</b>	<.0001	5.08
ACC	-0.126	.5886	2.22
Big I 2	1.765	<.0001	2.67
BIO	-0.590	.0350	4.66
BigEast	-3.058	<.0001	2.22
Pac I O	0.944	.0002	2.57
SEC	2.793	<.0001	2.15
Prev_ESPN_rep_bias	0.016	.6342	1.73
ESPN_rep_bias	<b>-0.028</b>	.4251	1.75
_Sigma	3.836	<.0001	