

College football, elections, and false-positive results in observational research

Anthony Fowler¹ and B. Pablo Montagnes²

Harris School of Public Policy Studies, University of Chicago, Chicago, IL 60637

Edited by John A. Ferejohn, New York University, Los Altos Hills, CA, and approved September 4, 2015 (received for review February 7, 2015)

A recent, widely cited study [Healy AJ, Malhotra N, Mo CH (2010) *Proc Natl Acad Sci USA* 107(29):12804–12809] finds that college football games influence voting behavior. Victories within 2 weeks of an election reportedly increase the success of the incumbent party in presidential, senatorial, and gubernatorial elections in the home county of the team. We reassess the evidence and conclude that there is likely no such effect, despite the fact that Healy et al. followed the best practices in social science and used a credible research design. Multiple independent sources of evidence suggest that the original finding was spurious—reflecting bad luck for researchers rather than a shortcoming of American voters. We fail to estimate the same effect when we leverage situations where multiple elections with differing incumbent parties occur in the same county and year. We also find that the purported effect of college football games is stronger in counties where people are less interested in college football, just as strong when the incumbent candidate does not run for reelection, and just as strong in other parts of the state outside the home county of the team. Lastly, we detect no effect of National Football League games on elections, despite their greater popularity. We conclude with recommendations for evaluating surprising research findings and avoiding similar false-positive results.

elections | voting | false-positive results

In a recent study, Healy et al. (1) find that college football games influence voting behavior. Victories within 2 weeks of an election seem to increase the success of the incumbent party in presidential, senatorial, and gubernatorial elections in the home county of the team. This study has received significant media coverage [for one illustrative example, see a 2012 article published by *Slate* entitled “Will Ohio State’s football team decide who wins the White House?” (2)] and has been influential among scholars, receiving ~130 citations in 5 years. Several factors contribute to the impact of this particular study. First, the result is surprising and memorable. Second, the results are substantively important for several reasons. Previous studies have shown that bad weather and natural disasters can influence election results (3–5), but such findings do not necessarily show that voters are incompetent or irrational. Governments prepare for and respond to natural disasters, and therefore, disasters could reasonably influence election results by revealing more information about the quality of the incumbent (6). However, the finding that football games influence elections may suggest that voters are incompetent or irrational, because “[u]nlike aberrant weather, local sports outcomes are not something that citizens could expect government to prepare for nor to respond to” (ref. 7, p. 296). Some scholars have gone so far as to suggest that this finding challenges the health of democracy: “This research could describe the reality of democracies as being closer to the worst-case view” (ref. 8, p. 7).

The identifying assumptions of Healy et al. (1) are relatively weak and defensible. However, even with a perfect research design (e.g., a randomized experiment), false-positive results can arise by chance. Because the estimated effect is so striking and because its implications for voter competence are so great, we believe additional scrutiny is warranted. When an empirical result is particularly surprising, it may reflect bad luck on the part of the research

community rather than a real phenomenon, and our goal is to distinguish between these possibilities in this case.

In this study, we reassess the evidence on college football games and elections. Multiple independent sources of evidence suggest that this particular result is a false positive, despite the fact that Healy et al. (1) used a credible research design and followed the best practices of social science. There is no single smoking gun. Rather, we test several additional hypotheses that should hold if college football games truly influence elections, and in each case, the evidence fails to support the original finding. We detect no effect of college football games on elections when we leverage situations where multiple elections take place in the same county and year but the incumbent parties differ. We find that the estimated effect is greater in counties that are less interested in college football, just as great even when the incumbent does not run for reelection, and just as great in counties outside the home county of the team. Lastly, we find no effect of National Football League (NFL) games on elections, despite the greater popularity of the NFL over college football. If college football games indeed influence elections, we think it would be virtually impossible to explain all of these results. Rather, we conclude that there is no meaningful effect of college football games on elections, and a false-positive result arose simply by chance.

Our investigation, while focusing on a narrow question about college football games and elections, may hold broader lessons for empirical research. Social scientists have made significant progress in developing credible research designs for making causal inferences from observational data (9), but this study suggests that a good research design is not sufficient for credible results. Recently, researchers have pointed out that many published, experimental results in medicine (10) and psychology (11, 12) are false, and the same concerns apply broadly to experimental and non-experimental findings in all academic disciplines. In the case of

Significance

We reassess the surprising finding that college football games influence incumbent support in subsequent elections. Because independent replication is impossible for such nonexperimental findings, we proceed by testing multiple independent hypotheses that should hold if college football games indeed influence elections. In each case, the evidence suggests that the original finding was a false positive. For example, if college football games indeed influence elections, this effect should be greatest in places where the public is most interested in college football. However, we find the opposite. We conclude with general recommendations for evaluating surprising research findings and avoiding similar false-positive results—particularly for non-experimental work in the social sciences where independent replication is impossible.

Author contributions: A.F. and B.P.M. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. Email: anthony.fowler@uchicago.edu.

²Present address: Department of Political Science, Emory University, Atlanta, GA 30322.

experimental research, false-positive discoveries can be identified and corrected through replication, but this option is typically not available for the observational studies most commonly produced by empirical social scientists. Therefore, we conclude with recommendations for researchers and scholars who hope to identify and avoid false-positive results in these settings. Specifically, we discuss the importance of theory in generating ex-ante, comparative-static predictions that can serve as independent tests of a hypothesis and lend additional credibility to empirical findings.

Results

All empirical results are shown in Table 1. We independently assembled data on football results and county-level election returns from 1960 to 2012. [Healy et al. (1) analyze presidential elections from 1960 to 2004 and senatorial and gubernatorial elections from 1967 to 2006. To be thorough, we expand the period of analysis and include all presidential elections from 1960 to 2012 and all senatorial and gubernatorial elections from 1960 to 2006. None of our subsequent results are meaningfully affected by using our period of analysis vs. that of Healy et al. (1).] Healy et al. (1) obtain their strongest empirical result when pooling both games from the 2 weeks leading up to an election, typically played on the Saturdays 3 and 10 days before the election. For each team and week, we code a variable that takes a value of 1 if the team won its game, 0 if the team lost, and 0.5 if the team tied. Then, for each team and year, we calculate the average of these values for the two games preceding an election. This produces a variable that ranges from 0 to 1 and takes the possible values of 0, 0.25, 0.5, 0.75, and 1. A value of 1 indicates that the team won both games preceding the election, a value of 0 indicates that a team lost both games, and so on. We call this variable Football. In the end, our primary dataset includes 2,408 team–election observations where a major conference, Division 1A college football team played games in both weeks preceding a presidential, senatorial, or gubernatorial election in their home state (Table 2 shows the list of teams included in our analysis). Following the work by Healy et al. (1), we exclude the University of Southern California and the University of California, Los Angeles, because they reside in the same county.

To replicate the main result by Healy et al. (1), we conduct a differences-in-differences regression with fixed effects following the specification of Ansolabehere and Snyder (13). We code a binary variable that takes a value of 1 if the Democratic Party is the incumbent party and a value of -1 if the Republican Party is the incumbent party, and we call this variable Incumbent Party (IncParty). We then regress the Democratic Party's share of the two-party vote (ranging from 0 to 1) on IncParty, Football, and the interaction of these two variables. We include county fixed effects and year–office fixed effects to account for the partisanship of different counties and the fact that some years are good for a particular party across the country. The quantity of interest is the estimated coefficient associated with the interaction term—IncParty \times Football, because it tells us how support for the incumbent party varies with college football games. In column 1 of Table 1 (labeled “Baseline”), we estimate a coefficient of .015, suggesting that, as a team goes from losing both games to winning both games, the incumbent party receives an extra 1.5% of the two-party vote. This result is statistically significant at the 0.05 level and similar to the pooled results presented by Healy et al. (1). Substantively, this effect size implies that about 1 in 130 voters will support the incumbent party if their local college football team has won its last two games but will vote against the incumbent party if the team has lost its last two games. This effect size is comparable to about half the estimated effect of a large-scale television advertising campaign (14), an extra half a percentage point of real income growth (15), and about one-quarter of the incumbency advantage (13).

The baseline estimate could reflect a true phenomenon, whereby college football games influence elections. Alternatively, it could simply reflect bad luck, and a closer analysis of the data may allow us to distinguish these possibilities. One way that this result could have manifested itself, even if there is no real effect, is through a chance correlation between football games and state- or county-level partisan shocks. What if college football teams in Democratic places happened, by chance, to win in years that were good for Democrats or vice versa? In column 2 of Table 1 (labeled “Within county–year”), we estimate the effect of college football games with an alternate research design,

Table 1. College football and incumbent support

	Baseline	Within county–year	By interest	By incumbent running	By home county	NFL
IncParty	.058* (.005)	.056* (.007)	.044* (.004)	–.006 (.007)	.057* (.001)	.043* (.004)
Football	–.008 (.008)		.000 (.010)	–.013 (.012)	–.005* (.001)	–.003 (.008)
IncParty \times Football	.015[†] (.007)	.003 (.011)	.026* (.008)	.021 (.012)	.011* (.001)	.004 (0.007)
IncParty \times High Interest			.036* (.010)			
Football \times High Interest			–.019 (.016)			
IncParty \times Football \times High Interest			–0.029[†] (.014)			
IncRunning				–.019 [†] (.009)		
IncParty \times IncRunning				.092* (.008)		
Football \times IncRunning				.007 (.013)		
IncParty \times Football \times IncRunning				–.007 (.013)		
Home County					.002 (.005)	
IncParty \times Home County					–.007 (.005)	
Football \times Home County					–.003 (.009)	
IncParty \times Football \times Home County					.002 (.007)	
Year–office fixed effects	X	X	X	X	X	X
County fixed effects	X		X	X	X	X
County–year fixed effects		X				
Observations	2,408	2,408	2,408	2,408	227,936	905
R ²	0.472	0.832	0.478	0.516	0.610	0.640

County-clustered SEs are in parentheses; details are in the text. Substantively relevant coefficients are formatted in bold.

* $P < 0.01$.

[†] $P < 0.05$.

which only draws inferences from situations where multiple elections (e.g., senate and gubernatorial) take place in the same state and year and the incumbent parties differ between those races. There are 496 such cases in our data. We accomplish this by including county-year fixed effects instead of county fixed effects. If college football games really influence incumbent support, we should see it clearly in these cases. Wins should differentially help Democratic candidates when the Democratic Party previously held a seat and Republican candidates when the Republican Party previously held a seat. Alternatively, if the original result arose because football games happened to be correlated with local partisan shocks by chance, then we should not observe this phenomenon. Table 1 shows that we detect no such effect with this research design, and our statistically insignificant result is not the product of noise. The estimated coefficient of .003 is substantively small—one-fifth the magnitude of the baseline estimate.

In columns 3–5 of Table 1, we test for variation in the effect of college football games. If such an effect exists, it should be strongest in places where citizens care the most about college football. Table 2 presents an approximate measure of interest in each team's home county—the percentage of Facebook users in that county who express an interest in college football through their “likes.” We see that interest ranges significantly from 3% in the case of Boston College to 43% in the case of Auburn. The face validity of this measure is quite high. Teams that are historically less popular

or teams in large, urban counties are measured as relatively unpopular in their home counties. Teams in small, rural counties with strong football programs are significantly more popular in their home counties. To simplify our analysis, we define a variable—High Interest, which takes a value of 1 if a team is above the median in popularity and 0 if a team is below or equal to the median. We then conduct a triple-interactive regression, where all combinations of IncParty, Football, and High Interest are included in the regression, including all three multiplied together. (The coefficient associated with High Interest alone is subsumed by the county fixed effects and therefore, not shown in Table 2.) The coefficient associated with IncParty \times Football indicates the estimated effect of college football games in the low-interest counties, and the coefficient associated with IncParty \times Football \times High Interest indicates the differential effect in high- vs. low-interest counties. We estimate a 2.6 percentage point effect in low-interest counties, and we estimate that this effect is 2.9 percentage points lower (therefore, -0.3 percentage points) in high-interest counties. In other words, the estimated effect of college football games is only detectable for the teams and counties where we would least expect it, like Boston College, Northwestern, Pittsburgh, and Rutgers, whereas no effect is detectable for the teams and counties where we would most expect it, like Nebraska, Alabama, Mississippi, and Auburn. (We code interest as binary for simplicity, but alternative cutoffs or continuous measures of interest produce the same pattern of results.)

Using a similar strategy in column 4 of Table 1 (labeled “By incumbent running”), we test whether the estimated effect of college football games is greater when an incumbent runs for reelection. If college football games influence the welfare of voters and if they unfairly incorporate their euphoria or frustration into their evaluations of incumbents, this effect should be smaller when no incumbent runs for reelection. However, column 4 of Table 1 shows that the estimated effect of college football is 2.1 percentage points when the incumbent is not on the ballot, and this effect is 0.7 percentage points smaller when the incumbent actually runs. In other words, the estimated effect is the same, if not larger, when no incumbents are being evaluated. If the effect of college football on elections is real, it seems highly unlikely that this effect would be just as great for new candidates from the incumbent party as for incumbents themselves.

Following the work by Healy et al. (1), the previous analyses only included the home counties of college football teams, but in column 5 of Table 1 (labeled “By home county”), we estimate the effect of college football games for all counties within a team's state. Each observation represents a team-county-election, and county-elections are repeated in cases where multiple teams reside in the same state. We then test whether the effect of college football games is greater in the home county of the team than in other parts of the state. If college football games influence elections, we would expect the effect to be smaller in other parts of the state, especially because many states have multiple teams, and support for college football teams is highly regionalized, even within a state (16). Alternatively, if the original result arose, because for example, the local football team happened to win by chance when there was already a popular candidate from the incumbent party on the ballot, then we would expect to find a similar effect outside the home county. The coefficient associated with IncParty \times Football shows that we detect a large and statistically significant estimate of the effect of football games in other counties outside the home county of the state, and the coefficient associated with IncParty \times Football \times Home County shows that this estimated effect is not statistically greater in the home county of the team.

In column 6 of Table 1 (labeled “NFL”), we replicate the baseline regression for the NFL. If irrelevant events influence elections, we would expect NFL games to have a greater effect than college football games, because the NFL is significantly

Table 2. College football teams in sample by interest

Lower interest teams	Like (%)	Higher interest teams	Like (%)
Boston College	3	Baylor	25
California	4	Georgia	25
Minnesota	4	Kentucky	25
Northwestern	4	Ohio State	25
Rutgers	4	Tennessee	25
Stanford	4	West Virginia	25
Maryland	7	Wisconsin	25
Miami (FL)	7	Missouri	26
Indiana	8	Notre Dame	26
Southern Methodist	8	South Carolina	27
Washington	8	Texas A & M	27
Colorado	9	Iowa	28
Pittsburgh	10	Virginia Tech	28
Arizona State	11	Arkansas	29
Cincinnati	11	Florida	29
Texas Christian	13	Penn State	29
Brigham Young	15	Florida State	30
Georgia Tech	15	Louisville	30
Illinois	15	Oregon	30
Utah	15	Texas Tech	30
Vanderbilt	15	Nebraska	31
Arizona	16	Oregon State	31
Kansas State	16	Alabama	32
Duke	17	Michigan State	32
North Carolina State	17	Iowa State	33
Wake Forest	17	Clemson	34
Purdue	19	Mississippi	34
Texas	19	Oklahoma	35
Virginia	19	Mississippi State	36
Syracuse	20	Oklahoma State	36
Kansas	21	Auburn	43
Louisiana State	24		
Michigan	24		
North Carolina	24		
Washington State	24		

more popular—television ratings are ~10 times greater—and NFL teams receive strong regional support just like college teams. However, we detect no effect of NFL games on elections. Furthermore, the null result for the NFL is not the product of imprecision—our estimated standard error is identical to that for college football.

Discussion

We reassess the evidence that college football games influence elections to determine whether the originally published finding reflects a real phenomenon or a false positive. In an experimental setting, researchers can reassess published results through replication. However, in an observational setting, such as this one, we cannot replay the last 50 years of football games and elections. Therefore, we proceed by testing additional hypotheses that should hold if college football games indeed influence elections. Specifically, we posit that this effect should be detectable in cases where there are multiple elections in the same time and place but the incumbent parties differ. This effect should also be greater for counties with greater interest in college football, greater when the incumbent candidate runs for reelection, and greater in the home county of the team. Furthermore, if college football games influence elections, we would expect NFL games to exhibit a comparable, if not greater, effect. However, we find no support for any of these hypotheses. Our conclusion is that college football games do not meaningfully influence elections and that the original findings by Healy et al. (1) were a false positive that arose through bad luck.

Additional Evidence from Healy et al. (1). To their credit, Healy et al. (1) show several additional tests intended to lend additional credibility to their design and results. How should we interpret these additional tests in light of our results? First, Healy et al. (1) present placebo tests showing that football games after the election seem to exhibit no effect on the previous election—a finding that we replicate. These tests lend additional credibility to the research design and inferential strategy by showing that football games, in expectation, are unrelated to the typical voting behavior of the county. However, these placebo tests say nothing about the possibility that their result is a false positive. These null placebo results are equally likely whether the original finding reflects a real phenomenon or arose by chance. If the original finding was obtained through bad luck, we would not expect the same misfortune to manifest itself with another independent draw of games.

Again, lending additional credibility to their research design and assumptions, Healy et al. (1) show that their results are largely unchanged when they include demographic controls, fixed effects, or control for game expectations. These tests mitigate concerns about confounding variables, but again, they say nothing about the possibility of a false-positive result. If football games are exogenous but happened to be correlated with election results by chance, we would expect these correlations to persist across specifications as Healy et al. (1) control for additional covariates.

Healy et al. (1) also test for variation in the effect of football games across the timing of the game. Healy et al. (1) show that the effect of the game 3 days before the election is actually weaker than the effect of the game 10 days before the election, and they explain this finding by arguing that voters may have already made up their minds 3 days before the election. However, before seeing their results, if football games do influence voter behavior, we would have most likely expected to detect the greatest effect immediately before the election. If anything, the campaign literature suggests that events closer to the election should be more influential, because persuasion effects tend to decay quickly (14). This example illustrates the value of theory and the generation of ex-ante predictions that are not subject ex-post rationalization. Furthermore, by separately testing each week and the average of the 2 weeks without subsequently adjusting their *P* values for multiple testing,

Healy et al. (1) increased their chances of obtaining a false-positive result (11).

Lastly, Healy et al. (1) conduct several tests in the spirit of our exercise. Theoretically, as we explain above, if college football games do influence elections by affecting voter mood, we would expect to see the greatest effects for teams in which a larger share of the county residents cares about the local team. This prediction is the underlying motivation for our test in column 3 of Table 1. Healy et al. (1) do not collect data on the proportion of county residents who care about the local team, but they do find that their estimated effects are greater for championship teams and high-attendance teams. However, winning a national championship is weakly correlated with county-wide interest, partly because teams in large, urban counties, like Pittsburgh, Georgia Tech, and Miami, have won national championships. Therefore, finding that the estimated effect is greater for championship teams is not necessarily consistent with their purported mechanism. Additionally, we are unable to replicate their result for high-attendance teams. With our data, we actually obtain smaller estimates for high-attendance teams than for low-attendance teams.

Implications. Substantively, our results suggest that voters are more competent than previously thought. Because voting behavior seems to be influenced by the economy (17), natural disasters (3–5), and the performance of elected officials (8) but not by football games, voters may be reasonably capable of distinguishing irrelevant factors from those for which the government can influence, prepare for, or respond. Of course, voters are imperfect; they may improperly weigh relevant factors (4, 18), and they can be influenced by irrelevant factors in laboratory settings (19). Nonetheless, we have little evidence that irrelevant factors meaningfully influence voting in real-world elections.

Our results also hold general methodological implications for empirical research in social science. Healy et al. (1) used a credible research design and followed the best practices in their field. Nonetheless, we conclude that their result reflects a chance false positive rather than a true phenomenon. False-positive results can arise in well-executed studies through bad luck, and they are more likely when the research community tests many hypotheses where little effect is expected a priori. These concerns become even greater as the “big data revolution” and improved computational ability allow researchers to test more hypotheses than ever before (20), and publication bias skews the reporting of positive and null results (21). Furthermore, widespread testing for unlikely effects is virtually guaranteed so long as there is demand for exciting, surprising results among editors, referees, scholars, journalists, and the public.

How should researchers proceed when they worry that an empirical result may be a false positive? For experimental studies, replication is the standard way to eradicate skepticism, but this option is not available for observational work in the social sciences. For example, we cannot rerun the past 50 years of college football games and elections to see if the same pattern manifests itself again. Therefore, we recommend that researchers proceed by conducting the kinds of theory-informed tests presented in this paper. There are often additional hypotheses that should hold if an estimated effect is genuine. These kinds of additional tests can lend significant credibility to empirical results or suggest that the original finding was spurious. (Of course, our recommendations are not a panacea. The research community could still obtain false-positive results through particularly bad luck or the selective presentation of confirmatory results. Nonetheless, if authors, referees, and readers think carefully about theory and ex-ante predictions, many false-positive results may be avoided.) These tests should be motivated by careful theorizing and clear ex-ante predictions that can lead the original hypothesis to be supported or falsified. Importantly, these tests should be largely independent of the original finding. For example, although alternative specifications and placebo tests may assess the validity of the research

design, they do provide independent tests of a hypothesis. Under some conditions, the additional tests recommended in this paper may hold the same value for observational researchers that replication does for experimentalists. Accordingly, we recommend that editors, referees, scholars, and the broader audience demand this kind of additional evidence when evaluating empirical results.

Materials and Methods

College football results were obtained through www.sports-reference.com, NFL results were obtained through www.profootball-reference.com, and election results were obtained from Jim Snyder, who collected official election

returns from the websites of state secretaries. College football teams were matched to counties using the official mailing addresses of the respective universities. Estimates of college football fandom by county come from Irwin and Quely (22), who tabulated the percentage of Facebook users in each county who, based on their “likes,” are fans of a college football team. Table 1 presents the results of seven different ordinary least squares regressions. The details of each regression are presented in *Results* and Table 1.

ACKNOWLEDGMENTS. We thank Scott Ashworth, Chris Berry, David Brookman, Ethan Bueno de Mesquita, Andy Hall, Will Howell, Gabe Lenz, Max Palmer, and John Patty for valuable feedback, and we thank Jim Snyder for providing election data.

1. Healy AJ, Malhotra N, Mo CH (2010) Irrelevant events affect voters' evaluations of government performance. *Proc Natl Acad Sci USA* 107(29):12804–12809.
2. Cowen T, Grier K (October 24, 2012) Will Ohio State's football team decide who wins the White House? *Slate*, Politics, Who's Winning, Who's Losing, and Why. Available at www.slate.com/articles/news_and_politics/politics/2012/10/how_the_presidential_race_between_barack_obama_and_mitt_romney_could_be.html.
3. Bechtel MM, Hainmueller J (2011) How lasting is voter gratitude? An analysis of the short- and long-term electoral returns to beneficial policy. *Am J Pol Sci* 55(4):851–867.
4. Healy A, Malhotra N (2009) Myopic voters and natural disaster policy. *Am Polit Sci Rev* 103(3):387–406.
5. Healy A, Malhotra N (2010) Random events, economic losses, and retrospective voting: Implications for democratic competence. *Quart J Polit Sci* 5(2):193–208.
6. Ashworth S, Bueno de Mesquita E, Friedenberg A (2015) *Learning About Voter Rationality*. Working paper (University of Chicago, Chicago). Available at home.uchicago.edu/~sashwort/shocks.pdf.
7. Healy A, Malhotra N (2013) Retrospective voting reconsidered. *Annu Rev Polit Sci (Palo Alto)* 16:285–306.
8. Lenz GS (2012) *Follow the Leader? How Voters Respond to Politicians' Policies and Performance* (Univ of Chicago Press, Chicago).
9. Angrist JD, Pischke J (2010) The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *J Econ Perspect* 24(2):3–30.
10. Ioannidis JPA (2005) Why most published research findings are false. *PLoS Med* 2(8):e124.
11. Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 22(11):1359–1366.
12. Simonsohn U, Nelson LD, Simmons JP (2014) P-curve: A key to the file-drawer. *J Exp Psychol Gen* 143(2):534–547.
13. Ansolabehere S, Snyder JM (2002) The incumbency advantage in U.S. elections: An analysis of state and federal offices, 1942–2000. *Election Law J* 1(3):315–338.
14. Gerber AS, Gimpel JG, Green DP, Shaw DR (2011) How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment. *Am Polit Sci Rev* 105(1):135–150.
15. Hibbs DA (2012) Obama's reelection prospects under “bread and peace” voting in the 2012 US presidential election. *PS Polit Sci Polit* 45(4):635–639.
16. Giratikanon T, Katz J, Leonhardt D, Quealy K, Tracy M (October 3, 2014) NCAA fan map: How the country roots for college football. *The New York Times*, The Upshot. Available at www.nytimes.com/interactive/2014/10/03/upshot/ncaa-football-fan-map.html?_r=0.
17. Bartels LM (2008) *Unequal Democracy: The Political Economy of the New Gilded Age* (Princeton Univ Press, Princeton).
18. Healy A, Lenz GS (2014) Substituting the end for the whole: Why voters respond primarily to the election-year economy. *Am J Pol Sci* 58(1):31–47.
19. Huber GA, Hill SJ, Lenz GS (2012) Sources of bias in retrospective decision-making: Experimental evidence on voters' limitations in controlling incumbents. *Am Polit Sci Rev* 106(4):720–741.
20. Efron B (2013) *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction* (Cambridge Univ Press, Cambridge, United Kingdom).
21. Gerber A, Malhotra N (2008) Do statistical reporting standards affect what is published? Publication bias in two leading political science journals. *Quart J Polit Sci* 3(3):313–326.
22. Irwin N, Quely K (November 8, 2014) The places in America where college football means the most. *The New York Times*, The Upshot. Available at www.nytimes.com/2014/11/08/upshot/the-places-in-america-where-college-football-means-the-most.html.