Matching Pennies on the Campaign Trail: An Empirical Study of Senate Elections and Media Coverage*

Camilo García-Jimeno[1] and Pinar Yildirim[2]

[1]Department of Economics, University of Pennsylvania and NBER

[2]Wharton School, University of Pennsylvania

This Version: September 6, 2017, First Version: October, 2015

Abstract

We study the strategic interaction between the media and Senate candidates during elections. While the media is instrumental for candidates to communicate with voters, candidates and media outlets have conflicting preferences over the contents of the reporting. In competitive electoral environments such as most US Senate races, this can lead to a strategic environment resembling a matching pennies game. Based on this observation, we develop a model of bipartisan races where media outlets report about candidates, and candidates make decisions on the type of constituencies to target with their statements along the campaign trail. We develop a methodology to classify news content as suggestive of the target audience of candidate speech, and show how data on media reports and poll results, together with the behavioral implications of the model, can be used to estimate its parameters. We implement this methodology on US Senatorial races for the period 1980-2012, and find that Democratic candidates have stronger incentives to target their messages towards turning out their core supporters than Republicans. We also find that the cost in swing-voter support from targeting core supporters is larger for Democrats than for Republicans. These effects balance each other, making media outlets willing to cover candidates from both parties at similar rates.

Keywords: Matching Pennies, Political Campaigns, Senate Elections, Media

JEL Codes: C3, D72

*We thank Jeff Groesback, Scott Wang, and Shawn Zamechek for excellent research assistance, and Noah Veltman for sharing with us the data on NFL fandom. We are also grateful to Daron Acemoglu, Frank DiTraglia, Gregory Martin, Maria Petrova, Carlo Prato, Jesse Shapiro, Noam Yuchtman, and to seminar participants at Tufts University, CEMFI, Pompeu Fabra, Stockholm University, the Wharton School, Princeton University, NYU Abu Dhabi, UC Berkeley, the Wallis Political Economy conference at the University of Rochester, the Southern Economic Association 2015 Conference, and the Stanford University SITE conference for their valuable comments. We gratefully acknowledge the financial support of the NET Institute, the Dean's Research Fund at the Wharton School, and the Hal Varian Visiting Assistant Professorship research fund at MIT.

1 Introduction

Scholars agree that free speech, and a free media as its main conduit, are necessary for a well-functioning liberal democracy. One key role attributed to the media is that of serving as a watchdog monitoring leaders' behavior while in office. The media also plays a role in shaping the behavior of candidates during campaigns. It is especially involved in reporting about politics during elections, and invests heavily in covering them. Candidates are particularly aware of the way and extent to which the media reports on them. Most modern campaigns invest heavily in media relations, and hire specialized staff focused on dealing and communicating with reporters. It is not uncommon for politicians to be escorted into campaign locations with political consultants who curate speeches by choosing or omitting words, phrases, and issues.

Electoral campaigns are not just a game between candidates, but a highly strategic one between candidates and the media. Candidates use the media to communicate with their constituencies, and the media uses candidates as sources of news (Bartels (1996); Prat and Strömberg (2013)). However, this relationship is not purely symbiotic. Although candidates and media outlets both share the objective of generating news, their preferences can be misaligned regarding the contents of such reporting. First, there are political scandals, which the media is particularly eager to report about (Fonseca et al. (2014)). The breaking of such events depends on the previous history of the candidates, and on the increased spotlight at which the campaign itself puts them. Second, and of more relevance for this paper, throughout their campaigns candidates need to target a heterogeneous electorate. Particularly in competitive and bipartisan races, candidates require support from centrist or non-ideological voters as much as from more ideological or core supporters. The literature often calls the former swing voters because they are likely to switch their vote across candidates. Core supporters, in contrast, are unlikely to switch party allegiances, making turnout the relevant margin of their decision (Cox and McCubbins (1986), Lindbeck and Weibull (1987)).

As a result, candidates have incentives to differentiate their message, especially if it can be targeted towards core or swing voters. The media, in turn, produces public information which constrains the targeting ability of candidates. If messages targeted to core supporters happen to be widely reported by the media, the cost in electoral support among swing voters may be larger than any benefits that may be reaped from additional core supporters.[1] On the other hand, following and reporting on candidates is costly for media outlets. When the media devotes more attention to a candidate it may reduce the candidate's incentive to

[1] A recent example of this was Mitt Romney's "the other 47%" statement during a private fundraiser in Boca Raton during the 2012 U.S. presidential election. Although intended for a very narrow audience of wealthy individuals, its public revelation led to a significant backlash and widespread media coverage.

target core supporters, making the gains from such a strategy relatively low. In this paper we highlight this trade off and explore empirically how it can shape the campaigning behavior of candidates and the reporting strategies of media outlets, in the context of U.S. Senate races.

We begin by arguing that these considerations make bipartisan campaigns closely resemble the strategic environment of a classic matching pennies game between each candidate and the media. Regardless of whether they desire to inform the public or simply earn revenue, as long as there is sufficient heterogeneity in media slant across outlets, the average media outlet is likely to profit more by reporting or covering statements directed towards core supporters. These are especially informative for swing voters, and also of interest to core supporters. In contrast, campaign messages targeted towards swing voters are likely more 'centrist' and as such, of little interest to core supporters and of low informational value to swing voters. Candidates, on the other hand, may benefit electorally from campaign speech targeted towards core supporters as long as it is not widely reported by the press. The implication of these conflicting objectives is that candidates' incentives to target core constituencies will be determined by how profitable it is for the media to report on them. Likewise, the media's incentives to invest in covering campaigns will be determined by the candidates' relative profitability from targeting core supporters relative to swing voters with their campaign speech. In such a strategic environment, both candidates and media outlets will have strong incentives to behave in a relatively unpredictable way (to the opponent).

In this paper we develop a simple model of bipartisan electoral races with media coverage and unidimensional policy which we then test empirically. Policy positioning by candidates happens through their campaign statements, but differences in the core supporter and swing-voter responses to campaign-trail statements put a limit to full policy convergence. Voters perceive a candidate as more or less centrist as a function of the history of statements they have had access to, either through the media or through direct contact with the candidate. Voters express their preferences throughout the campaign by responding to polls, and lastly, by voting on election day. Media outlets decide on the intensity with which they will cover each candidate, and candidates decide the type of statements to make at every date during the campaign. Modeling electoral races in this way suggests a new meaning for the role of the media in constraining politicians' behavior, different from the one emphasized in the standard political agency framework (e.g., Ferraz and Finan (2011); Snyder and Strömberg (2010)). Moreover, it suggests a novel channel through which politicians can influence the media's behavior that is unrelated to the corruption or influence-buying channels emphasized in the literature.

We implement this methodology on U.S. Senatorial races for the period 1980-2012. U.S.

Senate races offer an ideal empirical setting; most Senate races are high profile, and thus, enjoy ample media and polling coverage. We show how data on media reports, electoral and poll results, and an exogenous source of variation in the profitability of campaign reporting by the media, together with the behavioral implications of the model, can be used to estimate its structural parameters. We estimate a discrete game of complete information (see Bajari et al. (2010)) with several novelties. First, the nature of the environment allows us to study a repeated (and subsequently dynamic) game in a very parsimonious way. This is because matching pennies games have a unique Nash equilibrium in mixed strategies, and naturally, electoral campaigns are finite in time as they end on election day. Thus, an unraveling argument implies that the repeated (and dynamic) game will also have a unique sub-game perfect equilibrium, hugely simplifying estimation. Second, our empirical strategy allows us to overcome a pervasive problem faced by the literature on identification in discrete games of this nature, when frequencies for a subset of game outcomes are unobserved.[2]

In practice, we estimate a linear model (for the technology mapping equilibrium play to the electoral outcome of the race) where we allow media outlets to exhibit asymmetric payoffs from covering Republican and Democratic candidates. We call this media bias, and show that although it is only partially identified, we can nevertheless recover and compute its identified set, the equilibrium strategies of all players, and all candidate payoff parameters governing the game.[3] These directly measure the average impact that different types of statements and media reports have on the poll standings and electoral performance of candidates. As will become clear below, we make the linearity assumption only to make our identification arguments transparent. Our identification strategy closely relies on mapping conditional probabilities to observed frequencies, and on the availability of an exogenous shifter of the media's payoffs. Following ideas in Eisensee and Strömberg (2007) among others, we collected detailed and high frequency data on sports events from the four most important sports leagues in the US (the NCAA, NBA, MLB, and NFL). We use these data as sources of exogenous variation in the media's willingness to cover and report on politics during campaigns.

[2] For example, in entry games, data on the number of firms which decide not to enter in a given period is necessarily unobserved. Similarly, in our model the media chooses whether or not to cover a given candidate. In periods when the media does not report, we do not observe the type of statement made by the candidate. As a result, we observe the frequencies for different types of statements only conditional on the media reporting. Nevertheless, we overcome this difficulty tracking changes in candidate support based on poll data over the campaign, which we argue are responsive to the full distribution of statements made, and thus, allow us to to recover all relevant payoff parameters.

[3] The literature has emphasized how media bias or slant affect the media market (e.g.,Baron (2006); Mullainathan and Shleifer (2005)). In these papers media bias refers to an outlet's preference for 'spinning' the information they receive when reporting it. Here our definition of media bias refers only to the media's relative preference for the extent of reporting about either political party.

Our model and estimation require that we measure a signal of the *intended target* of each instance of a candidate's campaign trail speech as reported by the media. As such, our paper also contributes to the literature by developing and implementing a novel text-analysis methodology in the spirit of Gentzkow and Shapiro (2010). Our proposed methodology allows us to assess whether the contents of a given news piece mentions candidate speech suggestive of swing voter or core supporter targeting. The key idea is to create a self-referential measure. We first compute the most commonly used phrases related to policy issues found within the universe of written media coverage of each Senate race (state-x-year). We also compute the relative counts of candidate name mentions in each article to measure the extent to which it reports coverage of the Democratic or the Republican candidate. We then assign a score to each article based on the frequency with which it contains phrases that are relatively commonly used in articles that mention more heavily one or the other candidate. We can then use this index to classify each news piece as likely reporting on a core-supporter or a swing-voter targeted candidate statement. We use different classification criteria to explore the robustness of our results.

Any attempt to estimate the electoral or poll responsiveness of different types of voters to candidates' campaign promises faces a severe identification challenge: unobservables may drive both the types of statements candidates make and their electoral performance. This empirical challenge is even more serious when the researcher's ability to measure candidates' campaign statements depends on the endogenous decisions of media outlets on whether to report on those statements or not. Our methodology allows us to overcome these identification challenges.

We use our parameter estimates and empirical findings to assess the impact of partisan bias in the media, the importance of race characteristics such as the ideological distribution of voters, and technological innovations altering the cost of media reporting. Our findings point to a large asymmetry in the strategic environment faced by Democratic and Republican Senate candidates: on average, the turnout responsiveness of Democratic core supporters is much larger than that of Republican core supporters, possibly because of the lower turnout rates of traditionally Democratic demographics.[4] This gives strong incentives for the media to cover Democratic candidates more intensely. Moreover, we also find that swing-voters punish Democratic candidates more strongly for widely reported core supporter-targeted campaign speech. As a result, Democratic candidates' speech is sufficiently disciplined that in equilibrium Democratic and Republican candidates face similar rates of media coverage

[4]This finding is consistent with previous empirical studies showing the effect of the media on voting behavior takes places especially through increased turnout (see George and Waldfogel (2006); Oberholzer-Gee and Waldfogel (2009); Strömberg (2004a)).

and reporting. We also find little evidence of significant media preference for reporting on candidates of one party over the other.

Although we are unaware of any other study modeling the relationship between politicians and the media in the way we do here, nor estimating the effect of media campaign coverage on electoral outcomes within a structural model, our paper relates to several research areas. Foremost, this paper is related to the literature on media coverage (Gentzkow and Shapiro (2006); Puglisi and Snyder (2008); Strömberg (2004a)). Most of this literature separately endogenizes policy choices by politicians or coverage decisions by the media. In contrast, we directly explore the simultaneous determination candidates' choices and media coverage strategies. In a recent related contribution, Durante and Zhuravskaya (forthcoming) show that in the context of the Israel-Palestine conflict, the Israeli government targets the timing of its attacks strategically as to coincide with prominent and predictable events in the U.S. The empirical literature also has measured the impact of media coverage on policy outcomes (Snyder and Strömberg (2010); Strömberg (2004b)). Instead, we focus on the impact of media coverage on candidate behavior. Thus, our model is close in spirit to the ideas in Ansolabehere et al. (1992), according to whom "... some of the most crucial interactions in campaigns are those between candidates and reporters... campaign organizations seek to spoon-feed the press in order to control the news coverage their candidates receive. Journalists react by striving to keep candidates off balance through independent reporting" (pg. 72). Another related paper is Fonseca et al. (2014), who study the partisan bias in newspaper coverage of political scandals in the late 19th Century U.S. They find significant bias in reporting depending on newspaper partisanship. While they focus on political scandals only, here we focus on the media's coverage choices over any candidate-related content. From the theory side, the literature on issue selection has emphasized how informational frictions between voters and candidates may affect equilibrium campaign message choices (Egorov (2015)).

Our paper also relates to the literature on transparency, which asks how increased information affects policy outcomes (Maskin and Tirole (2004); Prat (2005)). Most insights in this literature follow closely those from the contract-theory literature on agency. In our model, an increase in the amount of information generated in equilibrium comes from more intense media reporting, which happens when candidates' payoffs from speech targeted to core supporters are higher. Thus, more information may be correlated with more extreme platform choices by politicians. Of course, if there is no relationship between what candidates say during campaigns and what they do while in office, understanding the forces shaping campaign speech would be uninformative about the media's role in shaping policy. Nevertheless, voters appear to care significantly about what candidates say, and the literature does suggest there is a close relationship between campaign speech and policy choices (Budge and Hofferbert (1990); Kurkones (1984)).

Also within the tradition of political agency, Besley and Prat (2006) develop a model where the media plays the key role of supplying voters information on incumbent behavior they use when deciding whether to retain or dismiss him. When the incumbent is able to influence the media's information supply decision, it can undermine democracy's ability to exert agency control. In this literature, competition in the media market limits the extent of media capture, and thus, of selection of the information supplied to the public (See for example Chiang and Knight (2011); Corneo (2006)). In contrast, we show that highly selected media content can arise even in highly competitive media markets. Our paper also directly fits within the literature studying how the media affects citizens' opinions and electoral choices (e.g.,Campante and Hojman (2010); DellaVigna and Kaplan (2007); Enikolopov et al. (2011)), and is related to the strand of the literature measuring the media's ideological positions (Gentzkow and Shapiro (2010); Puglisi (2006)). Instead of attempting to measure the ideological positions of different media outlets, we measure the extent of reporting on speech targeted to different voters by the media as a whole. Although our paper does not directly study voter learning and the extent to which voters react to new information during campaigns, our ability to establish an empirical link between overall news reports and poll changes indirectly suggests voter responsiveness to information, similar to the findings in Hirano et al. (2015) who study voter learning during primaries and find strong effects for statewide offices.

Here we model the interaction between politicians and the media as a matching pennies game. Thus, our paper also is related to the literature that has empirically studied this kind of strategic environment and the mixed-strategy equilibria it is associated with. Walker and Wooders (2001) were the first to look for empirical evidence of mixed-strategy behavior by studying serving on Wimbledon tennis matches. In a very different context, Knowles et al. (2001) developed a test for racial profiling in motor vehicle searches. In their model, policemen randomize over searching and not searching potential suspects. Palacios-Huerta (2003) and Chiappori et al. (2002) similarly studied penalty kick data in soccer to look for evidence of mixing behavior. To the best of our knowledge, our paper is the first to use this game-theoretic framework for empirical analysis in a political economy context.

Lastly, our paper contributes to the literature estimating discrete games of complete information. Most of these have been Industrial Organization applications focused on the problem of entry, and on pure strategy equilibria (see Berry (1992); Bresnahan and Reiss (1990, 1991)). In contrast, we estimate a model where only mixed strategies are economically meaningful, and propose a different identification strategy. Moreover, for games where a subset of outcomes is unobserved (such as the tax auditing game), Bresnahan and Reiss (1990) pointed out a negative identification result for the game's payoff parameters. Our methodology shows how this issue can be overcome empirically.

The rest of the paper proceeds as follows. Section 2 provides a brief overview of electoral campaigns in the U.S., focusing on Senate races. Section 3 presents our benchmark model of the campaign trail, Section 4 describes the data, and Section 5 discusses identification and our empirical strategy. Section 6 goes on to present our main results, and Section 7 concludes. Appendices A and B contain proofs and a detailed description of data sources.

2 Background

In this section we briefly discuss U.S. Senate races and provide some institutional background on them. The U.S. Senate has been democratically elected for a century now, after the 17th Amendment to the Constitution was passed in 1913. Before the amendment, State Legislatures elected U.S. senators. The Senate is composed of 2 senators per state; hence 100 senate seats currently exist. Senate elections are held every two years in November of even years, and senators are elected by plurality within each state. Under this system, a third of the seats are up for election on each 2-year cycle, and each seat has a six-year term. As a result, there are about 33 elections every electoral cycle[5].

As in most other elections for public office in the U.S., general elections are preceded by a period of campaigning, which in practice begins well before each party in each state has chosen its candidate in either a primary election or a convention. Most states hold primaries, which vary in how close to the general election they happen. Even during the primaries pollsters track hypothetical electoral outcomes for the general election. This is facilitated by the large fraction of Senate races including an incumbent senator, who is very likely to become his party's candidate in the general election, and often runs unopposed in the primary.

Technological change in the media industry has transformed in major ways how electoral politics operates in the U.S. Both the quantity and the kind of information received by voters changed as newspapers, television, and more recently the internet, arose and spread. Early on, direct contact between candidates and voters was reduced: printed news and television made the media an unavoidable middleman in the transmission of political messages. Direct contact between politicians and voters, for example through town-hall meetings and campaign-trail speeches, allowed candidates great control over the exact contents of their messages. Moreover, during the 19th and early 20th Centuries, the extent of direct con-

[5] After the resignation or death of an incumbent senator, special elections can be held at different times.

trol of media outlets by politicians' families also contributed to their ability to determine which constituencies were reached by different messages. In contrast, candidates now have little direct control over how the media will report on their actions and statements, both because of competition in the media market and the reduced extent of direct media control by politically involved families. Second, information has become increasingly public. Before the advent of these new information technologies, candidates had the ability to target their messages narrowly to specific groups. This ability has been significantly curtailed by the broad reach of modern media technologies. According to Ansolabehere et al. (1992, p. 71), "The importance of the mass media and the growth of television in particular have forced candidates to respond to the routines and incentives of news organizations. Candidates and their staffs devote a great deal of energy to influencing the decisions of reporters and editors. Successful candidates and campaigns also adjust their behavior to exploit the media environment in which they operate."

These changes manifest themselves in the key role that public and media relations play within the organizational structure of political campaigns. This is especially so in U.S. Senate races, which by their nature are quite salient and, as a result, are intensely covered by both state and national-level media outlets. Interestingly, the very recent emergence of social media may be allowing candidates to have more direct access to their constituencies once again. It may also partially allow increased message differentiation and targeting. In practice, technological change has altered both the costs of campaign coverage by the media, and the costs and benefits for candidates of producing differentiated messages.

This discussion also motivates our focus on U.S. Senate races. While the number of U.S. House races is significantly larger, House electoral districts are small relative to most media markets. This limits the extent to which the media will be directly following individual races. Furthermore, polling data for House races is scarce. In contrast, U.S. presidential races have extensive media and poll coverage, but there are too few of them for a satisfactory statistical analysis. Senate races are, thus, an ideal compromise. Moreover, their state-level nature implies that the electorate is diverse enough for candidates to have incentives to target different types of voters.

3 A Simple Model of the Campaign Trail

In this section we describe the simple model of campaign speech and media coverage that we subsequently estimate. The model captures what we consider key features of the interaction between two candidates running against each other, $p \in \{D, R\}$, and the distribution of media outlets m covering the race. In the model, candidates make statements over time that can be targeted to core or swing voter constituencies. The media decides on coverage of the campaigns every period, and obtain different payoffs from reporting on either type of campaign speech. The key assumption we maintain and implicitly test is that payoffs to the media are higher when reporting news on campaign speech targeted to core supporters. Candidates benefit electorally (in their poll standing) from media reports on their swing voter-targeted statements. Moreover, although they may also benefit from unreported core voter-targeted statements, they suffer from media reports of these kinds of statements as this leads swing-voters to shift towards their opponent.

Time is discrete, $t = 0, ..., T$, where $t = T$ is election day and $t = 0$ is the beginning of the campaign. Both candidates begin their campaigning on the same date. We also assume that each candidate makes a campaign statement every period. Each media outlet decides on whether to follow the Democratic candidate D, the Republican candidate R, or both. Conditional on following a candidate, the media successfully reports on their statements with an exogenous probability that may vary across parties. Candidate statements and media reports then map period by period onto changes in poll standings. Our focus on this paper is not on voters, so we model their behavior in a simple way; voters poll support decisions at any point in time respond to the amount of information they receive during the campaign, either directly from the candidates or through the media.

Players' Actions

Candidates make a statement every period to increase their electoral support. The underlying environment is such that candidates do not fully converge to the median voter's ideological stance. This can be easily micro-founded in a model where the turnout of voters in the extremes of the ideological distribution (core supporters) is sensitive to their distance to the candidates' position, and the density of voters is high in the extremes. Standard incentives to move towards the median would have to be traded-off against the loss in turnout from the margins of the distribution of voters. To capture the electoral support of core voters, candidates are tempted to make ideological statements c, directly targeted to this audience. Nevertheless, core-targeted statements may decrease the electoral support they receive from swing voters. Candidates may, instead, make relatively centrist, swing voter-targeted statements s, which generate little excitement in the extremes, but increases or maintains the electoral support in the center.

Candidates and the media have partially aligned preferences: candidates benefit from being reported by the media, and the media profits from reporting news about candidates. Nevertheless, their preferences are also partly misaligned: candidates benefit from the media reporting on their relatively centrist statements -those targeted to swing voters-, and are possibly hurt when the media reports on their ideological statements –those targeted at core voters. In contrast, the media profits more from reporting on core-targeted statements than from reporting on swing voter-targeted ones. Candidates also take each others' strategies as given when deciding their campaign-trail speech. This gives rise to a matching-pennies strategic environment between each candidate and the media. The reason is simple: the candidate's best response is to generate a swing voter-targeted statement when covered by the media, and to generate core-targeted speech when not covered by it. Similarly, the media's best response is to report core-targeted statements and to ignore statements targeted at swing voters.

Following these ideas, we model the candidates' action space as follows: each can take one of two actions every period; either to make a swing voter-targeted statement, or a riskier core voter-targeted statement: $a^p \in \{s, c\}$. Simultaneously, the media can take one of three possible actions: to follow both candidates, to follow only D, or to follow only R: $a^m \in \{(F_D F_R), (F_D N_R), (N_D F_R)\}$. In either case, after having taken its action, the media outlet successfully reports (denoted by $\chi = 1$) with probability $\mathbb{P}(\chi = 1|p) = \eta_p$. This modeling choice allows us to keep the action space of the media three dimensional, while still allowing for realizations of periods in which no news reports are observed. Moreover, it allows us to introduce a parameter, (η_D, η_R), directly capturing the media's overall propensity to report differentially about candidates from one or the other party. When $\eta_D \neq \eta_R$, we refer to this differential treatment as media bias.

Payoffs

The payoff structure is very simple. Every period, the media, m, must pay a cost k per candidate followed. The per-period gains from reporting on candidate p are:

$$\pi_p(a^p) = \begin{cases} 0 & if\ a^p = s \\ \pi_p & if\ a^p = c \end{cases} \tag{1}$$

where we have chosen to normalize the gain from reporting a swing voter-targeted statement to zero. Nevertheless, we allow the gain for the media to differ between a report about the Democrat or the Republican.

To simplify the payoff structure of the game, we make some behavioral assumptions about potential voters. The arrival of media reports can have two effects on voters' decisions. First, it can make them shift support from one candidate to the other. Second, it can alter their turnout decision. This distinction is important because the first margin leads to a zero-sum setting from the point of view of the candidates, while the second margin does not.

The payoff structure we present below implicitly assumes that core voters only react on the turnout margin, and never switch party allegiances. In contrast, centrist swing voters only react on the party support margin, and do not react on the turnout margin (their turnout rate is constant). Voters report truthfully to pollsters.

For candidates, instantaneous payoffs depend on whether their statements are reported or not, and whether these are targeted to swing voters or to core supporters. Candidates care about their poll standing, and players' actions directly map into changes in electoral and poll support. We suppose that unreported swing voter-targeted statements have no effect on either core or swing voters. When reported, these kinds of statements do have an effect on swing voters; they shift support from the candidate not reported to the candidate reported. Because the turnout rate for swing voters is unaffected, the gain for one candidate is exactly the loss for his opponent. We also suppose core voter-targeted statements increase the turnout of core constituencies. When they are unreported, these statements do not have an effect on swing voters. When they are reported, in contrast, they swing centrist voters away from the candidate making these statements and towards his opponent. For the stage game to have a matching pennies structure, this loss must be larger than the gain on the turnout margin. In such a case, the net effect on the polls from a reported core-targeted statement is negative for the candidate making it.

We denote by Δ^T_{ap} the average change in electoral support to candidate p on the turnout margin when he chooses action a^p, and by Δ^S_{ap} the average change in electoral support on the swing voter margin when candidate p chooses action a^p. We can express the change in poll support for each candidate $p \in \{D, R\}$ between periods t and $t+1$ as:

$$V_p(t+1) - V_p(t) = \Delta^T_{cp}\mathbf{1}\{a^p(t) = c, \chi(t) = 0\} + \left(\Delta^T_{cp} - \Delta^S_{cp}\right)\mathbf{1}\{a^p(t) = c, \chi(t) = 1\}$$

$$+\Delta^S_{c\sim p}\mathbf{1}\{a^{\sim p}(t) = c, \chi(t) = 1\} + \Delta^S_{sp}\mathbf{1}\{a^p(t) = s, \chi(t) = 1\} - \Delta^S_{s\sim p}\mathbf{1}\{a^{\sim p}(t) = c, \chi(t) = 1\} + \epsilon^p \quad (2)$$

Above, (ϵ^D, ϵ^R) are other unobserved shocks to the change in electoral support of Democrats and Republicans. We impose the following parameter restrictions:

Assumption 1. *The following inequalities hold:*

$$\begin{array}{llll} \Delta^T_{cD} < \eta_D\left(\Delta^S_{cD} + \Delta^S_{sD}\right), & \Delta^T_{cD} > 0, & \Delta^S_{cD} > 0, & \Delta^S_{sD} > 0 \\ \Delta^T_{cR} < \eta_R\left(\Delta^S_{cR} + \Delta^S_{sR}\right), & \Delta^T_{cR} > 0, & \Delta^S_{cR} > 0, & \Delta^S_{sR} > 0 \end{array}$$

These inequalities are sufficient for the stage game to have a unique Nash equilibrium in mixed strategies. Equations (1) and (2) and the parameter restrictions in Assumption 1 are fairly natural. They take into account the zero-sum nature of swing support, and also make explicit the assumptions that (i) unreported actions by a candidate do not have an effect on his opponent's support, and (ii) unreported s statements by a candidate do not have any effect on his own support. They also imply that candidates gain support from c statements that go unreported, but expect to lose support when these are reported. Finally, they imply that reported own swing voter-targeted statements increase own support (at the expense of the opponent), and reported opponent's swing-targeted statements decrease own support (and are a gain to the opponent). We further assume that candidates maximize their poll standing (which, in a bipartisan race, is equivalent to maximizing the winning probability at every t). In summary, we have a total of eleven structural parameters in this model: $\boldsymbol{\theta} = (\Delta^T_{cD}, \Delta^S_{cD}, \Delta^S_{sD}, \Delta^T_{cR}, \Delta^S_{cR}, \Delta^S_{sR}, \eta_D, \eta_R, \pi_D, \pi_R, k)$ characterizing the game. Given the uniqueness of equilibrium we establish below, these also pin down the joint distribution of players' actions and poll changes over time.

Indeed, the payoff structure of the game outlined above gives rise to a matching pennies stage game G. Its normal form representation is presented in Appendix A.

Proposition 1. (Equilibrium Strategies) *Suppose $\eta_p \pi_p > k$. The normal form game described above does not have a pure-strategy equilibrium. The unique mixed strategy equilibrium is given by:*

$$\gamma^*_R = 1 - \frac{\Delta^T_{cD}}{\eta_D \left[\Delta^S_{cD} + \Delta^S_{sD}\right]} \tag{3}$$

$$\gamma^*_D = 1 - \frac{\Delta^T_{cR}}{\eta_R \left[\Delta^S_{cR} + \Delta^S_{sR}\right]} \tag{4}$$

$$q^*_D = \frac{k}{\eta_D \pi_D} \tag{5}$$

$$q^*_R = \frac{k}{\eta_R \pi_R} \tag{6}$$

where γ_D denotes the probability that m plays $F_D N_R$, γ_R denotes the probability that m plays $N_D F_R$, q_D denotes the probability that D plays c, and q_R denotes the probability that R plays c. Furthermore, because the stage-game has a unique Nash equilibrium, the only sub-game perfect equilibrium of the finitely repeated game G^T is to play the unique stage-game Nash equilibrium every period.

Proof. See Appendix A. □

As is standard in a matching pennies environment, mixing probabilities are pinned down by indifference. This makes them a function only of the opponents' payoffs. In our application, this consequence of equilibrium has a subtle testable implication. If we want to study candidate campaign speech, we must do comparative statics on the media's payoffs. Candidates' payoffs are in fact irrelevant to explain their equilibrium behavior. Media payoffs from reporting on core-targeted statements will be negatively correlated with the equilibrium rate at which candidates make such statements. This is the sense in which, in our setting, the media can constrain candidates' behavior. Moreover, the poll gains from a given campaign statement should have no predictive power for the rate at which the candidate makes such statements. Conversely, the frequency with which the media reports on the candidates should be independent of how profitable it is to report. It should depend only on the candidates' payoffs. Notice also that the ratio $q_D^*/q_R^* = \eta_R\pi_R/\eta_D\pi_D$ does not depend on k.

4 Data

This section describes our data, the construction of our main variables, and the data structure. First, we discuss the information on all Senate races included in our dataset, the poll and election outcomes data, and how we rely on poll availability to construct the time dimension of our panel. We then introduce our news coverage data and describe the proposed methodology to compute news article scores measuring the type of content reported. We conclude by discussing the sports events data we will rely on for our identification strategy.

4.1 Senate Races

For our empirical implementation, we built a dataset of all ordinary competitive races to the U.S. Senate taking place between 1980 and 2012 for which a Democrat and a Republican ran.[6] Our final sample includes 415 races (out of the 567 that took place in this 32 year period). For each Senate race we have data on its outcome (Democratic share and Republican share) from the Federal Elections Commission, the date and outcomes of the primaries for each party whenever a primary took place –or whether the candidate was chosen at a party convention for states electing their candidates that way–, information on whether the incumbent senator was running, and characteristics of the political environment such as the party of the President, the party of the incumbent senators in the state, and the share of Democratic and Republican registered voters in the state. For states without party registration, we use the vote share for President in the most recent election. Table 1 presents summary statistics for the variables in our study.

[6] We excluded races with three prominent candidates, races where a candidate ran unopposed (or in practice unopposed), non bipartisan races, and races where either candidate died or quit during the campaign. Appendix B contains a list of dropped races.

Descriptive Statistics

Panel A	2-Week Intervals			3-Week Intervals		
	Dem.	All	Rep.	Dem.	All	Rep.
Number of poll-to-poll intervals per race		5.63 (4.55)			4.86 (3.62)	
Length of poll-to-poll interval (days)		30.51 (34.32)			35.16 (35.47)	
Number of polls per interval		1.74 (1.66)			2.01 (2.12)	
Electoral support (poll results)	0.44 (0.11)		0.42 (0.11)	0.44 (0.12)		0.42 (0.11)
Number of articles per interval	56.11 (101.7)	90.22 (127.42)	41.50 (62.49)	61.98 (112.53)	98.83 (138.0)	45.43 (67.63)
Number of core-targeted articles per interval (0.25 cutoff)	35.82 (91.09)	58.79 (102.6)	22.97 (45.92)	41.75 (101.52)	68.43 (113.89)	26.67 (52.28)
Number of swing-targeted articles per interval (0.25 cutoff)	20.29 (30.21)	38.82 (57.67)	18.53 (30.52)	20.23 (30.75)	38.98 (59.30)	18.76 (31.17)
Number of core-targeted articles per interval (0.5 cutoff)	29.05 (89.46)	45.8 (97.95)	16.75 (43.50)	34.07 (98.21)	54.12 (107.93)	20.05 (50.83)
Number of swing-targeted articles per interval (0.5 cutoff)	27.07 (40.23)	51.81 (77.23)	24.75 (40.81)	27.12 (42.40)	52.58 (82.64)	25.46 (43.78)
Number of core-targeted articles per interval (0.75 cutoff)	21.18 (63.21)	33.41 (69.67)	12.23 (31.81)	25.86 (77.21)	40.87 (84.21)	15.02 (38.08)
Number of swing-targeted articles per interval (0.75 cutoff)	33.92 (49.91)	63.01 (87.93)	29.09 (46.03)	35.33 (51.11)	65.82 (90.85)	30.49 (46.88)
Number of NFL games per interval (fan weighted)		4.22 (6.34)			4.91 (6.98)	
Number of MLB games per interval (fan weighted)		14.91 (25.43)			17.17 (27.12)	
Number of NBA games per interval (fan weighted)		8.91 (27.65)			10.15 (28.62)	
Number of NCAA games per interval (playoffs)		0.04 (0.29)			0.05 (0.31)	

Descriptive Statistics (cont.)

Panel B	Dem.	All	Rep.
Number of races		415	
Number of races per election cycle		24.41 (7.91)	
Number of polls		4076	
Number of polls per election cycle		239.76 (208.15)	
Number of polls per race		10.01 (11.93)	
Number of news articles	131131	210848	96984
Number of news articles per race	315.97 (488.13)	508.07 (687.04)	233.70 (358.47)
Article score	-0.52 (0.44)	-0.005 (0.70)	0.52 (0.43)
Number of media outlets per race		124.17 (85.40)	
Observations	2337		2033

Table 1: Descriptive Statistics: The table reports means and standard deviations for our main variables. Panel a reports summary statistics for the 2-week poll-to-poll interval panel and the 3-week poll-to-poll interval panel. Panel b reports overall summary statistics. Please see the text and the data description Appendix B for variable definitions and sources.

4.2 Polls

Our empirical strategy relies on information on the evolution of partisan support throughout the campaign. We collected detailed polling data for Senate races from a variety of sources. To the best of our knowledge, the earliest systematic compilation of polls goes back to 1998. We obtained polls from PollingReport.com for 1998-2004, and from Pollster.com for 2006-2012. For pre-1998 poll data, we did an exhaustive newspaper search using the Dow Jones/Factiva news database, focusing on all polling reported within a one-year window before election day.[7] We collected a total of 4076 polls. As Table 1 illustrates, we obtain

[7]For example, for the 1998 election we began our search on November 1, 1997. In a few cases we encountered discrepancies in the reported polling results across articles from different newspaper sources referring to the same poll, in which case we averaged the results. The 1998 poll data from PollingReport.com was sparse, so we also did an online newspaper search for polls for that year. When only the month of the poll was reported we imputed the date to be the fifteenth of the month except for November polls, in which

an average of 240 polls per election cycle, and of 10 polls per Senate race. Naturally, the frequency of Senate race polls becomes higher in more recent years and in states with larger populations.

Our empirical strategy also relies on our ability to compute frequencies of news reporting over time. To do this, we rely on our dataset of polls to construct what we call "poll-to-poll" intervals. By using the poll dates to create poll-to-poll time intervals within each race, we effectively create the time dimension of our panel. Subsequently, we use the dates of the news articles to assign them to their corresponding poll-to-poll interval, within which we measure the different news article statistics required by our empirical strategy, and which we describe in section 4.3. Because the definition of these periods is arbitrary, we explore two alternative criteria for the construction of the intervals, by grouping nearby polls using two-week or three-week windows, and averaging –weighting by poll sample size– all polls falling within the time window. We then assign the median date among the polls in the window as the period marker. This strategy is convenient because the frequency and spacing of polls is uneven across states and years, and because aggregating nearby polls helps us average out the inherent measurement error in poll reports. Figure 1 graphically illustrates the construction of the poll-to-poll intervals.

The construction of time periods in this way introduces an unavoidable precision/bias trade-off because the statistics we construct are based on observed news article relative frequencies. On the one hand, the longer a poll-to-poll interval, the smaller the sampling error in the measured statistics of news reports falling within it, and the closer these statistics will be to the probabilities with which they are generated. On the other hand, if the actual probabilities change significantly over time, –for example because payoff parameters depend on a time-varying state variable–, the longer a poll-to-poll period, the larger the bias from a statistic based on frequency counts within the interval. To deal with this issue, we explore the robustness of our results to alternative definitions of a poll-to-poll interval, and we perform a robustness exercise where we estimate a version of the model where the game is not repeated but dynamic. In that case, we allow the payoff parameters to evolve as a function of a state variable, namely the current relative poll standing of the candidates.

4.3 Measuring News Reporting

Estimation of our model requires us to classify different types of observed news article content related to the political campaigns. This allows us to establish a link between reported candidate speech and electoral performance. More specifically, we require a classification

case we imputed the date to be the first of the month.

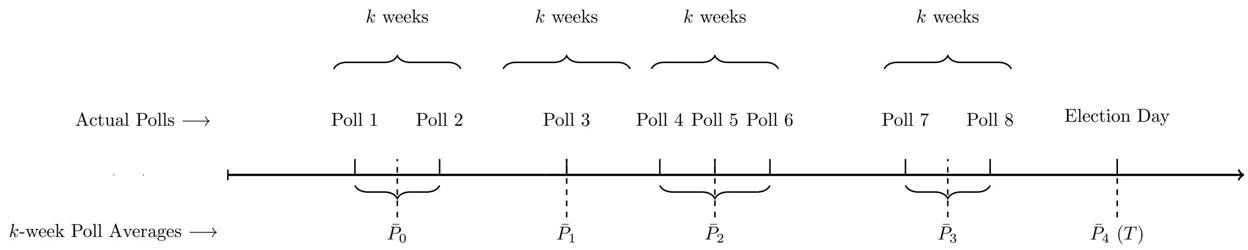


Figure 1: Construction of the poll-to-poll intervals.

criterion of each news piece as suggestive of turnout (core supporter)-targeted or swing voter-targeted campaign speech. Naturally, such a distinction is empirically meaningful only in relation to the ideological distribution of the relevant population of potential voters –the state in our setting. For example, the same statement may be considered moderate and targeted to centrist voters when expressed by a Democratic candidate in Massachusetts, but it may only appeal to core Democratic voters when expressed by a Democratic candidate in Utah. Moreover, the ideological distribution of the population within a state may change over time, making a statement that could be considered core-targeted in 1980, appealing to swing voters in 2012. A sensible classification criterion for the reported content of media reports must be race-specific.

With this in mind, we follow some of the ideas in the seminal work of Gentzkow and Shapiro (2010) for computing measures of media slant, to develop a novel index of media content. For each race, we conducted a comprehensive search of news reporting from two major news databases, *Lexis Nexis* and *Factiva*, which cover national and local newspapers. The search criteria was based on the names of both candidates in each race, during the year prior to election day.[8] We collected all news articles mentioning either candidate in a given race. Our initial search recovered more than 300,000 articles covering 560 races and 1120 candidates. For the set of articles mentioning either candidate in a given race, we implemented a text search algorithm to parse the HTML tags and gather information about the articles (publication date, source, subjects, and persons mentioned in the article). These tags allowed us to further weed out irrelevant articles and omit repeated articles. As Table 1 illustrates, our estimation sample contains information from 210,848 news articles, with an average of 508 articles per race. Although articles often mention both candidates, the average article is usually centered on reporting about one of them. The name of the opponent is reported as part of the context only. A few articles, of course, discuss the race as a whole and would be harder to classify as reporting about the Democrat or the Republican. We

[8]We downloaded these articles in HTML (for *Lexis Nexis*) and .rtf (for *Factiva*) formats

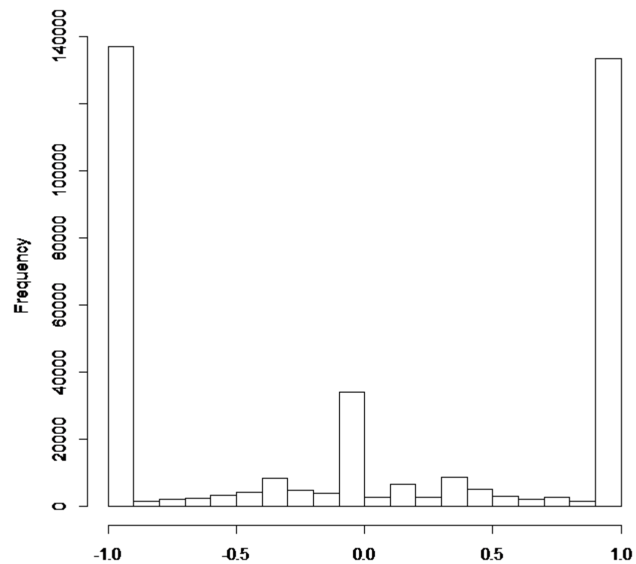


Figure 2: Distribution of article name assignments τ_i.

rely on the candidate name information in the articles themselves for the construction of our index of news reporting.

We proceed in the following way. First, to assess the extent to which an article reports on the Democratic or the Republican candidate, we count the number of times the name of each appears in the article. We then compute the candidate assignment statistic τ_i:

$$\tau_i = \frac{x_i^R - x_i^D}{x_i^R + x_i^D} \in [-1, 1]$$

where x_i^p is the count of party p's candidate name in article i. Values closer to $+1$ imply the article is more heavily reporting on the Republican, and values closer to -1 imply the article is more heavily reporting on the Democrat. Figure 2 presents the distribution of τ_i across all articles and races. The distribution is multi-modal, with most articles referring heavily to just one candidate. There is also some significant density of articles mentioning both candidates evenly (with scores close to 0). Table 1 reports the number of articles we classify as referring to the Democratic ($\tau_i < 0$) and Republican ($\tau_i > 0$) candidates. The τ_i provides us with a continuous measure that allows us to classify the contents of all articles in each race.

Within the set of all articles corresponding to a given race, we identify the 500 most commonly used 2 word phrases (2-grams), and the 500 most commonly used 3-word phrases (3-grams). We then proceed by giving a score $s_j \in [-1, 1]$ to each phrase $j \in \{1, 2, ..., 1000\}$, related to how Republican-specific vs. Democratic-specific the phrase is within the set of

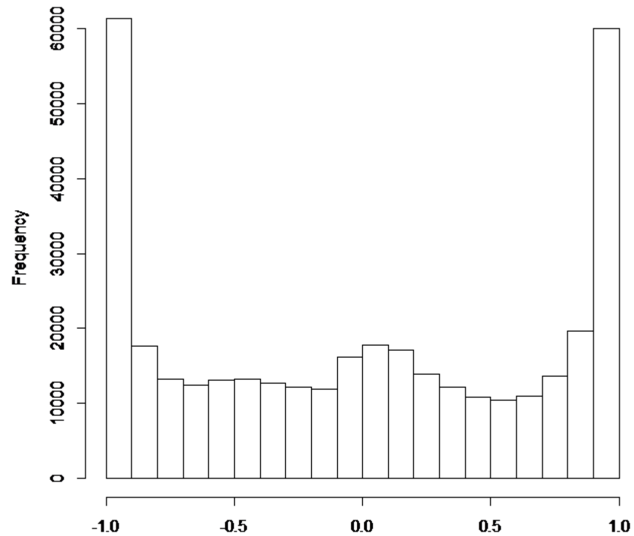


Figure 3: Distribution of articles scores σ_i.

articles covering the race. We do this by computing a weighted average of the τ_i's corresponding to articles containing phrase j, where the weights are the frequencies with which each phrase appears in each article, relative to all articles covering the race. Formally, for each j,

$$s_j = \frac{\sum_i \tau_i f_{ij}}{\sum_i f_{ij}} \in [-1, 1]$$

Here f_{ij} represents the frequency with which phrase j appears in article i. For example, if a given phrase appears only in articles that only mention the Republican candidate, then that phrase will have a score of $s_j = 1$. s_j gives gives us information regarding the extent to which phrase j is more commonly associated to one candidate or to the other. Endowed with the score s_j for each phrase in the race, we then compute a score for each news article in the race, building a weighted average of the scores of phrases appearing in the article, where the weights are the frequencies with which each phrase appears in each article, relative to all phrases in the article. Formally for each i,

$$\sigma_i = \frac{\sum_j s_j f_{ij}}{\sum_j f_{ij}} \in [-1, 1] \tag{7}$$

Articles with more phrases which, within the race coverage, are more closely associated with articles more heavily covering the Republican (Democratic) candidate will get higher (lower) scores. This measure has the advantage of being completely self-referential; we do not use any information from outside the coverage of the specific race to assess the extent to which a given

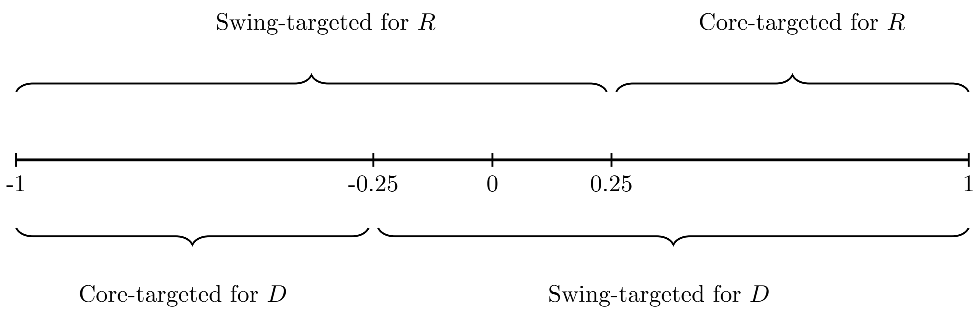


Figure 4: Illustration of the Article Type Classification (0.25 score cutoff case).

news piece is likely to be reporting about core supporter or swing voter-targeted statements by the candidates. σ_i is a continuous index which we use together with τ_i, to classify each article both as covering either the Democrat or the Republican, and whether the content is more swing voter targeted –s– or core supporter targeted –c– (depending on the value of σ_i). Figure 3 presents the distribution of the article scores σ_i for our sample of news pieces. Our benchmark specification classifies articles as signaling core supporter-targeted speech when $\sigma_i < -0.25$ for the Democrat and when $\sigma_i > 0.25$ for the Republican. It classifies the remaining articles as signaling swing voter-targeted speech (those with scores $\sigma_i \in [-0.25, 1]$ for the Democrat and with scores $\sigma_i \in [-1, 0.25]$ for the Republican). Figure 4 illustrates graphically the article classification criterion for the ± 0.25 cutoff. In our robustness analysis we present additional results that reclassify all articles using alternative cutoffs $\sigma_i = \pm 0.5$ and $\sigma_i = \pm 0.75$.[9]

Our collection of news articles also allowed us to obtain information on the number of different media outlets covering each race. We obtained this information based on the media outlet name and date tags in the articles. As a result, we have data on the count of different outlets reporting on a race within each poll-to-poll interval.[10] Finally, to compute overall reporting frequencies, we defined the total effective number of periods or stage games within each poll-to-poll interval as the number of days between polls times the total number of media outlets ever reporting on the particular race. This is equivalent to assuming that the candidates play a stage game against each media outlet every day during the campaign.

[9] Table 13 in Appendix B presents an illustrative sample of bigrams and trigrams from our text analysis with their corresponding scores.

[10] In order to remove misclassifications due to the occasional use of "the" in front of an outlet name (e.g., *The New York Times* could occasionally be classified as *New York Times*), we processed the text to remove the word "the" in front of all outlet names.

Conditional Correlations of Article Word Counts and Article Scores

Dependent variable:	Number of Words		Log Number of Words	
	(1)	(2)	(3)	(4)
Absolute value of article score ($\|\sigma_i\|$)	36.83	40.10	0.052	0.036
	(29.98)	(17.69)	(0.031)	(0.016)
Race Fixed Effects	N	Y	N	Y
R^2	0.02	0.06	0.09	0.29
No. of Races	417	417	417	417
No. of Observations	176034	176034	176034	176034

Table 2: Article Word Counts: The table presents OLS panel regressions of a measure of article word count on the absolute value of the article's score σ_i. The dependent variable in columns (1) and (2) is the word count of the article. The dependent variable in columns (3) and (4) is the log of the word count of the article. All models control for the article candidate assignment score τ_i, the article's days to election and its square, and include year fixed effects and a constant. Columns (2) and (4) include race fixed effects. Standard errors are robust to arbitrary heteroskedasticity and clustered at the Senate-race level.

4.3.1 A Cross-validation exercise

Our theoretical model of campaign-trail speech is based on the premise that the media profits relatively more from reporting on candidate speech targeted to core supporters. To the extent that this premise is correct, a revealed-preference argument would suggest that written media outlets should be willing to allocate more space to news pieces covering these kinds of campaign speech. As a cross-validation exercise of our index of media content σ_i, in Table 2 we look at the relationship between the number of words in an article in our sample, and the absolute value of its score σ_i. The table presents results from OLS specifications using either the number of words or its log, with and without race fixed effects. All specifications control for the article's candidate assignment score τ_i, a quadratic in the article's date, and year fixed effects. The conditional correlation between article length and σ_i is always positive, and is highly significant in the models including race fixed effects which exploit within-race variation only. The mean word count of articles in our sample is around 800 words. From column (2) in Table 2, moving from a score of 0 to a score of 1 increases the article's length by 40 words, or around 5% of the average article length. This suggests that our proposed index is a reliable signal of the article content relevant to our model.

4.4 Sports news data as media-payoff shifters

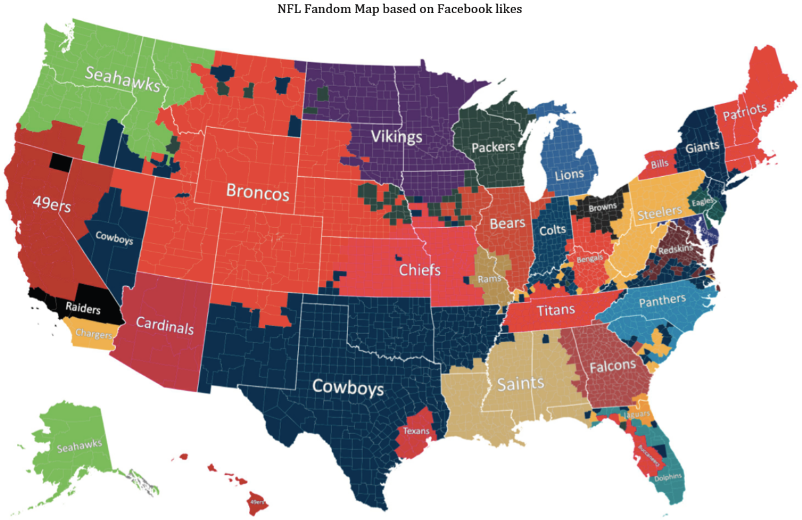
In our empirical strategy we exploit the observed correlations between frequencies of news reporting and changes in poll support for both candidates. Changes in electoral support along the campaign may be due to a host of unobservables (to the econometrician) which may, in turn, be correlated with candidates' incentives to make different kinds of statements and the media's incentives to cover campaigns. To overcome this difficulty we rely on the occurrence of major sports events as exogenous shifters of the media's attention, similar to Eisensee and Strömberg (2007). More specifically, we collected daily information on all games from the NFL, MLB, and NBA, and all playoffs games from the $NCAA$ between 1979 and 2012.[11] This constitutes a dataset with more than 600,000 observations. For each day we have information on whether a team played or not, and won or lost the game. We then match teams to their respective states, which gives us daily state-level variation in the media's payoff from reporting on political campaigns. This source of variation is unlikely to be related to unobservables driving candidate behavior along the campaign trail. Because most games for each sports league take place during a specific season of the year (e.g., football is concentrated in the winter, and baseball in the summer, for example), having information from the four leagues provides us with year-round variation.

Some states do not have teams in these leagues, or their teams seldom make it to the playoffs with enough frequency during the 33 year period. To obtain exogenous variation in media campaign coverage also for these states, we additionally collected information from *Facebook*. *Facebook* collected county-level information on the distribution of "likes" among its users in 2013, for each NFL, MLB, NBA, and $NCAA$ team. We use this information as a proxy for the extent to which the media covering a race in a given state may vary its behavior in response to salient sports events from teams of other states, which have a significant support in the state where the race is taking place. We computed the matrices $\mathbf{W}^{NFL}$, $\mathbf{W}^{MLB}$, and $\mathbf{W}^{NBA}$, where entry w_{ij}^{l}, $l \in \{NFL, MLB, NBA\}$ records the total population of counties in state i, as a fraction of total state population, where a plurality of *Facebook* users supports a team from state j in the sports league l. For states without teams in our data, these matrices provide us with variation in media payoffs, coming from a large fan base rooting for out-of-state sports teams that may lead to local media attention. Figure 5 illustrates the geographic distribution of fans of the teams in these four leagues, illustrating the straddling of fans across states which we exploit.[12]

[11] NFL is the National Football League, MLB is the Major League Baseball, NBA is the National Basketball Association, and $NCAA$ is the National Collegiate Athletic Association.

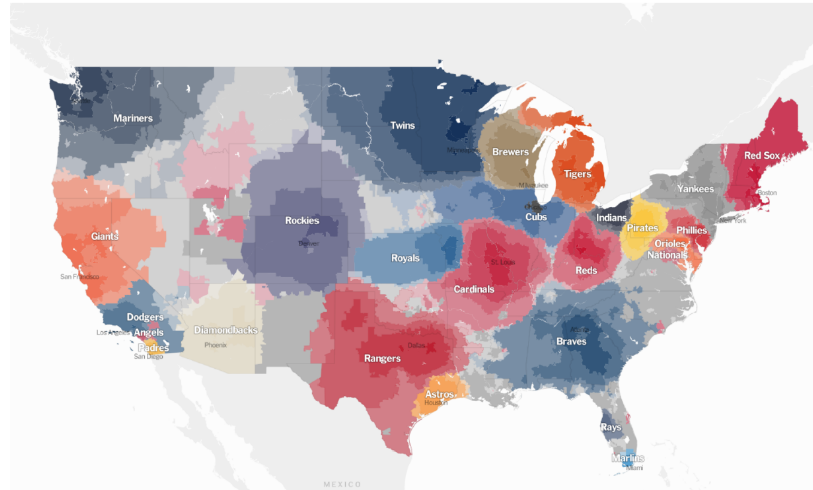
[12] The *Facebook* fan map for the $NCAA$ reveals that fandom for College Football is very highly correlated with state boundaries, thus giving us no additional variation. For this reason, we do not weight $NCAA$ sports events by the cross-state fandom weighs.

NFL Fandom Map based on Facebook likes



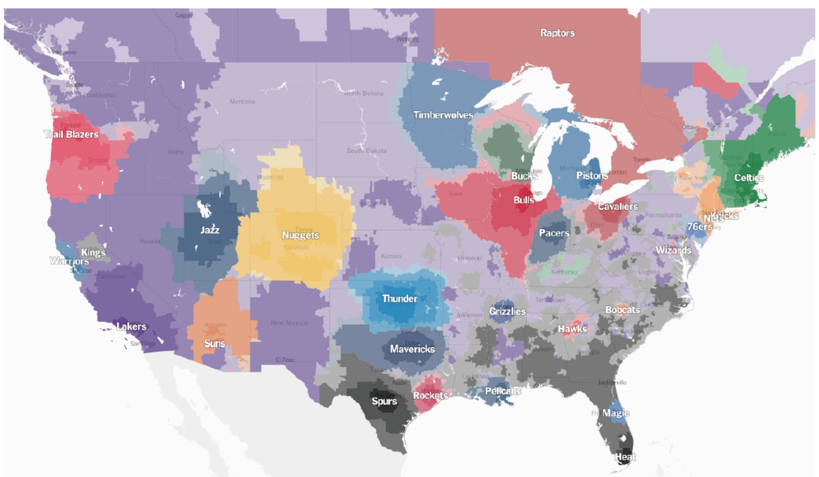
Source: The Atlantic online at: http://www.theatlantic.com/technology/archive/2014/09/the-geography-of-nfl-fandom/379729/

MLB Fandom Map based on Facebook likes



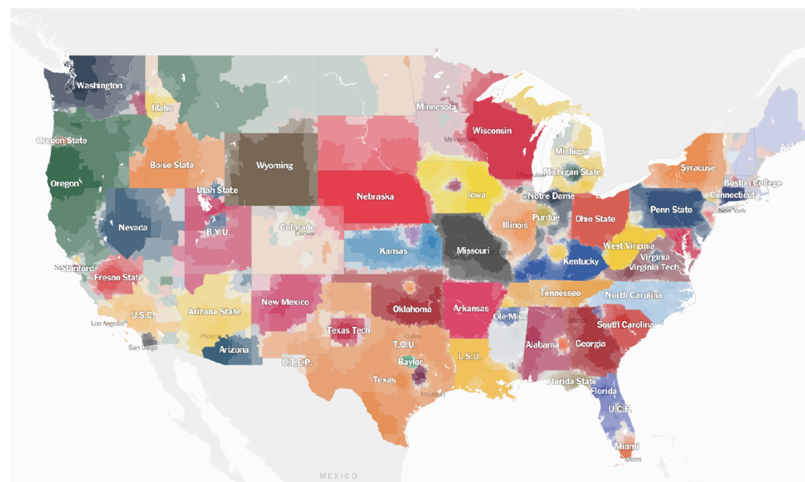
Source: New York Times online at: http://www.nytimes.com/interactive/2014/04/24/upshot/facebook-baseball-map.html

NBA Fandom Map based on Facebook likes



Source: The New York Times online at: http://www.nytimes.com/interactive/2014/05/12/upshot/12-upshot-nba-basketball.html?abt=0002&abg=1

NCAA Fandom Map based on Facebook likes



Source: The New York Times at: http://www.nytimes.com/interactive/2014/10/03/upshot/ncaa-football-fan-map.html?abt=0002&abg=1

Figure 5: *Facebook* Sport-Team Fans Distribution Maps.

5 Empirical Strategy and Identification

We now discuss our empirical strategy. Our purpose is to recover the payoff parameters governing the matching pennies game described above. The empirical strategy has several components. We first discuss the non-parametric identification of the equilibrium mixing strategies of all players based on the counts of the different types of news articles in our dataset. We then show how, relying on these mixing strategies, on polling data, and on the exogenous source of variation in campaign news coverage induced by the sports events, we can identify the electoral response elasticities that map directly on to the game's payoff parameters for the candidates. This is despite our inability to observe a subset of the equilibrium outcomes of the game, namely realizations in which the media does not report on the campaign.

5.1 Identification of the mixing strategies

Our first task is to develop a methodology allowing us to identify the model parameters $\boldsymbol{\theta}$ using the equilibrium conditions (3)-(6), and the data on media reports and poll and election results. The main difficulty in identifying payoff parameters from behavior reflecting mixed strategies in a setting such as this one is the nature of the game itself; we do not observe counts of type c or s candidate statements in periods when the media does not report on the campaign.

We first introduce some notation. Define $X_p^a(t, t+\tau)$ as the count of type $a \in \{c, s\}$ media reports on candidate $p \in \{D, R\}$ appearing between stage games t and $t+\tau$. Also define $N_p^a(t, t+\tau)$ as the count of type a campaign-trail statements by candidate p between stage games t and $t+\tau$ that do not get reported by the media. While the $X_p^a(t, t+\tau)$ are observed (with error), the $N_p^a(t, t+\tau)$ are unobserved. If within a time interval $[t, t+\tau)$ payoff parameters $\boldsymbol{\theta}$ are constant, the repeated matching pennies game directly gives us the joint distribution for the four observables and the four unobservables $(X_D^c, X_D^s, X_R^c, X_R^s, N_D^c, N_D^s, N_R^c, N_R^s)$. In fact, for each candidate, $(X_p^c(t, t+\tau), X_p^s(t, t+\tau), N_p^c(t, t+\tau), N_p^s(t, t+\tau))$ is a draw from a multinomial distribution with success and failure probabilities determined by the equilibrium mixing strategies of candidates and media outlets.

Correspondingly, each observed count X_p^a has a binomial marginal distribution given by

$$\mathbb{P}(X_p^a(t, t+\tau) = k) = \begin{pmatrix} \tau \\ k \end{pmatrix} \left(\varphi_p^a\right)^k \left(1 - \varphi_p^a\right)^{\tau - k} \tag{8}$$

where $\varphi_p^c \equiv q_p^*[1 - \gamma_{\sim p}^*]\eta_p$ and $\varphi_p^s \equiv (1 - q_p^*)[1 - \gamma_{\sim p}^*]\eta_p$. The expressions in equation (8) result from equilibrium play: q_p^* is the probability that candidate p chooses a core-targeted statement each period, and $[1-\gamma_{\sim p}^*]\eta_p$ is the unconditional probability that the media reports on candidate p each period.[13] Because the X_p^a are observable, equation (8) allows us to non-parametrically identify the conditional probabilities generating the observed media reports, by noticing that the MLE estimator for a binomial random variable is simply its sample analogue:

$$\hat{\varphi}_p^a(t, t+\tau) \equiv \frac{X_p^a(t, t+\tau)}{\tau} \tag{9}$$

The estimator in equation (9) gives us four equations in six unknowns $(q_D^*, q_R^*, \gamma_D^*, \gamma_R^*, \eta_D, \eta_R)$. Nevertheless, by taking the quotients of these conditional probabilities for each p, we can recover the equilibrium mixing strategies of both candidates:

$$\frac{\hat{\varphi}_p^s}{\hat{\varphi}_p^c} = \frac{(1-q_p^*)[1-\gamma_{\sim p}^*]\eta_p}{q_p^*[1-\gamma_{\sim p}^*]\eta_p} \Rightarrow \hat{q}_p^*(t, t+\tau) = \frac{\hat{\varphi}_p^s(t, t+\tau)}{\hat{\varphi}_p^s(t, t+\tau) + \hat{\varphi}_p^c(t, t+\tau)} \tag{10}$$

The efficient estimator of the $q_p^*(t, t+\tau)$'s will be the one taking $t = 0$, $\tau = T$. Moreover, from the definition of the φ_p^a's and equation (9) for $p \in \{D, R\}$, we can also recover the conditional probabilities of a news report of any type about candidate p:

$$\hat{\phi}_p(t, t+\tau) \equiv [1-\gamma_{\sim p}^*]\eta_p = \hat{\varphi}_p^c(t, t+\tau) + \hat{\varphi}_p^s(t, t+\tau) \tag{11}$$

Equation (11) also illustrates that in this model, and based on the observed article counts, we cannot separately identify the average media biases (η_D, η_R) from the equilibrium mixing strategies (γ_D^*, γ_R^*) of the media. The reason is that in stage games where a news report is not observed, we cannot distinguish whether this happened because the media did not invest in following the candidate, or because despite doing so, a news piece was not generated. We do not observe the realizations of the unobserved speech $(N_D^c, N_D^s, N_R^c, N_R^s)$. However, our ability to recover estimates of equilibrium reporting rates and of the equilibrium mixing strategies of candidates implies we have estimates of the parameters governing the

[13] We can additionally calculate the *conditional* distribution of the unobserved counts N_p^a using the information from the observed ones. This distribution is also binomial and takes the form:

$$\mathbb{P}(N_p^c(t, t+\tau) = k | X_p^c(t, t+\tau), X_p^s(t, t+\tau)) = \binom{\tau - X_p^c(t, t+\tau) - X_p^s(t, t+\tau)}{k} \times$$

$$\left[\frac{q_p^*[1-(1-\gamma_{\sim p}^*)\eta_p]}{(1-\gamma_{\sim p}^*)\eta_p}\right]^k \left[1 - \frac{q_p^*[1-(1-\gamma_{\sim p}^*)\eta_p]}{(1-\gamma_{\sim p}^*)\eta_p}\right]^{\tau - X_p^c(t, t+\tau) - X_p^s(t, t+\tau) - k}$$

This conditional distribution can be derived by noticing that $1 - (1-\gamma_{\sim p}^*)\eta_p$ is the total probability that the media does not report on candidate p each period. There is an analogous expression for the conditional distribution of $N_p^c(t, t+\tau)$ which we omit here to save space.

conditional distributions they are drawn from. Below we show this will be sufficient to recover the candidates' payoff elasticities $\mathbf{\Delta}$, arguably the parameters of most economic interest. Furthermore, as we will show below, exploiting the variation in polls and election results, we will also be able to provide an identified set for the η_p's, the γ_p's, and relative media payoffs π_R/π_D.

5.2 Identification of the electoral performance technology

We now discuss our identification strategy for the elasticities mapping equilibrium outcomes to changes in polls over time. Although a linear model for the changes in candidate shares of electoral support is necessarily misspecified, it makes our identification arguments transparent. It also illustrates clearly what the sources of variation we exploit are, to identify the candidates' payoff parameters. Recall the construction of our poll-to-poll intervals, and consider the realizations of all game outcomes within stage games t and $t+\tau$.[14] The proposition below establishes how the core-supporter turnout elasticities $(\Delta^T_{cD}, \Delta^T_{cR})$ can be recovered.

Proposition 2. (Identification of Turnout Elasticities) *Consider the following linear specification:*

$$\frac{V_D(t+\tau) - V_D(t) + V_R(t+\tau) - V_R(t)}{\tau} = \Delta^T_{cD}\frac{\hat{\varphi}^c_D(t,t+\tau)}{\hat{\phi}_D(t,t+\tau)} + \Delta^T_{cR}\frac{\hat{\varphi}^c_R(t,t+\tau)}{\hat{\phi}_R(t,t+\tau)} + \omega(t,t+\tau) \tag{12}$$

where $\hat{\varphi}^a_p(t,t+\tau)$ and $\hat{\phi}_p(t,t+\tau)$ are defined in equations (9) and (11), and $\omega(t,t+\tau)$ is an error term derived in Appendix A. Suppose instrumental variables $\mathbf{z}(t,t+\tau)$ are available, such that, (i) $\pi_p(\mathbf{z})$ vary with $\mathbf{z}$, (ii) $Cov(\mathbf{z},\omega)=0$, and (iii) the dimension of $\mathbf{z}$ is at least 2. Then a 2SLS regression of equation (12) using $\mathbf{z}$ as instruments for $\hat{\varphi}^c_D(t,t+\tau)/\hat{\phi}_D(t,t+\tau)$ and $\hat{\varphi}^c_R(t,t+\tau)/\hat{\phi}_R(t,t+\tau)$ identifies the elasticities $(\Delta^T_{cD}, \Delta^T_{cR})$.

Proof. See Appendix A. □

The dependent variable in equation (12) is the change between two time periods in the fraction of polled individuals reporting support for neither the Democratic nor the Republican candidate. In most polls, these are people who have not made up their mind about whether to turn out or about which candidate they prefer. Proposition 2 shows that the

[14] In practice, the length of a panel period, τ, will be determined by the frequency of polls for the race as we described in section 4.2. As long as pollsters' poll-timing decisions are not dependent on how the media is covering the campaigns or how the campaign is developing, choosing the panel periods this way will introduce no additional sources of bias when estimating equation A.5. In section 6.3.2 we test the plausibility of this assumption.

covariation in these changes with the ratios of news reports indicative of core-targeted statements to all news reports for each candidate can identify the average turnout response of the polled electorate to core voter-targeted statements. The result hinges on aggregating the changes in poll standings of both candidates; the zero-sum nature of swing voter-voter support implies that all swing voter effects cancel out after this aggregation.

The proof of this result proceeds by first writing down the net change in the poll standing of a given candidate between stage games t and $t+\tau$ from equation (2), as a function of the counts of all realizations of game actions that affect the evolution of electoral support $(X_p^c, X_{\sim p}^c, X_p^s, X_{\sim p}^s N_p^c)$ in that time interval. After adding the net changes for both candidates and dividing by the number of stage games considered, all terms related to swing voters cancel and we obtain an expression that depends only on $(N_p^c(t,t+\tau)+X_p^c(t,t+\tau))/\tau$ for $p \in \{D, R\}$ and the average of the shocks in the time interval (see equation (A.6) in Appendix A). A first empirical challenge is that this equation cannot be estimated directly; the counts of unreported core-targeted statements $(N_D^c(t,t+\tau), N_R^c(t,t+\tau))$ are unobserved. This difficulty can be overcome by noticing that $\mathbb{E}\left[(N_p^c(t,t+\tau)+X_p^c(t,t+\tau))/\tau\right] = q_p^*$ is simply the expected fraction of stage games in which candidate p will target his core supporters. We can then use our non-parametric estimate for the equilibrium mixed strategy of the candidate from equation (10), and express $(N_p^c(t,t+\tau)+X_p^c(t,t+\tau))/\tau$ as our estimate of q_p^*, namely $\hat{\varphi}_p^c(t,t+\tau)/\hat{\phi}_p(t,t+\tau)$, plus sampling noise that goes to zero at rate $\sqrt{\tau}$ as the poll-to-poll interval size increases.[15]

The second challenge in estimating equation (12) is the endogeneity of $\hat{\varphi}_D^c(t,t+\tau)/\hat{\phi}_D(t,t+\tau)$ and $\hat{\varphi}_R^c(t,t+\tau)/\hat{\phi}_R(t,t+\tau)$. Each of these is likely correlated with other unobservables that also determine the evolution of electoral support during a campaign, so we require at least two instrumental variables. These need to be sources of variation for the relative frequencies of core-targeted statements made by candidates, which do not, simultaneously, covary with any other determinants of the evolution of electoral support during the campaign. Our model suggests what the natural instruments for these variables should be. From the equilibrium mixing probabilities in equations (5) and (6), the mixing probabilities chosen by the candidates are pinned down by the media's payoffs from reporting:

$$q_p^*(\mathbf{z}) = \frac{k}{\eta_p \pi_p(\mathbf{z})} \tag{13}$$

A shifter of the media's payoffs to reporting on the campaign, which is otherwise unrelated to other campaign outcome determinants, will generate variation in the candidates' choices.

[15] Expressing the regressors in this way in equation (A.6) amounts in practice to including an explanatory variable with classical measurement error, which should create no additional issues as long as our instruments are valid.

As long as such an instrument varies across poll-to-poll intervals, it can be used to identify the parameters of interest from equation (12). Equation (13) also points out that the model makes an unambiguous prediction about the expected *sign* of the first stage; if larger values of the instrument reduce the media's profitability of reporting on politics, this should increase the rate at which the candidates target their core constituencies. Thus, we should expect a positive sign for the first stage. As we described in section 4.4, we rely on salient sports events which may crowd out the media's attention (thus, lowering its payoff from reporting on the campaigns). Eisensee and Strömberg (2007) use time variation generated by the occurrence of the Olympic Games to study media coverage of natural disasters. In a similar spirit, we use daily data on the occurrence of games in any of the four major sports leagues in the U.S. (MBL, NFL, NBA, $NCAA$). We match the games to the poll-to-poll intervals where they fall, including games with teams from the race's state or from other states with a significant local fan base as proxied by the *Facebook* fandom data (see section 4.4). The exclusion restriction is thus that games in any of these sports are uncorrelated to any unobserved determinants of the evolution of electoral support besides how they alter the media's relative payoffs from covering the campaigns. We believe this is a plausible exclusion restriction.[16] Moreover, because the model predicts the sign of the first stages, we consider the first stages as implicit specification tests of our model.

Equipped with estimates of Δ^T_{cD} and Δ^T_{cR}, we now describe the identification of the remaining electoral support elasticities.

Proposition 3. (Identification of Swing Elasticities) *Consider the following linear specification:*

$$\frac{\hat{D}(t,t+\tau)-\hat{R}(t,t+\tau)}{2} = \Delta^S_{sD}\hat{\phi}_D(t,t+\tau)\tau - \Delta^S_{sR}\hat{\phi}_R(t,t+\tau)\tau + \zeta(t,t+\tau) \tag{14}$$

where

$$\hat{p}(t,t+\tau) \equiv [V_p(t+\tau)-V_p(t)] - \hat{\Delta}^T_{c\sim p}\frac{\hat{\varphi}^c_{\sim p}(t,t+\tau)}{\hat{\phi}_{\sim p}(t,t+\tau)}\tau$$

is the change in electoral support for candidate p net of the turnout effects of the opposing candidate, and $\zeta(t,t+\tau)$ is an error term derived in Appendix A. Suppose instrumental variables $\mathbf{z}(t,t+\tau)$ are available, such that, (i) $\tau(\mathbf{z})$ varies with $\mathbf{z}$, (ii) $Cov(\mathbf{z},\zeta)=0$, and (iii) the dimension of $\mathbf{z}$ is as least 2. Then an IV regression of equation (14) using $\mathbf{z}$ as

[16] One possible channel through which the exclusion restriction may fail is if the occurrence of these sports events directly either lowers the turnout or changes the voters' electoral choices. Healy et al. (2010), for example, find that college football wins around election day increase the vote share of incumbent Senators. This effect is restricted to matter only around election day, thus only for the last poll-to-poll interval in each race. As robustness checks we estimate the model excluding the last period of each race, and using only variation in games won instead of variation in games taking place.

instruments for $\hat{\phi}_D(t,t+\tau)\tau$ *and* $\hat{\phi}_R(t,t+\tau)\tau$ *identifies the elasticities* $(\Delta^S_{sD},\Delta^S_{sR})$*. Finally,* $(\Delta^S_{cD},\Delta^S_{cR})$ *are identified from*

$$\Delta^S_{cp} = \frac{\hat{\Delta}^T_{cp}}{\hat{\phi}_p} - \hat{\Delta}^S_{sp} \tag{15}$$

Proof. See Appendix A. □

Proposition 3 shows that the covariation between appropriately "corrected" changes in the competitiveness of the race over time and the rate at which candidates are being reported about any type of statements can identify the average swing voter response of the polled electorate to swing voter-targeted statements. The correction term, $\hat{\Delta}^T_{c\sim p}\frac{\hat{\varphi}^c_{\sim p}(t,t+\tau)}{\hat{\phi}_{\sim p}(t,t+\tau)}\tau$ captures poll changes induced by core-targeted statements from the opponent. Although these do not directly affect the poll standing of the candidate (see equation (2)), in equilibrium they do so indirectly. As equation (15) shows, the poll gains to a candidate from core supporter-targeted statements by his opponent, Δ^S_{cp}, are fully pinned down by equilibrium play and the remaining elasticities. The proof of Proposition 3 shows that relying on this equilibrium relationship we can use the estimates of $(\Delta^T_{cD},\Delta^T_{cR})$ from Proposition 2 to appropriately correct for poll changes stemming from core-targeted candidate speech.

Subsequently, by scaling the net poll change in the interval by its size (the corresponding number of stage games), we obtain an expression where the regressors depend on τ but the error term does not. This is key to our identification strategy. The reason is that in equation (14) both $\hat{\phi}_p(t,t+\tau)\tau$'s will be correlated with other unobservables in $\zeta(t,t+\tau)$ driving the evolution of the campaign. We require the use of appropriate instruments once more, and we again rely on exogenous variation induced by sports events. The variation in this case is of a different nature, however. In contrast to the identification idea for the elasticities in Proposition 2, where we leveraged the dependence of q^*_p on the media's payoff from reporting on politics π_p, our model implies that the equilibrium reporting rate $\mathbb{E}\left[\hat{\phi}_p(t,t+\tau)\right] = \eta_p(1-\gamma_{\sim p})$ is independent of π_p. As equations (3) and (4) show, in equilibrium the reporting rate depends only on the candidates' payoff parameters, which are unlikely to respond to variation in sports events.

Nevertheless, the right-hand side variables in equation (14) are scaled by the size of the poll-to-poll interval. If the occurrence of sports events leads to variation across media outlets in their willingness to report on politics, then sports events can be valid shifters of $\tau(\mathbf{z})$, and thus valid instruments.[17] Sports events induce no intensive-margin response by a given media outlet (whose reporting strategy is pinned down by indifference). They can, nevertheless, induce an extensive margin response across the distribution of media outlets covering a race.

[17] Recall that in equation (14), $\zeta(t,t+\tau)$ is independent of τ.

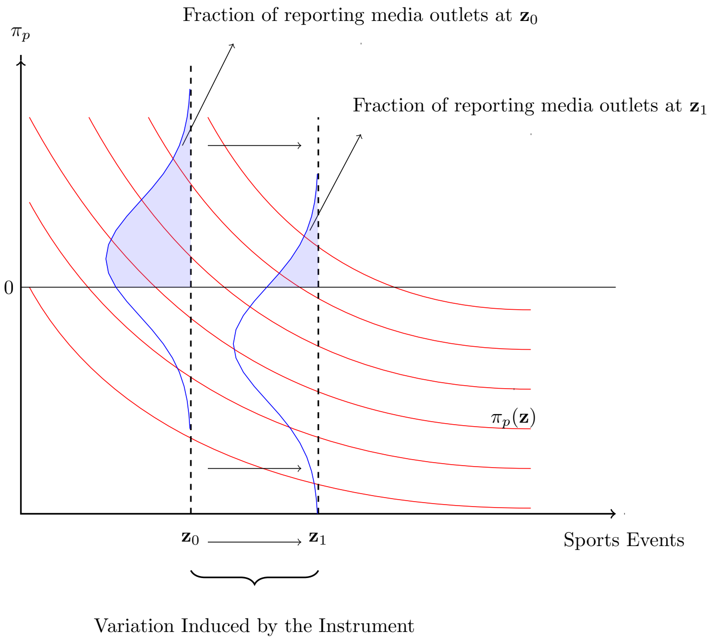


Figure 6: Instrument variation and media coverage in the extensive margin.

Figure 6 illustrates how variation in sports events can lead to infra-marginal outlets (those barely not covering the race in a given period) to enter coverage, or to supra-marginal outlets (those barely covering the race in a given period) to drop out from coverage. In the figure we plot a hypothetical distribution of media outlets who are heterogeneous in their payoffs from covering the campaign. Across the board, their payoff from campaign coverage is decreasing in the occurrence of relevant sports events, and only those outlets with a positive payoff invest in covering the campaign. When more sports events take place in a given period, some media outlets stop covering the campaign, effectively lowering the number of stage games τ being played. We exploit this source of variation to instrument for the endogenous variables in equation (14). As this discussion points out, our model once again makes an unambiguous prediction about the expected sign of the first stages for equation (14). In this case, the model predicts a *negative* relationship between the intensity of sports events and the two endogenous variables in the structural equation. The model also predicts a *positive* estimate for the IV second-stage coefficient on $\hat{\phi}_D(t, t+\tau)\tau$, and a *negative* estimate for the coefficient on $\hat{\phi}_R(t, t+\tau)\tau$. These sign predictions are further specification tests of our model.

In Table 3 we directly test this mechanism in our data, by looking at the correlation between sports events and the number of distinct media outlets from which we observe news pieces over time. We find evidence that the number of media outlets covering a senate

Testing Model Assuptions: Media Coverage and Sports Events on the Extensive Margin

Dependent variable:	No. of reporting media outlets in poll-to-poll interval / Total no. of reporting media outlets in the race									
	2 week poll-to-poll intervals					3 week poll-to-poll intervals				
Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log NFL games/τ	−0.008					−0.005				
	(0.029)					(0.036)				
Log MLB games/τ		−0.021					−0.042			
		(0.019)					(0.021)			
Log NBA games/τ			−0.035					−0.041		
			(0.018)					(0.021)		
Log NCAA games/τ				−2.335					−2.046	
				(0.609)					(0.658)	
Log all games/τ					−0.062					−0.085
					(0.017)					(0.020)
R^2	0.69	0.69	0.69	0.70	0.70	0.70	0.70	0.70	0.70	0.70
No. of Races	415	415	415	415	415	415	415	415	415	415
No. of Observations	2134	2134	2134	2134	2134	1865	1865	1865	1865	1865

Table 3: Testing Model Assuptions: Media Coverage and Sports Events on the Extensive Margin. The table presents OLS panel regressions. The dependent variable in all columns is the number of media outlets reporting on a race in a poll-to-poll interval as a fraction of all media outlets ever reporting on the race. All models include a full set of Senate-race fixed effects, month fixed effects, a dummy variable for the last poll-to-poll interval in the race, and a constant. The first five columns of the table are estimated on the 2 week poll-to-poll interval panel. The last five columns of the table are estimated on the 3-week poll-to-poll interval panel. Columns (1) and (6) include the log number of NFL games per day, columns (2) and (7) include the log number of MLB games per day, columns (3) and (8) include the log number of NBA games per day columns (4) and (9) include the log number of $NCAA$ games per day, and columns (5) and (10) include the log number of NFL, MLB, NBA, and $NCAA$ games per day. All regressions are weighted by the square root of the length in days of the poll-to-poll interval (relative to the longest interval). Standard errors are robust to arbitrary heteroskedasticity.

race does vary systematically with sports events relevant to the race's state. We test this assumption relying on our dataset of news articles, which includes the media outlet names for each news piece. This allows us to compute the number of different outlets producing news for a given race over time. The table reports OLS results of a regression where the dependent variable is the number of distinct media outlets reporting on a senate race in a given poll-to-poll interval as a fraction of all media outlets ever reporting on that race, on each of our sports events instruments. These models include Senate-race fixed effects, exploiting only within-race variation. The table presents results both for the 2-week and 3-week poll-to-poll interval datasets we described in section 4. All regressions show evidence of a significant and negative within-race correlation between game frequencies and media outlet coverage on the extensive margin.

6 Estimation Results

In this section we present our main empirical findings and discuss several robustness exercises. We first present our point estimates for the poll change elasticities $\mathbf{\Delta}$, and conclude with a partial identification exercise for the media biases (η_D, η_R). Overall, we find that the turnout margin response to core-targeted speech is more responsive for Democratic candidates and that the swing voter response to swing-targeted speech is quantitatively similar for Democrats and Republicans. We also find that the responsiveness of swing voters to swing-targeted speech is lower in states with a more uneven partisan distribution of voters. Nevertheless, we do not find evidence suggesting that voter responsiveness to media coverage significantly changes as the campaigns develop. Our results also indicate that Democratic candidates suffer larger swing voter losses when their core-targeted campaign speech is widely reported by the media compared to those of Republican candidates. Finally, we find that a large bias of the average media outlet towards either party is unlikely.

6.1 Poll Change Elasticities

Our empirical strategy consists of several steps. On our state-x-race-x-poll-to-poll interval dataset, we compute the average counts of reported core-targeted statements for each candidate within a poll-to-poll interval $\hat{\varphi}^c_{p,r,t} \equiv \frac{X^c_{p,r,t}}{\tau_{r,t}}$, and the average counts of total reported statements for each candidate within a poll-to-poll interval $\hat{\phi}_{p,r,t} \equiv \frac{X^c_{p,r,t}}{\tau_{r,t}} + \frac{X^s_{p,r,t}}{\tau_{r,t}}$. $p \in \{D, R\}$ denotes the candidate's party, r denotes the race, $t \in \{1, 2, ...T_r\}$ denotes the poll-to-poll interval, T_r is the last poll-to-poll interval of race r, and $\tau_{r,t}$ denotes the number of stage games within poll-to-poll interval t for race r (computed as the days in the poll-to-poll interval times the number of total media outlets ever reporting on the race).

Following Proposition 2, we estimate the turnout effects of core-targeted statements by 2SLS using the following model:

$$\frac{\triangle_t v^D_r + \triangle_t v^R_r}{\tau_{r,t}} = \Delta^T_{cD} \frac{X^c_{D,r,t}}{X^c_{D,r,t} + X^s_{D,r,t}} + \Delta^T_{cR} \frac{X^c_{R,r,t}}{X^c_{R,r,t} + X^s_{R,r,t}} + \delta_r + \sum_{m=1}^{12} \varrho^m_{r,t} + \omega_{r,t} \qquad (16)$$

Here the δ_r are race fixed effects. These will capture any unobservable systematic differences that are constant within a state or within an election year, such as the state's average ideology, or any specific features of a given electoral year such as the party in power, or whether it is a midterm election year. As such, we exploit exclusively within-race variation in media reporting and electoral support changes along the campaign trail. The $\varrho^m_{r,t}$ are month-of-the-year fixed effects. These are important in this setting because the sports events we use as instruments are highly seasonal. As a robustness exercise, we also estimate equation (16) above with a full set of state, year, and state-x-year fixed effects instead of race fixed effects.

Estimation of equation (16) requires instruments that vary across poll-to-poll intervals within a race, for both right-hand side regressors. As mentioned above, we rely on the occurrence of major sports events. More specifically, we compute our instruments $z^l_{r,t}$ as the fan-weighted log number of games per day from sports league $l \in \{NFL, MLB, NBA, NCAA\}$ relevant to state r falling within the poll-to-poll interval t:

$$z^l_{r,t} = \log\left[\frac{1}{\tau_{r,t}}\sum_j w^l_{rj} l_{r,t}\right]$$

where the w^l_{rj} are the fraction of state r's population in counties where a plurality of *Facebook* users are fans of a team from state j playing in sports league l. We do not use the *Facebook* fan weights for $NCAA$ games (see section 4.4). This amounts to making the $w^{NCAA}_{rj} = 0$ if $r \neq j$, and $w^{NCAA}_{rr} = 1$.[18]

Table 4 presents our main estimates of equation (16) together with the coefficients on our four instruments in each of the two first stages. Recall that our model predicts that the occurrence of sports events, by lowering the profitability of campaign reporting, should lead to an *increase* in core-targeted reported statements relative to total reported statements. Reassuringly, there is a systematically positive first-stage relationship between our instruments and each endogenous right-hand side variable in the main equation.[19] The first stage diagnostic statistics reveal that sports events are jointly good predictors of the fraction of core-targeted to total news articles on a candidate.

Table 4 reports estimates based on the 2-week poll-to-poll interval dataset in the first four columns, and on the 3-week poll-to-poll interval dataset in the last four columns. In both cases we report results using the ± 0.25 article score cutoff classification described in section 4.3. We also present estimates from OLS models which illustrate the importance of appropriately controlling for the endogeneity of equilibrium news coverage. Columns (1), (2), (5) and (6) present results that include race fixed effects, while columns (3), (4), (7) and (8) present results that include the state, year, and state-x-year fixed effects instead. In practice, results are unchanged when using either set of fixed effects. The standard errors we present throughout allow for heteroskedasticity and serial autocorrelation of up to order two, which

[18] As additional robustness exercises available upon request, we estimated our main equations using the number of winning games per day as instruments instead. Results are very similar.

[19] The partial correlation coefficient for $NCAA$ games on the first stage for $\frac{X^c_{R,r,t}}{X^c_{R,r,t}+X^s_{R,r,t}}$ is negative. Nevertheless, the unconditional correlation (without controlling for the remaining sports) is positive.

Electoral Share Response Elasticities (0.25 score cutoff)

Panel A: Structural equation		Dependent variable: $(\Delta V_D + \Delta V_R)/\tau$							
		2 week poll-to-poll intervals				3 week poll-to-poll intervals			
Explanatory variable	Param.	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)	OLS (7)	2SLS (8)
$X^c_D/(X^c_D + X^s_D)$	Δ^T_{cD}	0.024	0.16	0.026	0.16	0.032	0.15	0.032	0.18
		(0.006)	(0.06)	(0.006)	(0.06)	(0.006)	(0.08)	(0.07)	(0.089)
$X^c_R/(X^c_R + X^s_R)$	Δ^T_{cR}	-0.003	0.05	-0.003	0.05	-0.005	0.09	-0.004	0.09
		(0.005)	(0.058)	(0.005)	(0.059)	(0.005)	(0.08)	(0.005)	(0.088)
Panel B: First Stages		Dependent variable: $X^c_D/(X^c_D + X^s_D)$							
Log NFL games/τ			0.076		0.077		0.104		0.119
			(0.034)		(0.035)		(0.041)		(0.041)
Log MLB games/τ			0.049		0.048		0.034		0.032
			(0.024)		(0.025)		(0.026)		(0.027)
Log NBA games/τ			0.060		0.060		0.058		0.058
			(0.020)		(0.020)		(0.022)		(0.022)
Log NCAA games/τ			1.112		1.135		0.695		0.722
			(0.588)		(0.590)		(0.674)		(0.685)
R^2			0.95		0.95		0.96		0.95
F test (p-value)			0.001		0.001		0.006		0.003
		Dependent variable: $X^c_R/(X^c_R + X^s_R)$							
Log NFL games/τ			0.088		0.089		0.033		0.014
			(0.034)		(0.034)		(0.041)		(0.042)
Log MLB games/τ			0.087		0.086		0.048		0.050
			(0.024)		(0.024)		(0.026)		(0.027)
Log NBA games/τ			0.023		0.023		-0.010		-0.012
			(0.020)		(0.020)		(0.022)		(0.022)
Log NCAA games/τ			-1.234		-1.234		-1.339		-1.352
			(0.579)		(0.579)		(0.673)		(0.688)
R^2			0.93		0.93		0.94		0.94
F test (p-value)			0.000		0.000		0.05		0.042
Race fixed effects		Y	Y	N	N	Y	Y	N	N
Year×State fixed effects		N	N	Y	Y	N	N	Y	Y
No. of Races		415	415	415	415	415	415	415	415
No. of Observations		2134	2134	2134	2134	1865	1865	1865	1865

Table 4: Turnout Support Gain Elasticities (0.25 score cutoff). The table presents OLS and 2SLS estimates of the turnout gain elasticities from equation (16) using a 0.25 article score cutoff. Even-numbered columns present OLS estimates and odd-number columns present 2SLS estimates. Panel A present estimates for the structural equation (second stage), and panel B presents estimates of the coefficients for the instruments in both the first stages for the Democratic and the Republican ratios of turnout-targeted to total news reports. The first four columns in the table are estimated on the 2-week poll-to-poll interval panel. The last four columns are estimated on the 3-week poll-to-poll interval panel. All regressions are weighted by the square root of the length in days of the poll-to-poll interval (relative to the longest interval). Columns (1), (2), (5), and (6) include Senate-race fixed-effects. Columns (3), (4), (7), and (8) include a full set of year, state, and year-x-state fixed effects. All models include a dummy variable for the last poll-to-poll interval in a race and month fixed effects. Standard errors are robust to arbitrary heteroskedasticity and to arbitrary serial correlation of up to order 2 following Newey and West (1987). Coefficients and standard errors in Panel A are multiplied by 1000.

we believe is important given the nature of our data. Because we measure the end-of-period electoral support for the last poll-to-poll interval of each race using the election outcome instead of a poll, we additionally include a dummy variable for the last poll-to-poll interval in each race. All of our estimated regressions are also weighted by the square root of the length in days of the poll-to-poll interval because longer intervals contain more information than shorter ones and there is significant variation in poll-to-poll interval sizes in our data.

All of our IV estimates for the Democratic turnout elasticity Δ_{cD}^{T} are positive and significant. Although the IV estimates for the Republican turnout elasticity Δ_{cR}^{T} are systematically positive across all of our specifications and robustness exercises, they are significantly smaller than the Democratic turnout elasticity, and their standard errors are large. This is not too surprising given the large amount of measurement error in our dependent variable, which relies on arguably quite noisy polls. Our estimates are also very similar when using the 2-week and the 3-week poll-to-poll intervals. We believe this first result is important, as it points out that Republican core supporters are much less responsive on the turnout margin to campaigning targeted towards them than Democratic core supporters. This may be because solidly Republican constituencies exhibit high baseline turnout rates. It is well known, for example, that senior white males in rural areas, who tend to favor the Republican party, turn out at much higher rates than other demographic groups. As a result, Republican candidates' incentives to target those sectors of the electorate may be weaker. Democratic campaigns, in contrast, often appear focused on mobilizing turnout among younger and minority demographic groups, possibly because these groups have lower average turnout rates, making the potential gains on this margin large.

Our estimates from Table 4 are informative about the partial equilibrium quantitative effects of candidate behavior on poll changes. In the bottom panel of Table 6 we report the average estimates of candidates' mixing strategies q_p, as measured by the ratio of core-targeted news reports to all news reports. Based on the 0.25 article score cutoff criterion, $\mathbb{E}[q_D] \approx 0.56$, and $\mathbb{E}[q_R] \approx 0.45$. A ten percent increase in these probabilities, which is within the range of variation induced by the sports events, if sustained during a month would translate, on average, into a 3.3 percentage point gain to the Democratic candidate, and a 0.8 percentage point gain to the Republican candidate stemming from their core supporters' increased turnout.[20] Because the margin of victory for most Senate races is around 5 percentage points, this simple exercise illustrates the importance of media coverage incentives on election outcomes.

With our estimates for $(\Delta_{cD}^{T}, \Delta_{cR}^{T})$ at hand, we then construct $\hat{D}_{r,t}$ and $\hat{R}_{r,t}$ as defined in

[20] $(0.1 \times 0.56) \times (0.16/1000) \times 124$ media outlets on average $\times$ 30 days ≈ 0.033 for Democrats, and $(0.1 \times 0.45) \times (0.05/1000) \times 124$ media outlets on average $\times$ 30 days ≈ 0.008 for Republicans.

Proposition 3. We then estimate the swing voter effects of swing voter-targeted candidate speech by 2SLS using the specification:

$$\frac{\hat{D}_{r,t} - \hat{R}_{r,t}}{2}\tau_{r,t} = \Delta^S_{sD}[X^c_{D,r,t} + X^s_{D,r,t}] - \Delta^S_{sR}[X^c_{R,r,t} + X^s_{R,r,t}] + \tilde{\delta}_r + \sum_{m=1}^{12} \tilde{\varrho}^m_{r,t} + \zeta_{r,t} \tag{17}$$

Once again, the $\tilde{\delta}_r$ are race fixed effects, and $\varrho^m_{r,t}$ are month-of-the-year fixed effects. Recall from Proposition 3, that through the lens of our model, the regressors in equation (17) are the sample analogues of $\tau\eta_p(1-\gamma_{\sim p})$. Total observed news reports on a candidate should not vary as a function of changes in the media's payoff –from equations (3)-(4), these conditional probabilities only depend on candidates' payoffs–. Our model suggests that a media outlet's reporting strategy is pinned down by indifference, and thus, is independent of its own payoff. Nevertheless, for the distribution of media outlets as a whole, a shift in the profitability of reporting on political campaigns can lead to an extensive margin response by outlets entering into or dropping out from coverage (see Figure 6). Following this idea, we use sports events as exogenous sources of variation for the two endogenous regressors in equation (17). In this case, the model predicts sports events should be *negatively* correlated with $[X^c_{p,r,t} + X^s_{p,r,t}]$. This is exactly the pattern we find in the first stage estimates, which we present in Table 5.

Table 5 presents our benchmark estimates of equation (17). The table has the same structure as that of Table 4. Its first four columns are based on the 2-week poll-to-poll interval dataset, and the last four are based on the 3-week poll-to-poll interval dataset. All models in the table are also based on the ± 0.25 article score cutoff classification. As discussed above, the first stage estimates in panel B show that our instruments are systematically negatively correlated with both the Democratic and the Republican total news reports counts. Panel A then presents our main estimates of the Democratic and Republican swing-voter elasticities in response to swing voter-targeted media contents. Quite reassuringly, across all models estimated by 2SLS we obtain a positive coefficient on $[X^c_{D,r,t} + X^s_{D,r,t}]$ corresponding to Δ^S_{sD}, and a negative coefficient on $[X^c_{R,r,t} + X^s_{R,r,t}]$ corresponding to $-\Delta^S_{sR}$, exactly as implied by equation (17). We consider this pattern of resulting signs to very strongly suggest the validity of our proposed model. The IV estimates in the table show that the swing voter electoral support elasticities are remarkably similar in magnitude for Democratic and Republican candidates. Column (4), for example, shows our estimates for both parameters to be 0.0018, although the one for the Democratic candidate is more precisely estimated. Both Δ^S_{sD} and Δ^S_{sR} are significant at the 5% level. Across specifications both the magnitudes and significance of the parameter estimates are very similar.

We can undertake a quantitative exercise based on our benchmark estimates of $(\Delta^S_{sD}, \Delta^S_{sR})$

Electoral Share Response Elasticities (0.25 score cutoff)

Panel A: Structural equation		Dependent variable: $\tau(\hat{D}-\hat{R})/2$							
		2 week poll-to-poll intervals				3 week poll-to-poll intervals			
Explanatory variable	Param.	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)	OLS (7)	2SLS (8)
$X_D^c + X_D^s$	Δ_{sD}^S	0.15	0.18	0.15	0.18	0.15	0.2	0.15	0.2
		(0.04)	(0.09)	(0.04)	(0.087)	(0.3)	(0.10)	(0.03)	(0.10)
$X_R^c + X_R^s$	$-\Delta_{sR}^S$	−0.08	−0.18	−0.08	−0.18	−0.10	−0.27	−0.10	−0.29
		(0.02)	(0.10)	(0.02)	(0.11)	(0.02)	(0.136)	(0.02)	(0.14)
Panel B: First Stages		Dependent variable: $X_D^c + X_D^s$							
Log NFL games/τ			1.73		1.08		19.47		18.30
			(24.07)		(24.05)		(30.50)		(30.48)
Log MLB games/τ			−34.72		−34.62		−36.12		−36.00
			(17.06)		(17.06)		(19.58)		(19.59)
Log NBA games/τ			−12.82		−12.82		−17.99		−19.60
			(13.93)		(13.93)		(16.43)		(16.45)
Log NCAA games/τ			−1322.8		−1321.3		−1348		−1347.5
			(410.13)		(410.02)		(503.09)		(504.2)
R^2			0.70		0.70		0.72		0.72
F test (p-value)			0.007		0.007		0.029		0.028
		Dependent variable: $X_R^c + X_R^s$							
Log NFL games/τ			−34.52		−34.44		−28.01		−29.04
			(12.93)		(12.92)		(16.24)		(16.19)
Log MLB games/τ			−6.99		−7.02		−7.35		−6.77
			(9.16)		(9.17)		(10.42)		(10.41)
Log NBA games/τ			−22.17		−22.13		−31.59		−31.88
			(7.49)		(7.48)		(8.75)		(8.74)
Log NCAA games/τ			−663.0		−663.8		−582.2		−581.7
			(220.34)		(220.31)		(267.91)		(267.84)
R^2			0.73		0.73		0.76		0.76
F test (p-value)			0.000		0.000		0.000		0.000
Race fixed effects		Y	Y	N	N	Y	Y	N	N
Year×State fixed effects		N	N	Y	Y	N	N	Y	Y
No. of Races		415	415	415	415	415	415	415	415
No. of Observations		2134	2134	2134	2134	1865	1865	1865	1865

Table 5: Swing-Voter Gain Elasticities (0.25 score cutoff). The table presents OLS and 2SLS estimates of the swing-voter gain elasticities from equation (17) using a 0.25 article score cutoff. Even-numbered columns present OLS estimates and odd-number columns present 2SLS estimates. Panel A presents estimates for the structural equation (second stage), and panel B presents estimates of the coefficients for the instruments in both the first stages for the Democratic and the Republican total news reports. The first four columns in the table are estimated on the 2-week poll-to-poll interval panel, and the dependent variable is constructed using the parameter estimates from the model in Panel A, column (4) of Table 4. The last four columns are estimated on the 3-week poll-to-poll interval panel, and the dependent variable is constructed using the parameter estimates from the model in Panel A, column (8), of Table 4. All regressions are weighted by the square root of the length in days of the poll-to-poll interval (relative to the longest interval). Columns (1), (2), (5), and (6) include Senate-race fixed-effects. Columns (3), (4), (7), and (8) include a full set of year, state, and year-x-state fixed effects. All models include a dummy variable for the last poll-to-poll interval in a race and month fixed effects. Standard errors are robust to arbitrary heteroskedasticity and to arbitrary serial correlation of up to order 2 following Newey and West (1987). Coefficients and standard errors in Panel A are multiplied by 100.

Electoral Share Response Elasticities and Equilibrium Mixing Strategies

Dependent variable:	2 week poll-to-poll intervals			
	0.25 article score cutoff (1)		0.5 article score cutoff (2)	
Panel A	Parameters (s.e.)			
Δ^T_{cD}	0.016	(0.006)	0.011	(0.005)
Δ^T_{cR}	0.005	(0.005)	0.004	(0.005)
Δ^S_{sD}	0.18	(0.060)	0.12	(0.046)
Δ^S_{sR}	0.18	(0.081)	0.15	(0.062)
Δ^S_{cD}	0.69	(0.15)	0.50	(0.10)
Δ^S_{cR}	0.18	(0.076)	0.15	(0.083)
Panel B	Average Equilibrium Mixing Strategies			
$\mathbb{E}[q_D]$	0.557		0.414	
$\mathbb{E}[q_R]$	0.449		0.299	
$\mathbb{E}[\eta_D(1-\gamma_R)]$	0.018		0.018	
$\mathbb{E}[\eta_R(1-\gamma_D)]$	0.014		0.014	

Table 6: Identified Parameter Estimates and Equilibrium Mixing Strategies. The table presents all the identified parameters (Panel A) and average equilibrium mixing probabilities (Panel B) in the model estimated using the 2-week poll-to-poll interval dataset. Electoral gain elasticities in Panel A are taken from the estimation of equations (16) and (17). Swing-voter loss elasticity parameters are computed according to equation (18) in the text. Column (1) is based on the 0.25 article score cutoff and the estimates in column (4) of Table 4 and column (4) of Table 5. Column (2) is based on analogous models using the 0.5 article score cutoff. Estimates in Panel B are computed directly from the sample analogues as weighted averages using relative interval lengths as weights. Parameter estimates reported in Panel A are multiplied by 100.

similar to the one we discussed above for $(\Delta^T_{cD}, \Delta^T_{cR})$. In the bottom panel of Table 6 we report the average estimates of the unconditional probabilities of observing a news piece, $\eta_p(1-\gamma_{\sim p})$, as measured by the ratio of observed news pieces relative to the number of relevant stage games. Based on the 0.25 article score cutoff criterion, $\mathbb{E}[\eta_D(1-\gamma_R)] \approx 0.018$, and $\mathbb{E}[\eta_R(1-\gamma_D)] \approx 0.014$. A ten percent increase in these probabilities sustained during a month would translate, on average, into a 1.2 percentage point gain to the Democratic candidate, and a 1 percentage point gain to the Republican candidate, stemming from increased swing voter support. These would be losses for the opposing candidate[21].

[21] $(0.1 \times 0.018) \times (0.18/100) \times 124$ media outlets on average $\times$ 30 days ≈ 0.012 for Democrats, and $(0.1 \times 0.014) \times (0.18/100) \times 124$ media outlets on average $\times$ 30 days ≈ 0.010 for Republicans.

The next step in our empirical strategy is to back out estimates of the swing voter responses to core-targeted campaign speech using the equilibrium mixing strategies of the media in equations (3)-(4), together with our estimates $\hat{\phi}_p$ of the conditional reporting probabilities $\eta_p(1-\gamma_{\sim p})$ as described in Proposition 3. We obtain average elasticities by integrating over our sample as follows:

$$\hat{\Delta}^S_{cp} = \frac{\hat{\Delta}^T_{cp}}{\frac{1}{N}\sum_r \sum_{t=1}^{T_r} \hat{\phi}_{p,r,t}} - \hat{\Delta}^S_{sp} \tag{18}$$

Panel A in Table 6 presents the estimates of all six poll-change elasticities in our model. The table presents estimates using the 2-week poll-to-poll interval dataset, both using the ± 0.25 article cutoff classification in column (1), and the ± 0.5 cutoff in column (2). The magnitude of the estimates is very similar for both cutoffs, showing that the specific criterion chosen to classify articles as c or s is not critical for our results. The full set of parameters is also similar when using the 3-week poll-to-poll intervals. These results are omitted to save space. As the table illustrates, the swing voter responsiveness to core-targeted statements is significantly larger for the Democratic candidate than for the Republican candidate. Using the ± 0.25 cutoff estimates, $\Delta^S_{cD} = 0.69$, $\Delta^S_{cR} = 0.18$. Our estimates suggest that swing voters are on average very sensitive to media content that signals relatively core-mobilizing campaign speech by Democrats. This difference in poll response elasticities across parties has substantial implications for the dynamics of the Senate races; although the turnout gains of core-targeted statements are larger for Democrats than for Republicans, the cost on the swing voter margin is even larger. If both centrist voters and the media play a role in moderating candidates' campaign trail speech, these results suggest that swing voters are especially important in fulfilling this role for Democratic candidates, while the media is relatively more important to constrain the behavior of Republican candidates. The equilibrium implication of this pattern of parameters is that candidates from both parties are covered by the media at similar rates. As we pointed out above, our average estimates for these probabilities, reported in Panel B of Table 6, are 0.018 for Democrats and 0.014 for Republicans. These should be understood as the average unconditional probabilities that a given media outlet generates a news piece on a candidate in a given day during the campaign.

6.2 Payoff Heterogeneity

The 2SLS estimates of the payoff parameters $(\Delta^T_{cD}, \Delta^T_{cR}, \Delta^S_{sD}, \Delta^S_{sR}, \Delta^S_{cD}, \Delta^S_{cR})$ described above are average effects across states and three decades, identified off the variation in media coverage and poll changes within races over time. In this section we explore the extent of heterogeneity in these payoff parameters across races. We do so in a straightforward parametric way by allowing the payoff parameters we recover from equations (16) and (17) to depend on race characteristics which may be important sources of heterogeneity. Here we discuss four sources of heterogeneity: the partisan distribution of voters across states and time, the time to election day, the competitiveness of the election at a given point in time, and the presence of an incumbent senator in the race. Specifically, we allow the payoff elasticities to be linear functions of one of these four characteristics $K_{r,t}$: $\Delta^T_{cp} = \alpha^T_{cp} + \beta^T_{cp} K_{r,t}$ and $\Delta^S_{sp} = \alpha^S_{sp} + \beta^S_{sp} K_{r,t}$ for $p \in \{D, R\}$.[22] We estimate equations (16) and (17) by 2SLS including the relevant interaction terms, instrumenting the interaction terms with the respective interactions between our sports events instruments and the source of heterogeneity in each case.[23]

6.2.1 The partisan distribution of voters

We first explore heterogeneity in electoral responses as a function of the partisan distribution of the electorate, which varies considerably across states. We proxy this distribution using the average of the Democratic registration share of the electorate and the most recent presidential election results. For states without partisan registration, we use only the presidential election returns. The results for this exercise are presented in column (1) of Table 7. These and all other estimates in the table use our benchmark 2-week poll-to-poll interval dataset based on the ± 0.25 article score cutoff and are estimated by 2SLS using all sports events and interactions of sports events and voter registration as instruments. Panel A presents the estimates for the turnout elasticities from equation (16), while panel B presents the estimates for the swing voter elasticities from equation (17). Although the pattern of signs implies that Δ^T_{cD} decreases while Δ^T_{cR} increases with Democratic registration, we cannot estimate these effects precisely. In contrast, we find a significant decreasing relationship between Democratic registration and Δ^S_{sD}. In states with relatively few Democratic voters, these

[22] An additional reason to explore heterogeneity in this context is the potential bias of our estimates if parameters vary substantially over time because we base our empirical strategy on the computation of probabilities based on relative frequencies. On the one hand, if the underlying probabilities vary substantially over time, the sample analogue estimators of the mixing probabilities will be biased. This would make shorter poll-to-poll intervals preferable. On the other hand, longer poll-to-poll intervals reduce sampling error, as long as the $\mathbf{\Delta}$'s are constant within a time interval. This is an unavoidable bias-precision trade-off.

[23] To recover the remaining elasticities $\Delta^S_{cp}(K)$ when allowing for heterogeneity, we construct decile bins for $K_{p,r}$ and compute the integration in equation (18) restricted to the set $\Gamma_K = \{(r,t) : K_{r,t} \in K\}$ of observations falling in each decile:

$$\hat{\Delta}^S_{cp}(K) = \frac{\hat{\Delta}^T_{cp}(K_{p,r})}{\frac{1}{|\Gamma_K|}\sum_r \sum_{t=1}^{T_r} \hat{\phi}_{p,r,t}} - \hat{\Delta}^S_{sp}(K_{p,r}), \qquad (r,t) \in \Gamma_K.$$

Testing for Heterogeneity in Electoral Response Elasticities (2 week poll-to-poll intervals, 0.25 article score cutoff)

Panel A Parameter	Regressor	Coefficient	Dem. registration (1)	Log days to election (2)	Race tightness (3)	Incumbent running (4)
			Interaction term (K): Dependent variable: $(\Delta V_D + \Delta V_R)/\tau$			
Δ^T_{cD}	$X^c_D/(X^c_D + X^s_D)$	α^T_{cD}	0.075 (0.037)	0.001 (0.025)	0.019 (0.010)	0.018 (0.009)
	$X^c_D/(X^c_D+X^s_D)\times K$	β^T_{cD}	−0.13 (0.074)	0.001 (0.005)	−0.031 (0.039)	−0.003 (0.011)
Δ^T_{cR}	$X^c_R/(X^c_R + X^s_R)$	α^T_{cR}	−0.060 (0.048)	0.047 (0.041)	0.004 (0.009)	0.005 (0.011)
	$X^c_R/(X^c_R+X^s_R)\times K$	β^T_{cR}	0.14 (0.10)	−0.011 (0.009)	0.008 (0.032)	0.00 (0.014)
Panel B			Dependent variable: $\tau(\hat{D} - \hat{R})/2$			
Δ^S_{sD}	$X^c_D + X^s_D$	α^S_{sD}	0.037 (0.009)	0.0028 (0.008)	0.0014 (0.001)	0.0007 (0.001)
	$(X^c_D + X^s_D) \times K$	β^S_{sD}	−0.078 (0.018)	−0.0001 (0.001)	−0.002 (0.002)	0.0008 (0.001)
$-\Delta^S_{sR}$	$X^c_R + X^s_R$	α^S_{sR}	0.0020 (0.014)	−0.0163 (0.014)	0.0009 (0.002)	0.000 (0.002)
	$(X^c_R + X^s_R) \times K$	β^S_{sR}	−0.007 (0.028)	0.003 (0.003)	−0.014 (0.005)	−0.0015 (0.002)
	$\mathbb{E}[K]$		0.47	4.56	0.17	0.76

Table 7: Heterogeneity in Electoral Response Elasticities. The table presents parameter estimates from 2SLS models that include an interaction between a race characteristics and the endogenous explanatory variables. All models are estimated on the 2-week poll-to-poll interval panel using 2134 observations from 415 Senate races. Models in column (1) allow for an interaction with the Democratic registration in the state as defined in the text. Models in column (2) allow for an interaction with the log of days to the general election. Because this variable varies across poll-to-poll intervals and races, log days to the general election is also included as a covariate. Models in column (3) allow for an interaction with a proxy for the competitiveness of the race measured as the absolute value of the difference between the Democratic and Republican poll results in the beginning on the poll-to-poll interval. Because this variable varies across poll-to-poll intervals and races, race competitiveness is also included as a covariate. Models in column (4) allow for an interaction with a dummy variable for races where an incumbent senator is running. The dependent variable for the equation in Panel B uses the parameter estimates from the corresponding column of Panel A. All models include Senate-race fixed effects, month fixed effects, and a dummy variable for the last poll-to-poll interval in a race. The set of instruments includes the log of NFL games per day, the log of MLB games per day, the log of NBA games per day, the log of $NCAA$ games per day, and interactions of each of these variables with the corresponding interaction variable. All regressions are weighted by the square root of the length in days of the poll-to-poll interval (relative to the longest interval). Standard errors are robust to arbitrary heteroskedasticity and to arbitrary serial correlation of up to order 2 following Newey and West (1987). Coefficients and standard errors in Panel A are multiplied by 100.

voters appear to be more responsive to swing-voter targeted media coverage favoring the Democratic candidates. Except for this result, the partisan distribution of the electorate does not appear to be a major source of heterogeneity in electoral response elasticities across states or over time.

6.2.2 Days to Election

In a second exercise we explore the possibility that the electoral responsiveness of voters varies during the campaign. For example, if voters pay more attention to media coverage as November approaches, they may become more responsive to the news over time. We explore this possibility by allowing the payoff parameters in equations (16) and (17) to depend on the time between the initial date of the poll-to-poll interval and the general election date. Because the time to election day varies across poll-to-poll intervals within race, we also include the time to election as a covariate. Our main results for this exercise are presented in column (2) of Table 7. They show no statistically significant evidence of heterogeneity on time to election day. Overall, the Δ's appear to be stable along the campaign trail.

6.2.3 State of the Race: A Dynamic Game

We also explore whether poll response elasticities vary as a function of an endogenous state variable, making the game in practice a dynamic one rather than a repeated one. It is possible that both candidates' and the media's incentives change along the campaign trail as a function of the political environment and the previous evolution of the race itself. On the one hand, we may expect a candidate to become more willing to take risks when he is behind in the polls. On the other hand, the electoral cost of bad press may grow as election day approaches, making politicians more cautions late in the race. Similarly, the media's campaign coverage profitability may grow as election day approaches. To explore these possibilities and their implications for the robustness of our results, we allow payoff parameters to depend on the current state of the race as measured by the poll margin between candidates at the beginning of the corresponding poll-to-poll interval.[24]

We now have a dynamic game where payoffs depend on a state variable, and where the state variable itself evolves over time as a function of the players' previous choices. Even in this case, the finite horizon of the game and the uniqueness of Nash equilibrium in its stage game imply that the dynamic game only has one sub-game perfect equilibrium. It prescribes playing the mixed-strategy Nash equilibrium of the stage game given the value

[24] In principle, the relevant state variable may be a high-dimensional vector of time-varying characteristics. In practice, our sample size requires us to limit the dimensionality of the state variable we consider.

of the state variable at every period. As a result, equilibrium play is independent across periods conditional on the state variable, and we can replicate our estimation strategy from above. Similar to the time-to-election exercise, the poll margin varies over time within a race, so we also include it separately as a covariate.

Results for this exercise are presented in column (3) of Table 7. Overall, we do not find a strong relationship between the state of the race and the electoral support elasticities. The only exception arises for the swing-voter response elasticity to swing-targeted speech for Republicans, Δ^S_{sR}. This parameter is higher in more competitive periods of a race. Taken at face value, it suggests that incentives to target swing voters become stronger for Republican candidates as races become closer. The results from this exercise should be taken with caution because the poll margin is an endogenous outcome which we are including as a covariate.

6.2.4 Incumbent Running

Our final exercise looking at payoff heterogeneity explores whether poll responsiveness differs in races where incumbents are running. We allow the Δ's to depend on a dummy variable for elections with a running incumbent. Results for this exercise are presented in column (4) of Table 7. We find no evidence of differences in candidate payoff parameters in races with or without incumbents. We should notice, however, that this test may not have much power: 75% of all Senate races in our sample have an incumbent running.

6.3 Robustness Exercises and Specification Tests

6.3.1 Robustness Exercises

In Tables 8 and 9 we present a subset of additional econometric exercises exploring the robustness of our main findings. Table 8 reports 2SLS results for alternative specifications based on the 2-week poll-to-poll interval dataset. First, we estimate equations (16)-(17) excluding the last poll-to-poll interval for each race. We do this for two reasons. First, our last poll-to-poll interval for each race is constructed using the general election result as the end-of period value. This is in contrast to all other periods in which beginning and end-of-period electoral support are measured using averages of polls. Second, the validity of our instruments relies on the assumption that sports events are shifters of the media's reporting payoffs, but do not otherwise affect the evolution of the polls. If sports events that happen very near election day –thus falling on the last poll-to-poll interval– directly lead to lower turnout in elections, the exclusion restriction would not be satisfied.[25] Excluding these

[25]We believe this is unlikely given that poll-to-poll intervals cover an average of 30 days.

Robustness exercises (2 week poll-to-poll intervals)

		Robustness exercise:			
Panel A		Excluding last poll-to-poll interval	Controlling for post-primary dummy	0.5 article score cutoff	0.75 article score cutoff
Parameter	Regressor	(1)	(2)	(3)	(4)
		Dependent variable: $(\Delta V_D + \Delta V_R)/\tau$			
Δ^T_{cD}	$X^c_D/(X^c_D + X^s_D)$	0.015	0.014	0.011	0.015
		(0.005)	(0.005)	(0.005)	(0.006)
Δ^T_{cR}	$X^c_R/(X^c_R + X^s_R)$	0.003	0.008	0.004	0.005
		(0.005)	(0.006)	(0.005)	(0.006)
Panel B		Dependent variable: $\tau(\hat{D} - \hat{R})/2$			
Δ^S_{sD}	$X^c_D + X^s_D$	0.21	0.13	0.12	0.12
		(0.08)	(0.056)	(0.056)	(0.06)
$-\Delta^S_{sR}$	$X^c_R + X^s_R$	−0.24	−0.16	−0.15	−0.14
		(0.12)	(0.082)	(0.076)	(0.078)
No. of races		415	415	415	415
No. of observations		1871	2134	2134	2134

Table 8: Robustness Exercises. The table presents 2SLS estimates of the electoral support elasticities from equations (22) and (23). All models are estimated on the 2 week poll-to-poll interval panel, and include a full set of Senate-race fixed effects, and month fixed effects. The dependent variable in Panel B is constructed using the parameter estimates from Panel A. All regressions are weighted by the square root of the length in days of the poll-to-poll interval (relative to the longest interval). Column (1) excludes all observations consisting of the last poll-to-poll interval in a race. Columns (2), (3), and (4) include a dummy variable for the last poll-to-poll interval in a race. All models use log of NFL games per day, log of MLB games per day, log of NBA games per day, and log of $NCAA$ games per day as instruments. Standard errors are robust to arbitrary heteroskedasticity and to arbitrary serial correlation of up to order 2 following Newey and West (1987). Coefficients and standard errors in Panels A and B are multiplied by 100.

observations reduces the sample size from 2134 to 1871. As column (1) in Table 8 shows, the magnitude and significance of the estimated parameters is almost unchanged relative to our baseline estimates.

In column (2) we then include a dummy variable for poll-to-poll intervals after the primary election for the race. If the strategic environment is significantly different before and after the primaries have taken place, it may be important to distinguish between both regimes. For most races, even during primary campaign days, pollsters are already collecting polls asking for the candidates who eventually become the Democratic and Republican nominees. This suggests that in most cases, the bipartisan race is already implicitly taking place before the primary outcome is known. As column (2) in Table 8 shows, controlling for a post-primary dummy variable does not alter any of our benchmark estimates either.

Finally, in columns (3) and (4) of Table 8 we estimate our main specification using two alternative article score cutoffs. Column (3) presents estimates using a ± 0.5 cutoff, and column (4) presents estimates using a quite extreme $\pm .75$ cutoff. Because our classification cutoff for core-targeted versus swing-targeted news content is arbitrary, it is reassuring that our main results are unaltered.

In Table 9 we turn to a sensitivity analysis of our estimates to the inclusion of alternative subsets of our sports events instruments. These, in practice, amount to over-identification exercises. We report the results from models using the 2-week (columns (1)-(5)) and the 3-week (columns (6)-(10)) poll-to-poll interval datasets, using the ± 0.25 article score cutoff classification. Panel AI presents the parameter estimates for equation (16). Panel BI presents the parameter estimates for equation (17). Panels AII and BII present diagnostic statistics for the respective first stages that include different subsets of instruments. We present results that omit one by one each of the four sports events from the instrument set in columns (1)-(4) and (6)-(9). In columns (5) and (10) we also include a more demanding specification where we omit both MLB and NBA games simultaneously, making these models exactly identified. The F-tests for the excluded instruments across the table do suggest that we lose some of the joint predictive power of our instruments when excluding some of them. However, we fail to reject the null of no joint significance in only 4 out of the 40 first stages reported in the Table. Standard errors for the parameter estimates are also somewhat larger, but in most cases the parameter estimates that are significant in our benchmark specification using all instruments remain significant at the 5% level when using only a subset of them. More importantly, the table shows that the magnitude and pattern of signs for the estimated parameters remain unchanged relative to our baseline model estimates.

Overidentification Exercises

		2 week poll-to-poll intervals					3 week poll-to-poll intervals				
Excluded Instruments:		NFL	MLB	NBA	NCAA	MLB, NBA	NFL	MLB	NBA	NCAA	MLB, NBA
Panel AI: Structural Equation		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Regressor	Param.	Dependent variable: $(\Delta V_D + \Delta V_R)/\tau$									
$X_D^c/(X_D^c + X_D^s)$	Δ_{cD}^T	0.014	0.019	0.011	0. 024	0.015	0.009	0.018	0.020	0.02	0.021
		(0.006)	(0.008)	(0.006)	(0.010)	(0.009)	(0.007)	(0.010)	(0.010)	(0.007)	(0.012)
$X_R^c/(X_R^c + X_R^s)$	Δ_{cR}^T	0.003	0.013	0.004	0.00	0.013	0.005	0.0013	0.009	0.01	0.012
		(0.006)	(0.010)	(0.004)	(0.01)	(0.009)	(0.007)	(0.012)	(0.008)	(0.01)	(0.012)
Panel AII: First Stages		Dependent variable: $X_D^c/(X_D^c + X_D^s)$									
R^2		0.78	0.78	0.78	0.78	0.78	0.80	0.80	0.80	0.80	0.80
F test (p-value)		0.014	0.011	0.058	0.011	0.037	0.071	0.015	0.082	0.014	0.042
		Dependent variable: $X_R^c/(X_R^c + X_R^s)$									
R^2		0.80	0.80	0.80	0.80	0.80	0.82	0.82	0.82	0.82	0.82
F test (p-value)		0.002	0.076	0.000	0.001	0.032	0.021	0.067	0.019	0.166	0.053
Panel BI: Structural Equation											
Regressor	Param.	Dependent variable: $\tau(\hat{D} - \hat{R})/2$									
$X_D^c + X_D^s$	Δ_{sD}^S	0.10	0.24	0.18	0.10	0.30	0.12	0.27	0.19	0.15	0.32
		(0.07)	(0.10)	(0.08)	(0.11)	(0.14)	(0.07)	(0.11)	(0.09)	(0.12)	(0.17)
$X_R^c + X_R^s$	$-\Delta_{sR}^S$	−0.02	−0.21	−0.20	−0.19	−0.33	−0.15	−0.30	−0.29	−0.27	−0.46
		(0.16)	(0.11)	(0.12)	(0.11)	(0.17)	(0.013)	(0.14)	(0.17)	(0.12)	(0.28)
Panel BII: First Stages		Dependent variable: $X_D^c + X_D^s$									
R^2		0.60	0.60	0.60	0.60	0.60	0.63	0.63	0.63	0.63	0.63
F test (p-value)		0.062	0.095	0.046	0.495	0.047	0.090	0.134	0.088	0.672	0.062
		Dependent variable: $X_R^c + X_R^s$									
R^2		0.60	0.60	0.60	0.60	0.60	0.64	0.64	0.64	0.64	0.64
F test (p-value)		0.009	0.001	0.003	0.013	0.001	0.016	0.006	0.048	0.037	0.023
No. of races		415	415	415	415	415	415	415	415	415	415
No. of observations		2134	2134	2134	2134	2134	1865	1865	1865	1865	1865

Table 9: Overidentification Exercises. The table presents 2SLS estimates of the electoral support elasticities from equations (16) in Panel A, and (17) in Panel B. All models are estimated using the 0.25 article score cutoff, and include a full set of Senate-race fixed effects, month fixed effects, and a dummy variable for the last poll-to-poll interval in a race. The dependent variable for the structural equation in Panel BI is constructed using the benchmark parameter estimates from the structural equation (16). Columns (1)-(5) are based on the estimate from column (4) of Table 4. Columns (6)-(10) are based on the estimate from column (8) in Table 4. Panels AI and BI present estimates for the structural equations (second stages), and Panels AII and BII report the corresponding R-squared and p-value for the F-tests on the excluded instruments for each first stage. The dependent variables in the first stages of Panel AII are the Democratic and Republican ratios of turnout-targeted to total news reports. The dependent variables in the first stages of Panel BII are the Democratic and Republican total news reports. The first five columns in the table are estimated on the 2 week poll-to-poll interval panel. The last five columns are estimated on the 3 week poll-to-poll interval panel. All regressions are weighted by the square root of the length in days of the poll-to-poll interval (relative to the longest interval). Columns (1) and (6) exclude the log number of NFL games per day from the instrument set. Columns (2) and (7) exclude the log number of MLB games per day from the instrument set. Columns(3) and (8) exclude the log number of NBA games per day from the instrument set. Columns (4) and (9) exclude the log number of $NCAA$ games per day from the instrument set. Columns (5) and (10) present a just-identified model excluding the log number of MLB games per day and the log number of NBA games per day from the instrument set. Standard errors are robust to arbitrary heteroskedasticity and to arbitrary serial correlation of up to order 2 following Newey and West (1987). Coefficients and standard errors in Panels AI and BI are multiplied by 100.

Testing Model Assuptions: Poll Timing and Race Tightness

Dependent variable:	Polls in poll-to-poll interval		$\frac{\text{Polls in poll-to-poll interval}}{\text{Length of poll-to-poll interval}}$	
Poll-to-poll interval size:	2 week	3 week	2 week	3 week
	(1)	(2)	(3)	(4)
Race tightness ($\lvert V_D - V_R \rvert$)	0.041	0.100	0.001	0.025
	(0.506)	(0.697)	(0.066)	(0.077)
R^2	0.39	0.48	0.42	0.47
No. of Races	415	415	415	415
No. of Observations	2134	1865	2134	1865

Table 10: Testing Model Assumptions: Poll Coverage Intensity and Race Competitiveness: The table presents OLS panel regressions of a measure of poll coverage intensity on the tightness of the Senate race as measured by the absolute value of the difference between the Democratic candidate's electoral support and the Republican candidate's electoral support. The dependent variable in columns (1) and (2) is the number of polls in the poll-to-poll interval. The dependent variable in columns (3) and (4) is the number of polls per day in the poll-to-poll interval. All models include a full set of Senate-race fixed effects, month fixed effects, a dummy variable for the last poll-to-poll interval in a race, and a constant. All regressions are weighted by the square root of the length in days of the poll-to-poll interval (relative to the longest interval). Standard errors are robust to arbitrary heteroskedasticity.

6.3.2 A Test for Poll Timing Independence

Finally, we are also able to indirectly test whether the timing of polls appears to be uncorrelated with the evolution of the Senate races. Recall from our discussion in section 4.2 that this underlies the validity of our method for building the poll-to-poll intervals which determine the panel structure of our dataset. We do this by exploring the correlation between the frequency of actual polls in our dataset and the competitiveness of the race at any given point in time. In Table 10 we report results from OLS regressions of the number of actual polls used to construct the average end-poll of each poll-to-poll interval, on the measure of race competitiveness we introduced in section 6.2. We present results with or without normalizing by the length of the interval in days, and for both the 2-week and the 3-week poll-to-poll interval datasets. As the table illustrates, we find no correlation between poll frequencies and the state of the race. Pollsters do not appear to be releasing polls as a function of how the race is evolving. We see these results, together with those using alternative poll-to-poll windows, as reassuring.

6.4 Partial Identification of the Media Reporting Biases

We conclude the discussion of our empirical findings showing that the media's partisan reporting biases are only partially identified in our model. We then compute their identified set.

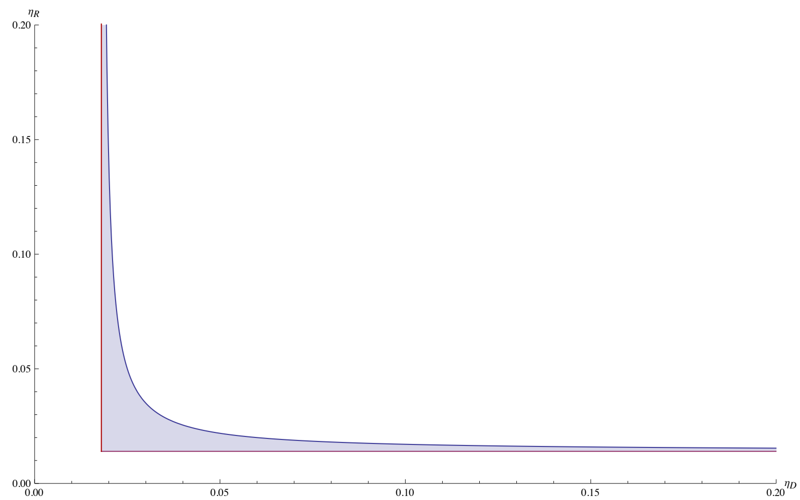


Figure 7: Identified set for media biases (η_D, η_R).

Although our identification strategy allows us to recover all of the payoff parameters governing the poll change technology $(\Delta^T_{cD}, \Delta^T_{cR}, \Delta^S_{cD}, \Delta^S_{cR}, \Delta^S_{sD}, \Delta^S_{sR})$ and the conditional probabilities of a media report $(\eta_p(1-\gamma_{\sim p}))$, we cannot separately identify the η_p's from the γ_p's. The reason is that observed counts of reports are the outcome of both media coverage and successful reporting, and these events are indistinguishable based on observed news reports alone. In equilibrium, observed media reporting on a given candidate results from the interaction between the reporting bias and the media's mixed strategy. Our model does provide, however, additional structure to partially identify the η_p's (and thus the γ_p's through equations (3)-(4)). Equation (11) directly allows us to solve for γ_p as a function of the media bias for the opposing candidate $\eta_{\sim p}$, and the reduced form parameter $\hat{\phi}_{\sim p}$: $\gamma_p = 1 - \frac{1}{\eta_{\sim p}}\hat{\phi}_{\sim p}$. Because $\gamma_p \in (0,1)$, it follows that $\eta_D > \hat{\phi}_D$ and $\eta_R > \hat{\phi}_R$. Furthermore, $\gamma_D + \gamma_R < 1$ implies that

$$\eta_R < \frac{\hat{\phi}_R}{1 - \frac{1}{\eta_D}\hat{\phi}_D} \tag{19}$$

These three inequalities determine the identified set for (η_D, η_R), which we illustrate in Figure 7 at our benchmark estimates for $(\hat{\phi}_D, \hat{\phi}_R)$. The hyperbola represents the constraint in the right-hand side of equation 19.[26] As Figure 7 illustrates, our estimates restrict con-

[26]Notice that we can exploit the candidates' equilibrium mixing probabilities from equations (5)-(6) together with our identified set for the η_p's to obtain an identified set for the relative payoff to the media from reporting about R and D:

$$\frac{\pi_R}{\pi_D} = \frac{\eta_D}{\eta_R}\frac{\hat{q}^*_D}{\hat{q}^*_R}$$

which we can trace on the identified set for (η_D, η_R) at the estimated q_p's. These can be interpreted as siderably the range of values that the η_p's can take. In particular, the region over which very asymmetric values for η_D and η_R are feasible is very small. For example, if we were to consider a uniform prior over this identified set, most of the density would fall over a region where η_D is close in magnitude to η_R. This is the basis of our assertion that a large media bias as measured by a large difference in the η_p's across parties is very unlikely.[27]

7 Concluding Remarks

In this paper we develop a framework to study how the interaction between the media's incentives to cover and report on electoral campaigns, and candidates' incentives to target different groups of voters, shape both campaign trail speech, and the evolution of the races themselves. We do this by proposing a simple game-theoretic model of the interaction between the media and candidates, where the media gains from reporting on core-supporter targeted campaign speech from candidates, while candidates benefit from reports about swing voter-targeted messages. Because candidates have incentives to target both types of constituencies, this strategic environment is similar to a standard matching pennies game.

The simple structure of the game allows us to propose an empirical strategy to estimate this discrete game of complete information, and to test for its empirical relevance. The significance of our results should be interpreted as a test of the empirical plausibility of the premises on which we build our structural model. We use information on U.S. Senate races from the last 30 years, which are politically salient and thus, systematically covered by the media and by pollsters. We show how polling data measuring the evolution of the campaigns, together with media coverage information based on a text analysis methodology, can be used to estimate the key parameters of the game. Our results suggest the mechanism we propose here is important for understanding the nature of bipartisan electoral competition in settings with ample media presence. Moreover, our model and results provide a novel way of thinking about how and why the media matters in politics, by highlighting not only that the media shapes candidate behavior, but also that candidates shape the way the media reports about politics. Our empirical findings suggest a large asymmetry across Democratic and Republican candidates in their incentives to target core supporters. While turnout appears more responsive to core-targeted speech by Democrats, swing-voters also appear more willing to change allegiances towards Republicans when the media reports widely on

bounds on the relative media payoffs from Democratic versus Republican coverage.

[27] Although we could average over the identified set for the η_p's, incorporating researcher uncertainty about these parameters, our data and empirical strategy does not allow us to identify the reporting costs κ_p of the average media outlet either. As a result, we are unable to use the identified parameters to perform counterfactual exercises to predict the evolution of Senate races.

core-targeted messaging by Democrats. Exploring the nature of the differential responses of the electorate to targeted campaign messages may be a fruitful area for future research.

References

Ansolabehere, S., R. Behr, and S. Iyengar (1992): *The Media Game: American Politics in the Television Age*, Longman.

Bajari, P., H. Hong, and S. Ryan (2010): "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica*, 78, 1529–1568.

Baron, D. (2006): "Persistent Media Bias," *Journal of Public Economics*, 90, 1–36.

Bartels, L. (1996): "Politicians and the Press: Who Leads, Who Follows," Princeton.

Berry, S. (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60, 889–917.

Besley, T. and A. Prat (2006): "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability," *American Economic Review*, 96, 720–736.

Bresnahan, T. and P. Reiss (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553.

——— (1991): "Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57–81.

Budge, I. and R. Hofferbert (1990): "Mandates and Policy Outputs: U.S. Party Platforms and Federal Expenditures," *The American Political Science Review*, 84, 111–131.

Campante, F. and D. Hojman (2010): "Media and Polarization," Harvard.

Chiang, C.-F. and B. Knight (2011): "Media Bias and Influence: Evidence from Newspaper Endorsements," *The Review of Economic Studies*, 78, 795–820.

Chiappori, P., S. Levitt, and T. Groseclose (2002): "Testing Mixed-Strategy Equilibria When Players are Heterogeneous: The Case of Penalty Kicks in Soccer," *American Economic Review*, 92, 1138–1151.

Corneo, G. (2006): "Media Capture in a Democracy: The Role of Wealth Concentration," *Journal of Public Economics*, 90, 37–58.

Cox, G. and M. McCubbins (1986): "Electoral Politics as a Redistributive Game," *Journal of Politics*, 48, 370–389.

DellaVigna, S. and E. Kaplan (2007): "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 122, 1187–1234.

Durante, R. and E. Zhuravskaya (forthcoming): "Attack when the World is Not Watching? U.S. News and the Israeli-Palestinian Conflict," *Journal of Political Economy*.

Egorov, G. (2015): "Single-Issue Campaigns and Multidimensional Politics," Northwestern.

Eisensee, T. and D. Strömberg (2007): "News Droughts, News Floods, and U.S. Disaster Relief," *Quarterly Journal of Economics*, 122, 693–728.

Enikolopov, R., M. Petrova, and E. Zhuravskaya (2011): "Media and Political Persuasion: Evidence from Russia," *American Economic Review*, 111, 3253–3285.

Ferraz, C. and F. Finan (2011): "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments," *The American Economic Review*, 101, 1274–1311.

Fonseca, A., J. M. S. Jr., and B. K. Song (2014): "Newspaper Market Structure and Behavior: Partisan Coverage of Political Scandals in the U.S. from 1870 to 1910," .

Gentzkow, M. and J. Shapiro (2006): "Media Bias and Reputation," *Journal of Political Economy*, 114, 280–316.

——— (2010): "What Drives Media Slant? Evidence from U.S. Daily Newspapers," *Econometrica*, 78, 35–71.

George, L. and J. Waldfogel (2006): "The New York Times and the Market for Local Newspapers," *The American Economic Review*, 96, 435–447.

Healy, A., N. Malhotra, and C. H. Mo (2010): "Irrelevant Effects Affect Voter's Evaluations of Government Performance," *Proceedings of the National Academy of Sciences*, 107, 12804–12809.

Hirano, S., G. Lenz, M. Pinkovskiy, and J. Snyder (2015): "Voter Learning in State Primary Elections," *American Journal of Political Science*, 59, 91–108.

Knowles, J., N. Persico, and P. Todd (2001): "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 109, 203–229.

Kurkones, M. (1984): *Promises and Performance: Presidential Campaigns as Policy Predictors*, University Press of America.

Lindbeck, A. and J. Weibull (1987): "Balanced Budget Redistribution as the Outcome of Political Competition," *Public Choice*, 52, 273–297.

Maskin, E. and J. Tirole (2004): "The Politician and the Judge: Accountability in Government," *American Economic Review*, 94, 1034–1054.

Mullainathan, S. and A. Shleifer (2005): "The Market for News," *American Economic Review*, 95, 1031–1053.

Newey, W. and K. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Oberholzer-Gee, F. and J. Waldfogel (2009): "Media Markets and Localism: Does Local News en Espanol Boost Hispanic Voter Turnout?" *The American Economic Review*, 99, 2120–2128.

Palacios-Huerta, I. (2003): "Professionals Play Minimax," *Review of Economic Studies*, 70, 395–415.

Prat, A. (2005): "The Wrong Kind of Transparency," *American Economic Review*, 95, 862–877.

Prat, A. and D. Strömberg (2013): "The Political Economy of Mass Media," .

Puglisi, R. (2006): "Being the New York Times: The Political Behavior of a Newspaper," LSE.

Puglisi, R. and J. Snyder (2008): "Media Coverage of Political Scandals," Mimeo.

Snyder, J. and D. Strömberg (2010): "Press Coverage and Political Accountability," *Journal of Political Economy*, 118.

Strömberg, D. (2004a): "Mass Media Competition, Political Competition, and Public Policy," *Review of Economic Studies*, 71, 265–284.

——— (2004b): "Radio's Impact on Public Spending," *Quarterly Journal of Economics*, 119, 189–221.

Walker, M. and J. Wooders (2001): "Minimax Play at Wimbledon," *American Economic Review*, 91, 1521–1538.

A Appendix A

A.1 Proof of Proposition 1

The normal form game G is presented in table 11 below. In each cell, the payoffs are written in the order (D, R, M).

Existence and Uniqueness:

Define the following parameters:

$$
\begin{array}{cc}
\triangle_1 \equiv \Delta^T_{cD} - \eta_D \Delta^S_{cD} + \eta_R \Delta^S_{cR} & \triangle_{13} \equiv \Delta^T_{cD} - \eta_D \Delta^S_{cD} - \eta_R \Delta^S_{sR} \\
\triangle_2 \equiv \Delta^T_{cR} - \eta_R \Delta^S_{cR} + \eta_D \Delta^S_{cD} & \triangle_{14} \equiv \eta_R \Delta^S_{sR} + \eta_D \Delta^S_{cD} \\
\triangle_3 \equiv \Delta^T_{cD} - \eta_D \Delta^S_{cD} & \triangle_{15} \equiv \Delta^T_{cD} - \eta_D \Delta^S_{cD} \\
\triangle_4 \equiv \Delta^T_{cR} + \eta_D \Delta^S_{cD} & \triangle_{16} \equiv \eta_D \Delta^S_{cD} \\
\triangle_5 \equiv \Delta^T_{cD} + \eta_R \Delta^S_{cR} & \triangle_{17} \equiv \Delta^T_{cD} - \eta_R \Delta^S_{sR} \\
\triangle_6 \equiv \Delta^T_{cR} - \eta_R \Delta^S_{cR} & \triangle_{18} \equiv \eta_R \Delta^S_{sR} \\
\triangle_7 \equiv \eta_D \Delta^S_{sD} + \eta_R \Delta^S_{cR} & \triangle_{19} \equiv \eta_D \Delta^S_{sD} - \eta_R \Delta^S_{sR} \\
\triangle_8 \equiv \Delta^T_{cR} - \eta_R \Delta^S_{cR} - \eta_D \Delta^S_{sD} & \triangle_{20} \equiv \eta_R \Delta^S_{sR} - \eta_D \Delta^S_{sD} \\
\triangle_9 \equiv \eta_D \Delta^S_{sD} & \triangle_{21} \equiv \eta_D \Delta^S_{sD} \\
\triangle_{10} \equiv \Delta^T_{cR} - \eta_D \Delta^S_{sD} & \triangle_{22} \equiv -\eta_D \Delta^S_{sD} \\
\triangle_{11} \equiv \eta_R \Delta^S_{cR} & \triangle_{23} \equiv -\eta_R \Delta^S_{sR} \\
\triangle_{12} \equiv \Delta^T_{cR} - \eta_R \Delta^S_{cR} & \triangle_{24} \equiv \eta_R \Delta^S_{sR}
\end{array}
$$

G is a game with finite action space, which is sufficient for existence of a Nash equilibrium. Checking the non-existence of a Nash equilibrium in pure strategies is straightforward. Thus, any equilibria must be in mixed strategies. Conditions for such an equilibrium are:

1. M must be indifferent between playing $a^M = F_D F_R$ and $a^M = F_D N_R$:

$$\mathbb{E}[U_M | F_D F_R] = q_D q_R(\eta_D \pi_D + \eta_R \pi_R - 2k) + (1-q_D) q_R(\eta_R \pi_R - 2k) + q_D(1-q_R)(\eta_D \pi_D - 2k) + (1-q_D)(1-q_R)(-2k)$$

Democrat's Action		Media's action: $a^M = F_D F_R$	$a^M = F_D N_R$	$a^M = N_D F_R$	Republican's Action
	$a^D = c$	$(\Delta_1, \Delta_2,\ \eta_D \pi_D + \eta_R \pi_R - 2k)$	$(\Delta_3, \Delta_4,\ n_D \pi_D - k)$	$(\Delta_5, \Delta_6,\ n_R \pi_R - k)$	$a^R = c$
	$a^D = s$	$(\Delta_7, \Delta_8,\ n_R \pi_R - 2k)$	$(\Delta_9, \Delta_{10},\ -k)$	$(\Delta_{11}, \Delta_{12},\ n_R \pi_R - k)$	$a^R = c$
	$a^D = c$	$(\Delta_{13}, \Delta_{14},\ \eta_D \pi_D - 2k)$	$(\Delta_{15}, \Delta_{16},\ n_D \pi_D - k)$	$(\Delta_{17}, \Delta_{18},\ -k)$	$a^R = s$
	$a^D = s$	$(\Delta_{19}, \Delta_{20},\ -2k)$	$(\Delta_{21}, \Delta_{22},\ -k)$	$(\Delta_{23}, \Delta_{24},\ -k)$	$a^R = s$

Table 11: Normal Form of the Stage Game.

$$= q_D q_R(\eta_D \pi_D - k) + (1-q_D)q_R(-k) + q_D(1-q_R)(\eta_D\pi_D - k) + (1-q_D)(1-q_R)(-k) = \mathbb{E}[U_M|F_D N_R]$$

$$\Leftrightarrow q_R^* = \frac{k}{\eta_R \pi_R} \tag{A.1}$$

2. M must be indifferent between $a^M = F_D F_R$ and $a^R = N_D F_R$:

$$\mathbb{E}[U_M|F_D F_R] = q_D q_R(\eta_D\pi_D + \eta_R\pi_R - 2k) + (1-q_D)q_R(\eta_R\pi_R - 2k) + q_D(1-q_R)(\eta_D\pi_D - 2k) + (1-q_D)(1-q_R)(-2k)$$

$$= q_D q_R(\eta_R\pi_R - k) + (1-q_D)q_R(\eta_R\pi_R - k) + q_D(1-q_R)(-k) + (1-q_D)(1-q_R)(-k) = \mathbb{E}[U_M|N_D F_R]$$

$$\Leftrightarrow q_D^* = \frac{k}{\eta_D \pi_D} \tag{A.2}$$

3. D must be indifferent between $a^D = c$ and $a^D = s$:

$$\mathbb{E}[U_D|c] = (1-\gamma_D-\gamma_R)q_R\triangle_1 + \gamma_D q_R\triangle_3 + \gamma_R q_R \triangle_5$$

$$+(1-\gamma_D-\gamma_R)(1-q_R)\triangle_{13} + \gamma_D(1-q_R)\triangle_{15} + \gamma_R(1-q_R)\triangle_{17}$$

$$= (1-\gamma_D-\gamma_R)q_R\triangle_7 + \gamma_D q_R\triangle_9 + \gamma_R q_R\triangle_{11}$$

$$+(1-\gamma_D-\gamma_R)(1-q_R)\triangle_{19} + \gamma_D(1-q_R)\triangle_{21} + \gamma_R(1-q_R)\triangle_{23} = \mathbb{E}[U_D|s]$$

$$\Leftrightarrow \gamma_R^* = 1 - \frac{\Delta_{cD}^T}{\eta_D\left[\Delta_{cD}^S + \Delta_{sD}^S\right]} \tag{A.3}$$

4. R must be indifferent between $a^D = c$ and $a^D = s$:

$$\mathbb{E}[U_R|c] = (1-\gamma_D-\gamma_R)q_D\triangle_2 + \gamma_D q_D\triangle_4 + \gamma_R q_D\triangle_6$$

$$+(1-\gamma_D-\gamma_R)(1-q_D)\triangle_8 + \gamma_D(1-q_D)\triangle_{10} + \gamma_R(1-q_D)\triangle_{12}$$

$$= (1-\gamma_D-\gamma_R)q_D\triangle_{14} + \gamma_D q_D\triangle_{16} + \gamma_R q_D\triangle_{18}$$

$$+(1-\gamma_D-\gamma_R)(1-q_D)\triangle_{20} + \gamma_D(1-q_D)\triangle_{22} + \gamma_R(1-q_D)\triangle_{24} = \mathbb{E}[U_R|s]$$

$$\Leftrightarrow \gamma_D^* = 1 - \frac{\Delta_{cR}^T}{\eta_R\left[\Delta_{cR}^S + \Delta_{sR}^S\right]} \tag{A.4}$$

Thus, the mixed-strategy Nash equilibrium is unique.

A.2 Proof of Proposition 2

Consider taking the difference between the poll outcomes for a candidate between stage games $t+\tau$ and t. From equation (2) across τ stage games, the change in electoral support to candidate $p \in \{D, R\}$ is given by

$$v^p(t+\tau) - v^p(t) = \Delta_{cp}^T N_p^c(t, t+\tau) + \left(\Delta_{cp}^T - \Delta_{cp}^S\right) X_p^c(t, t+\tau)$$

$$+ \Delta_{c\sim p}^S X_{\sim p}^c(t, t+\tau) + \Delta_{sp}^S X_p^s(t, t+\tau) - \Delta_{s\sim p}^S X_{\sim p}^s(t, t+\tau) + \epsilon^p(t, t+\tau) \tag{A.5}$$

where $\epsilon^p(t, t+\tau) = \sum_{\iota=t+1}^{t+\tau} \epsilon^p(\iota)$. Now add together the equations for each candidate, and divide by the number of stage games in the interval. The zero-sum nature of swing-voter support implies all swing voter effects cancel out, and we are left with an expression that only depends on the counts of events that generate electoral responses on the turnout margin:

$$\frac{v^D(t+\tau) - v^D(t) + v^R(t+\tau) - v^R(t)}{\tau} =$$

$$\Delta_{cD}^T \frac{N_D^c(t, t+\tau) + X_D^c(t, t+\tau)}{\tau} + \Delta_{cR}^T \frac{N_R^c(t, t+\tau) + X_R^c(t, t+\tau)}{\tau} + \tilde{\omega}(t, t+\tau) \tag{A.6}$$

where $\tilde{\omega}(t,t+\tau) \equiv \frac{\epsilon^D(t,t+\tau)}{\tau} + \frac{\epsilon^R(t,t+\tau)}{\tau}$. This specification cannot be estimated because the (N_D^c, N_R^c) are unobserved. Even if an instrument $\mathbf{z}$ that satisfies the exclusion restriction of being uncorrelated with other determinants of the evolution of electoral support $\tilde{\omega}$ is available, it will necessarily be correlated with N_p^c as long as it is correlated with X_p^c. This implies that it is not possible to leave N_D^c and N_R^c in the error term of equation (A.6) if we want to implement an instrumental variables strategy. Instead, notice that equilibrium play implies $q_p^* = \mathbb{E}_\tau\left[\frac{N_p^c(t,t+\tau)+X_p^c(t,t+\tau)}{\tau}\right]$, and thus, we can express each of the (endogenous and unobserved) regressors in equation (A.6) as the equilibrium mixing strategy of the candidate plus sampling noise $\xi_p(t,t+\tau)$ that converges in probability to zero at rate $\sqrt{\tau}$ and is uncorrelated with $\mathbf{z}$:

$$\frac{N_p^c(t,t+\tau)+X_p^c(t,t+\tau)}{\tau} = q_p^*(t,t+\tau) + \frac{1}{\tau}\xi_p(t,t+\tau)$$

Now use our non-parametric estimator for $q_p^*(t,t+\tau)$ from equations (10) and (11), and define $\omega(t,t+\tau) \equiv \tilde{\omega}(t,t+\tau) + \Delta_{cD}^T \frac{1}{\tau}\xi_D(t,t+\tau) + \Delta_{cR}^T \frac{1}{\tau}\xi_R(t,t+\tau)$ as a composite error term that includes all the shocks in the interval and the sampling error to obtain equation (12).

A.3 Proof of Proposition 3

We first use our non-parametric estimator for the unconditional probability of a media report from equation (11), together with the equilibrium mixing strategies for the media in equations (3) and (4) to express the swing voter elasticities to core-targeted statements for each candidate Δ_{cp}^S as functions only of observables and the estimated $\hat{\Delta}_{cp}^T$'s from the estimation of equation (12):

$$\Delta_{cp}^S = \frac{\hat{\Delta}_{cp}^T}{\hat{\phi}_p} - \Delta_{sp}^S \tag{A.7}$$

Using equation (A.7) we can now eliminate the Δ_{cp}^S from equation (A.5) to obtain:

$$v^p(t+\tau) - v^p(t) - \hat{\Delta}_{cp}^T[X_p^c(t,t+\tau) + N_p^c(t,t+\tau)] =$$

$$\left(\Delta_{sp}^S - \frac{\hat{\Delta}_{cp}^T}{\hat{\phi}_p}\right) X_p^c(t,t+\tau) + \left(\frac{\hat{\Delta}_{c\sim p}^T}{\hat{\phi}_{\sim p}} - \Delta_{s\sim p}^S\right) X_{\sim p}^c + \Delta_{sp}^S X_p^s(t,t+\tau) - \Delta_{s\sim p}^S X_p^s(t,t+\tau) + \epsilon^p(t,t+\tau)$$

Grouping terms,

$$[v^p(t+\tau) - v^p(t)] - \hat{\Delta}_{cp}^T[X_p^c(t,t+\tau) + N_p^c(t,t+\tau)] + \hat{\Delta}_{cp}^T \frac{X_p^c(t,t+\tau)}{\hat{\phi}_p} - \hat{\Delta}_{c\sim p}^T \frac{X_{\sim p}^c(t,t+\tau)}{\hat{\phi}_{\sim p}} =$$

$$\Delta_{sp}^S[X_p^c(t,t+\tau) + X_p^s(t,t+\tau)] - \Delta_{s\sim p}^S[X_p^s(t,t+\tau) + X_{\sim p}^c(t,t+\tau)] + \epsilon^p(t,t+\tau)$$

Multiplying and dividing by τ the second and third terms in the left-hand side of this expression, we have that

$$\frac{X_p^c(t,t+\tau) + N_p^c(t,t+\tau)}{\tau}\tau = \frac{\hat{\varphi}_p^c(t,t+\tau)}{\hat{\phi}_p(t,t+\tau)}\tau + \xi_p(t,t+\tau)$$

and

$$\frac{X_p^c(t,t+\tau)}{\hat{\phi}_p}\frac{\tau}{\tau} = \frac{\hat{\varphi}_p^c(t,t+\tau)}{\hat{\phi}_p(t,t+\tau)}\tau$$

so that these terms in the left-hand side cancel. Similarly, multiplying and dividing by τ the fourth term in the left-had side can be re-expressed as

$$\frac{X_{\sim p}^c(t,t+\tau)}{\hat{\phi}_{\sim p}}\frac{\tau}{\tau} = \frac{\hat{\varphi}_{\sim p}^c(t,t+\tau)}{\hat{\phi}_{\sim p}(t,t+\tau)}\tau$$

Now multiply and divide by τ the first and second terms of the right-hand side, to obtain:

$$[v^p(t+\tau) - v^p(t)] - \hat{\Delta}^T_{c\sim p}\frac{\hat{\varphi}^c_{\sim p}(t,t+\tau)}{\hat{\phi}_{\sim p}(t,t+\tau)}\tau =$$

$$\Delta^S_{sp}\hat{\phi}_p(t,t+\tau)\tau - \Delta^S_{s\sim p}\hat{\phi}_{\sim p}(t,t+\tau)\tau + \varpi^p(t,t+\tau) \tag{A.8}$$

where $\varpi^p(t,t+\tau) \equiv \epsilon^p(t,t+\tau) + \hat{\Delta}^T_{cp}\xi_p(t,t+\tau)$. Crucially, notice that the error term in this equation *does not* depend on τ. The left-hand side term in equation (A.8) is defined in Proposition 3 as $\hat{p}(t,t+\tau)$. Because equation (A.8) depends on the same slope parameters and observables for both parties, it is convenient to subtract the equation for candidate D from the equation for candidate R, which directly gives the result of the proposition by defining $\zeta(t,t+\tau) \equiv (1/2)\left[\varpi^D(t,t+\tau) - \varpi^R(t,t+\tau)\right]$.

With estimates of $(\Delta^T_{cD}, \Delta^T_{cR}, \Delta^S_{sD}, \Delta^S_{sR})$ at hand, equation (A.7) uniquely pins down the remaining two elasticities $(\Delta^S_{cD}, \Delta^S_{cR})$.

B Appendix B

News Processing

We followed several steps to process the news article texts. The data collection was conducted in *Lexis Nexis* and *Factiva*.[28] Our search terms included the name of the candidate (e.g., "Alan Kenneth Smith") as well as common abbreviations of the names (examples include "Senator Smith", "Al Smith", "Al K. Smith"). We downloaded all articles which with a succesfull hit for either search criterion.[29] We followed a clean-up procedure before computing our classification scores as follows: first we removed all common English words from the article (before the words are stemmed). Then using the Porter Stemming algorithm, we stemmed the words to their linguistic roots. The benefit of the stemming algorithm is that it allows us to reduce the words to workable roots which eliminate differentiations due to tense or subject.

To reduce the Type-I and Type-II error in the algorithm, we then eliminated articles irrelevant to our setting. In the first pass, after stemming the articles, we searched for candidate names (Here we looked for complete names, excluding any middle names or abbreviations) If the name of the candidate was mentioned in the article, we considered the article to be relevant to our data analysis. If there was no mention of the name in the article, we removed it into a secondary group over which we undertook a secondary search to prevent the unintentional removal of relevant articles.[30] We found our first pass categorizes about 25% of the articles as irrelevant. To reduce the potential for Type-II errors, we conducted a second manual search on the articles that failed the first pass. A research assistant investigated the common reasons for error on articles where a mistake arose, by looking at 10% of all removed articles. We then updated our algorithm to account for these common errors. This second pass reduced the percent of articles removed to 20%.

We carried out our search algorithm for the common words on the set of articles that passed our second test. For each set of candidate articles, after removal of common English words, punctuation, and stemming, we sought for the most commonly used two-word and three-word phrases. Single words may result in a high number of uninformative words and therefore they were not preferred for analysis here (see Gentzkow and Shapiro (2010) for another example of a similar choice).

[28] Due to the limits of search and downloads imposed on us by *Factiva*, we could not rely exclusively on this database.

[29] The article texts themselves are proprietary of these two companies.

[30] For example, a common failure reason in the first pass is a mis-typed character or string (e.g., instead of "Senator Elizabeth", the article would be stored in the newspaper database as "SenatorElizabeth". The missing character can prevent our algorithm from picking up the name of the candidate.

Dropped Senate Races

Candidate Died	Non-bipartisan Races	3-way Races	Unopposed Races	Unopposed Race (in practice)	Other Reason
MN 2002	LA 1990	AK 2010	ID 2004	VA 1990	NE 1988
	VT 2006	LA 1992	SD 2010	AZ 2000	IN 1990
	VT 2012	LA 2002	AR 1990	MA 2002	ND 1992
		CT 2006	GA 1990	MS 2002	TN 1994
		FL 2010	MS 1990	VA 2002	KS 1996
		ME 2012	KS 2002	IN 2006	GA 2000
				AR 2008	MO 2002
					WY 2008
					CO 2010
					DE 2010
					LA 2010
					WV 2010

Table 12: Dropped Senate Races.

Dropped Senate Races

We drop from our analysis some senate races either because they were 3-way races, unopposed races, in practice unopposed races (more than one candidate ran but other candidates were from third parties), not bipartisan races (not a Democrat and a Republican running against each other), or because a candidate died during the race. Table 12 presents a list of races for which data was available, but which we excluded from the analysis for the aforementioned reasons.

Bigram and Trigram Examples

Year	State	Phrase	Score σ_i
1982	CT	Tax cut	0.52
1982	CT	Environment protection	-0.83
1982	CA	Social security	-0.12
1982	CA	Tax increase	0.07
1986	OK	Natural gas	0.90
1986	OK	MPH speed limit	-0.47
1986	ID	Farm policy	0.02
1986	ID	Freedom and safety	-0.50
1990	HI	Death penalty	-0.78
1990	HI	Abortion right	0.65
1990	NC	Gasoline tax cent	-0.37
1990	NC	Depict Homosexual act	0.83
1994	RI	Extended health coverage	0.99
1994	RI	Total spent	0
1994	ND	Canadian wheat import	-0.99
1994	ND	TV violence	-0.99
1998	PA	Public utility commission	-0.53
1998	PA	Sexual relationship Lewinski	0.99
1998	CO	Gil Romero	-0.13
1998	CO	Denver Post	0.34
2002	NJ	Torricelli dropped	-0.19
2002	NJ	Prescription drug	0.32
2002	MA	Traffic safety administration	-0.99
2002	MA	Energy and natural resources	-0.99
2006	MS	USA patriot act	0.99
2006	MS	Dakota Thuner vote	0.99
2006	DE	american dream	-0.08
2006	DE	social security	0.51
2010	MD	Tea party	0.27
2010	MD	Fayetteville Observer	0.29
2010	VT	Law School	-0.99
2010	VT	Human rights	-0.99

Table 13: Examples for Phrases and Scores over Selected Years and States (1982-2010)