

U.S. Newspapers, Like Voters, Are Not Very Ideological*

David A. Beavers[†] António Câmara[‡] James M. Snyder, Jr.[§]

September 2, 2024

Please do not circulate without author permission

Abstract

We study statewide ballot measures across several states. We apply optimal classification, W-NOMINATE, and principal components analysis to a dataset of endorsements ($n = 28,043$) made by newspapers and interest groups, and the votes cast by citizens, to estimate the locations of newspapers, groups, and voters in a multidimensional policy space. We find that newspapers are much less “ideological” than previous research has assumed. We find evidence that, at least in California, newspapers have become more ideological over time, though not more extreme. Next, we study editorial board endorsement articles ($n = 5,210$) using two natural language processing methods: stance detection and document similarity. We find that editorial boards allocate minimal attention to providing opponents’ perspectives and roughly comparable attention to advocating for their preferred position and providing neutral context. Furthermore, we find that newspapers’ ideal points as estimated by optimal classification and W-NOMINATE do a poor job of predicting the one-sidedness of editorials. Finally, we find no evidence that California newspapers mirror expert group’s public communications on domain-relevant propositions.

*Presented at the 120th American Political Science Association Annual Meeting, September 2024. We thank Emily Mao for her significant contributions to data collection. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 2140743. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

[†]Harvard University, Department of Government. dbeavers@g.harvard.edu

[‡]Harvard University, Department of Government. acamara@g.harvard.edu

[§]Harvard University, Department of Government. jsnyder@gov.harvard.edu

1 Introduction

A large body of research assumes that media outlets have clear ideological or partisan biases. Some of this work attempts to measure the size and direction of this bias.¹ Other scholars study the theoretical causes and consequences of this bias.² Yet other work attempts to identify the causes or consequences of this bias empirically.³

In this paper, we show that U.S. newspapers are less “ideological” than previously assumed, in the sense that their behavior is not well-explained by a simple one-dimensional spatial model. We do this by studying newspaper endorsements on statewide ballot measures, comparing them to interest group endorsements and voting outcomes on these same ballot measures. Our analysis predominantly focuses on a set of states with sufficient numbers of ballot propositions and newspapers – Arizona, California, Colorado, Florida, Oregon, and Washington. The time period for our analysis is 1988 to 2022. We find that a one-dimensional model fits the endorsements of most interest groups relatively well, with an average proportional reduction in error (APRE) of 50-70 percent.⁴ For newspapers, the APRE is in the range of 15-25 percent.

Focusing on California, where the data is richest, we document several other interesting patterns. First, the average APRE of right-leaning newspapers – defined as those to the right of the state “median voter” – is significantly higher than the average APRE of left-leaning newspapers. This implies right-leaning newspapers in California more closely resemble the ideological consistency of interest groups than do left-leaning newspapers. Second,

¹See, for example, Groseclose and Milyo 2005, Ansolabehere, Lessem, and Snyder 2006, Baum and Groeling 2008, Ho and Quinn 2008, Gentzkow and Shapiro 2010, Puglisi and Snyder 2011, Larcinese, Puglisi, and Snyder 2011, Gans and Leigh 2012, Lott and Hassett 2014, Puglisi and Snyder 2015b, Kim, Lelkes, and McCrain 2022, and Bernhardt, Dewenter, and Thomas 2023. See Puglisi and Snyder (2015a), Groeling (2013) and Hamborg, Donnay, and Gipp (2019) for reviews of this literature.

²See, for example, Schulz and Weimann 1989, Strömberg 2004, Mullainathan and Shleifer 2005, Baron 2006, Besley and Prat 2006, Gentzkow, Glaeser, and Goldin 2006, Gentzkow and Shapiro 2006, Bernhardt, Krassa, and Polborn 2008, Chan and Suen 2008, Chan and Suen 2009, Duggan and Martinelli 2011, Stone 2011, and Anderson and McLaren 2012. See Gentzkow, Shapiro, and Stone (2015) for a review of this literature.

³See, for example, Druckman and Parkin 2005, DellaVigna and Kaplan 2007, Gerber, Karlan, and Bergan 2009, Ladd and Lenz 2009, Gentzkow and Shapiro 2010, Chiang and Knight 2011, Gentzkow, Shapiro, and Sinkinson 2011, Petrova 2011, Arceneaux, Johnson, and Murphy 2012, Levendusky 2013, Clinton and Enamorado 2014, Eberl, Boomgaarden, and Wagner 2017, Martin and Yurukoglu 2017, Arceneaux, Dunaway, et al. 2020, Garz, Sood, et al. 2020, Song 2020, Levendusky 2022, Garz and Rickardsson 2023. See Dellavigna and Gentzkow (2010) for a review of this literature.

⁴APRE is widely used as a goodness-of-fit measure for binary outcomes. As we discuss below, it is the most appropriate measure for the endorsement data.

newspapers in California have become increasingly ideological over time in the sense that the average APRE for newspapers has increased substantially over the time period we study. Third, on the other hand, newspapers have not become substantially more extreme over time, at least not relative to interest groups.

If newspaper endorsement behavior is not well explained by a unidimensional ideological model, what considerations do drive endorsement behavior? We perform several additional analyses of the endorsement data to investigate this. First, we examine whether newspapers' endorsements merely reflect the preferences of voters in their home counties. Second, we examine whether newspapers have easily measured preferences over a multidimensional issue space. Neither appear to be the case. Third, we consider whether newspapers behave strategically in their endorsement behavior – either by choosing certain positions to signal moderation to readers or by strategically *abstaining* from endorsing on particularly controversial initiatives. Though we are somewhat restricted in our ability to examine these questions due to our limitations in our data, we find little evidence of either type of strategic behavior.

To more fully understand newspaper endorsement behavior – and what non-ideological considerations drive it – we turn to the *text* of endorsements. We collected the text of 5,210 editorial endorsements from NewsLibrary and are in the process of digitizing more than 430,000 pages of newspaper text from Newspapers.com. We employ two natural language processing methods to analyze the NewsLibrary data. First, we employ a stance detection algorithm. We find that editorial boards allocate minimal attention to providing opponents' perspectives, and roughly comparable amount to space to advocating for their preferred position and providing neutral, background context. Newspapers' ideal points as estimated by the optimal classification and W-NOMINATE models perform very poorly at predicting the one-sidedness of editorial endorsements, further showing that these unidimensional ideological scaling methods poorly fit newspaper endorsement behavior. Second, in a more exploratory analysis, we examine semantic similarity between newspapers' endorsement text and the public communications of interest groups on those same texts. We find no evidence that California newspapers mimic domain experts' language.

Before proceeding, we must mention an important caveat to our analysis. Our results are

about newspaper *editorials* rather than news coverage.⁵ Previous work shows that ideological/partisan biases in news sections and editorial pages tend to be correlated, but these correlations are somewhat modest (Larcinese, Puglisi, and Snyder 2011; Puglisi and Snyder 2011; but see also Dalton, Beck, and Huckfeldt 1998). We would guess that newspapers are *more* explicitly ideological on the editorial pages than in the news sections – they are less constrained on the editorial pages and can “wear their ideology on their sleeve” if they so desire – but without further analysis this is just an intuition.

The rest of this paper proceeds as follows. In the next section, we describe our primary data sources and the goodness-of-fit measures we use in the endorsement analysis. In the third section, we detail the results of our optimal classification and W-NOMINATE models. In the fourth section, we detail the results of our text analysis. The final section concludes and discusses implications.

2 Data and Measures

2.1 Endorsement Data

We study the endorsements made by newspapers and interest groups on statewide ballot measures. We collect trichotomous endorsement data – whether a given endorser endorsed yes, no, or abstained on a given ballot proposition – for six states that have both a large number of ballot measures and of newspapers: Arizona, California, Colorado, Florida, Oregon, and Washington.⁶ The period studied is 1988-2022 for California, and 1996-2022 for the other states. We begin with the data set compiled in Puglisi and Snyder (2015b) and add to it by nearly doubling the years covered and filling in many missing cases. Table 1 summarizes the data. In total, we have 28,043 endorsements on 1,105 ballot propositions. For California, we also have county-level election results for all 58 counties for the period

⁵When describing the process underlying their endorsements, every newspaper we have consulted states that the endorsement decisions are made solely by member of the editorial board, and the news reporters are not involved.

⁶The legal process surrounding ballot measures, and by extension the number of ballot measures, varies greatly from state to state. Therefore, not all states have sufficient data to include in our analysis. Eventually, we hope to expand the analysis to other states with sufficient ballot measures and a strong culture of newspapers making editorial endorsements, including Massachusetts, Michigan, and Texas.

	Newspapers	Groups	# Endorsements	# Propositions
Arizona	9	11	1,155	128
California	82	26	17,431	378
Colorado	24	6	1,748	157
Florida	42	14	3,220	99
Oregon	19	16	2,603	188
Washington	18	12	1,886	155
Total	194	85	28,043	1,105

Table 1: Endorsement data summary

1990-2022.⁷ For each state, we employ several criteria to eliminate endorsers and propositions with insufficient data. We describe those criteria more below. The figures in Table 1 represent the sample size *after* eliminating endorsers with insufficient data.

The newspaper endorsements often appear as in Figure 1 below. When newspapers endorse in a given election year they tend to make endorsements on all, or nearly all, ballot measures. This is true for some groups as well, although many groups are more selective because of their particular agendas. Taxpayer groups, for example, generally endorse on all ballot measures that significantly impact taxes, spending, or debt, but often abstain on ballot measures on social issues that have minimal tax implications. We drop groups with agendas that are too narrow and that endorse too selectively, such as Equality California, which endorses on LGBTQ rights issues but little else.

Despite our best efforts over several years, there are holes in the data for a variety of reasons. We are missing some newspapers in some years, or we have only found their endorsements on a few ballot measures. We continue to collect data to fill in these holes where possible.

2.2 Text Data

We supplement the trichotomous endorsement data by collecting the full text of newspaper endorsements. As a first pass, we have collected 5,210 editorial endorsements from NewsLibrary from the same six states above plus several additional states for which we have more limited endorsement data – Massachusetts, Michigan, Missouri, and New Mexico. We are

⁷To treat county-level vote share as comparable to a group or newspaper endorsement, we dichotomize the vote shares to 1 if a majority supported the measure, 0 if a majority opposed the measure, and treat as missing instances in which the county-level vote was a tie. There were only four such occurrences.

The Tribune's endorsements	
STATE PROPOSITIONS:	
Proposition 1A (protect transportation funds)	Yes
Proposition 1B (\$19.9 billion transportation bond)	Yes
Proposition 1C (\$2.8 billion housing bond)	No
Proposition 1D (\$10.4 billion education facility bond)	No
Proposition 1E (\$4 billion levee, flood protection bond)	Yes
Proposition 83 (Sex offender monitoring initiative)	No
Proposition 84 (\$5.4 water, flood control bond)	No
Proposition 85 (parental notification of teen abortion)	No
Proposition 86 (\$2.60 per pack tax on cigarettes)	Yes
Proposition 87 (tax on oil extraction in California)	No
Proposition 88 (\$50 per year parcel tax for education)	No
Proposition 89 (tax on corporations to finance elections)	No
Proposition 90 (restrictions on eminent domain for private property)	No
STATE OFFICES	
U.S. Senator	Dianne Feinstein
Governor	Arnold Schwarzenegger
Lt. Governor	John Garamendi
Secretary of State	Bruce McPherson
Controller	John Chiang
Treasurer	Bill Lockyer
Attorney General	Jerry Brown
Insurance Commissioner	Steve Poizner
Alameda County Superior Court Judge	Sandra K. Bean
LOCAL RACES	
Oakland City Council District 2	Pat Kernighan
Oakland City Auditor	Courtney Ruby
Alameda Mayor	Doug deHaan
Alameda City Council	Frank Matarrese and Lena Tam
Alameda School Board	Tracy Lynn Jensen and Mike McMahon
Berkeley Mayor	Tom Bates
Berkeley City Council	Linda Maio, Dona Spring, Kriss Worthington, Gordon Wozniak
Peralta Community College District	Abel Guillen
AC Transit Board At-Large	Rebecca Kaplan
Assembly District 14	Loni Hancock
Assembly District 16	Sandre R. Swanson
U.S. Rep. District 9	Barbara Lee
LOCAL MEASURES	
Oakland Measure N (library improvement bond)	No
Oakland Measure O (instant runoff elections)	Yes
Oakland Measure M (police, fire retirement fund)	Yes
Berkeley Measure A (school district tax)	Yes

Figure 1: Endorsements by the *Oakland Tribune* in 2006

also in the process of digitizing more than 430,000 page scans from Newspapers.com. This process involves training an n -gram multinomial naive Bayes classifier to identify editorial pages (of which we have identified 73,125) and then applying a series of computer vision methods to extract articles from the pages. Appendix A.1 goes into more detail on this process. Table ?? breaks down the NewsLibrary articles and Newspapers.com pages by state.

2.3 Measures

We apply three different scaling techniques to the newspaper endorsement data: optimal classification, principal components analysis, and W-NOMINATE. Optimal classification is a nonparametric method for maximizing the correct classification of binary choice data (Poole 2005). The W-NOMINATE model is widely used for analyzing legislative roll-call data, though it can be applied more generally to the analysis of binary choice data (Poole and Rosenthal 1997).

Our preferred goodness-of-fit measure is the average proportional reduction in error (APRE). For principal components analysis, we also examine the proportion of variance explained.

	NewsLibrary	Newspapers.com pages	Newspapers.com ed pages
AZ	57	13,937	4,557
CA	1,640	93,281	22,790
CO	664	11,372	2,000
FL	731	99,065	17,819
MA	452	93,328	8,025
MI	410	37,978	5,122
MO	9	-	-
NM	5	-	-
NV	-	29,686	1,978
OR	800	27,814	3,894
WA	442	36,903	7,567
Total	5,210	432,364	73,125

Table 2: Text data summary

We use APRE because it is easy to predict a large share of endorsements correctly with the “naive” model that simply postulates that “everyone take the majority side all the time.” On average, 75.9 percent of endorsements are on the majority side. Newspapers are even more likely than groups to be on the majority side – 81.2 percent for newspapers compared to 72.4 percent for groups. In other words, the naive model would predict 81.2 percent of newspaper endorsements correctly.

For any given model, the APRE asks: How much better does the model do than the “naive” model? It can be written as

$$\begin{aligned}
 APRE &= \frac{Naive\ Errors - Model\ Errors}{Naive\ Errors} \\
 &= \frac{Minority\ Endorsements - Model\ Errors}{Minority\ Endorsements}
 \end{aligned}$$

where *Minority Endorsements* is the number of endorsements on the minority side of each ballot measure. The APRE takes a value of 0 if the model does no better than the naive model and a value of 1 if the model perfectly predicts all choices; APRE can also be negative in instances in which the naive model outperforms the given model.

As we will see below, as an empirical matter there is a strong relationship between the

estimated ideological positions of endorsers and their corresponding APREs. In particular, endorsers with more moderate positions have lower APREs. Importantly, as a theoretical matter this is not a “mechanical” relationship. To see this, consider the following. Suppose the endorsers’ preferences are single-peaked (and can be described by utility functions with slopes that are “similar enough”) or their preferences are order-restricted. If there is perfect information and endorsers always sincerely endorse their preferred alternative on all ballot measures, then we will be able to place all endorsers relative to one another and perfectly account for all endorsements using a one-dimensional model.⁸ In this case, the APRE will be 1 for all endorsers. Here is a simple example:

	Endorser						
	1	2	3	4	5	6	7
Measure A	Y	N	N	N	N	N	N
Measure B	Y	Y	N	N	N	N	N
Measure C	Y	Y	Y	N	N	N	N
Measure D	Y	Y	Y	Y	N	N	N
Measure E	Y	Y	Y	Y	Y	N	N
Measure F	Y	Y	Y	Y	Y	Y	N

Suppose we order the endorsers (1, 2, 3, 4, 5, 6, 7). Suppose further that we optimally choose the “cut-point” dividing the Y’s from the N’s on each ballot measure – i.e., between endorsers 1 and 2 on Measure A, between endorsers 2 and 3 on Measure B, and so on. Then, we will correctly predict 100% of the endorsements, and the APRE for each endorser will be 1 – both for endorsers who are ideologically “extreme” and those who are ideologically “moderate.”

3 Results

3.1 Are Newspapers Ideological?

We begin with examining the results from optimal classification, principal components analysis, and W-NOMINATE. In the analyses, we restrict the sample as follows. For the data from California, 1988-2022, we drop endorsements from years in which a group/newspaper endorsed on fewer than 80 percent of proportions. For the remaining states, with data for

⁸Roberts (1977) uses the term “hierarchical adherence” to describe the order-restriction condition.

	Optimal Classification			W-NOMINATE		
	Newspapers	Alt Press	Groups	Newspapers	Alt Press	Groups
AZ	16.1	–	65.9	13.3	–	61.4
CA	21.7	47.8	49.1	15.5	37.3	49.8
CO	14.8	82.9	75.0	17.9	74.3	55.0
FL	17.3	43.8	44.9	14.2	50.0	21.1
OR	38.9	59.3	80.7	13.0	7.4	55.1
WA	32.7	87.5	82.5	28.0	87.5	76.3
Avg	23.6	64.3	66.4	17.0	51.3	53.1

Table 3: APRE of one-dimensional model

1996-2022, we employ a threshold between 50 percent and 80 percent, depending on the particular state.⁹ Next, we keep only groups/newspapers with at least 30 endorsements remaining in California or 20 endorsements remaining in the other states. Finally, we drop propositions with fewer than 10 endorsements and for which the minority position represents less than 2.5 percent of endorsements. These thresholds are arbitrary, but the results are largely robust to various choices of thresholds.

The results are in Table 3, with APRE rescaled to have an upper bound of 100. There are several takeaways. First, groups that endorse are relatively well-explained by a one-dimensional model, in the sense that the APRE is reasonably high (average across states is 66.4 for optimal classification and 53.1 for W-NOMINATE). For newspapers, the APRE is significantly lower. For all six states, and for both optimal classification and W-NOMINATE, the average APRE for newspapers is lower than the average APRE for groups, often by a wide margin. The alternative press – tabloid-style publications like the *San Francisco Bay Guardian* and the *East Bay Express* – fall somewhere in between the traditional press and interest groups, but much closer to the latter.

The interpretation is largely the same if we instead look at a one-dimensional principal factor analysis. Table 4 displays the percentage of variance explained separately by state for newspapers, the alternative press, and groups. As was the case with optimal classification and W-NOMINATE, the one-dimensional principal factor analysis does a poor job of explaining the variance in newspapers’ endorsement behavior relative to groups.

⁹We use a threshold of 80 percent for Florida, 70 percent for Colorado, 60 percent for Oregon and Washington, and 50 percent for Arizona. We lower the thresholds to keep more endorsements for states in which data is sparser.

	Percent of Variance Explained		
	Newspapers	Alt Press	Groups
AZ	17.0	–	42.1
CA	5.4	22.4	21.6
CO	15.9	5.5	35.6
FL	9.5	4.9	12.1
OR	6.1	15.0	22.4
WA	10.7	41.9	29.0
Avg	10.8	17.9	29.0

Table 4: One-dimensional principal factor analysis

Figure 2 visualizes the relationship between endorsers’ optimal classification rank and APRE for California newspapers, groups, statewide voters, and voters from select counties.¹⁰¹¹ As we already saw in the table above, newspapers tend to have lower APREs than groups. Newspapers (in blue) are relatively clustered in the center of the x-axis; there are no newspapers as extreme, according to their optimal classification ranking, as the most extreme liberal or conservative groups. By contrast, groups are clustered at the extremes (with the notable exception of the California Chamber of Commerce). There is, however, an asymmetry in newspaper ideology, with the most conservative newspapers (higher optimal classification ranking) both located somewhat farther from the median optimal classification ranking and much better explained by the one-dimensional model (have higher APREs) than do the most liberal newspapers. The alternative press is concentrated near the far-left side of the ideological cut points and has relatively high APREs. Given the asymmetry in the mainstream press, this highlights how the liberal alternative press may develop where its proprietors perceive a need for an ideological counterbalance. We also see from Figure 2 that there is a clear relationship between endorsers’ ideology (their value on the x-axis) and endorsers’ APRE (their value on the y-axis), though this relationship is not one-to-one. This can be seen clearly if one glances at the grad line at 90 on the x-axis. This line roughly intersects the *Redding Record Searchlight*, with a weak APRE of 10.3, the *Long Beach Press Telegram*, with a modest APRE of 46.2, and the *Inland Valley Daily Bulletin*, which has

¹⁰We chose several counties – Los Angeles, San Francisco, Modoc, Shasta, Kern, orange, Contra Costa, and Alameda – to represent a range of liberal, moderate, and conservative counties and urban, suburban, and rural counties.

¹¹Several newspapers had negative APREs, meaning the optimal classification model performed *worse* than the naive model of guessing the newspaper always sided with the majority. For these papers – *Visalia Times-Delta* (-7.7), *Gilroy Dispatch* (-13.0), and *Kingsburg Recorder* (-25.0) – we collapsed their APREs closer to 0 (by dividing them by 10) for plotting purposes.

one of the highest APRE values of any newspaper in the sample – 77.8. There is clearly not a one-to-one relationship between endorser ideology and APRE.¹²

3.2 Ideological Shifts Over Time

For the analysis above, we pooled all endorsements from 1988-2022 (for California) or 1996-2022 (for the other states). This period, however, saw numerous large changes in the newspaper industry, including massive layoffs and newspaper closures, increasing concentration in media ownership, the digitization of news, the growth of social media, and the loss of advertising revenues to technology giants (e.g. Baker 2007; Seamans and Zhu 2014; Darr, Hitt, and Dunaway 2018; Peterson 2021; Peterson and Dunaway 2023). Given these tectonic shifts, is it possible that our pooled analysis is masking changes in newspaper ideology over time? To examine this possibility, we rerun the above analyses, breaking the data up into eight overlapping 6-year periods (i.e. 1988-1994, 1992-1996, etc.) to examine whether newspapers' APREs increase over time. We focus just on California for this analysis, where the data is richest.

Insofar as we observe changes in the average APRE of newspapers and groups over time, we want to be sure this is due to actual changes in endorser behavior and not due to changes in the composition of endorsers included in each time period. Therefore, we must apply different thresholds for inclusion in the analysis than those used in the pooled analysis. We limit the sample to those newspapers and groups that have made sufficient endorsements (at least 15) in at least 6 of the 8 time periods. This results in a sample of 43 newspapers and 26 groups. The individual analyses for each time period are therefore made up of the subsample of those 69 endorsers with 15 or more endorsements in that period. As above, these thresholds are arbitrary, but results are largely robust to various choices. Also as above, we drop all propositions in each time period on which there are fewer than 10 endorsements and for which the minority endorsers represent less than 2.5 percent of endorsers.

¹²One possible concern when interpreting these results is that if the cut-points all cluster in one location, then (in small samples) the positions of some endorsers might be poorly estimated, and the corresponding APREs large. Fortunately, the distribution of cut-points is relatively uniform and covers the full spectrum of endorser rankings, with a modest leftward skew.

The results are presented in Figure 3. The left panel presents results from the optimal classification model; the right panel presents results from the W-NOMINATE model. The black lines represent newspapers; the gray lines represent groups.

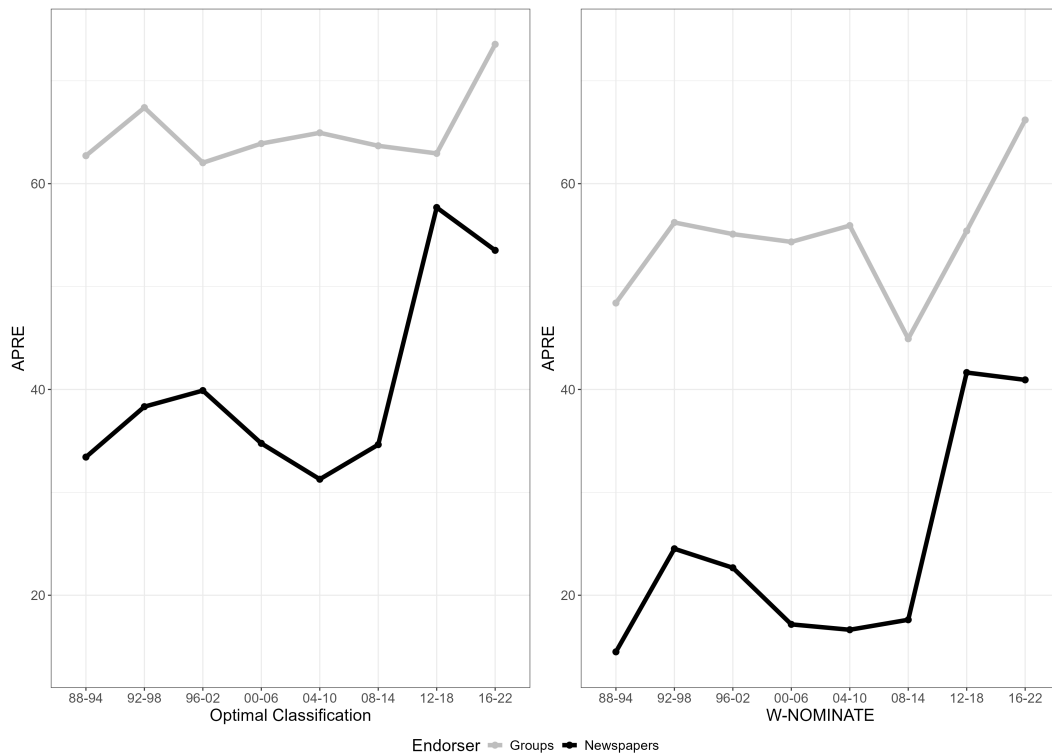


Figure 3: APRE by period for California, optimal classification model on left, W-NOMINATE model on right (newspapers in black, groups in gray)

Newspapers’ average APRE has increased since 1988 – with nearly all of the increase coming since 2008. While groups’ average APRE has also increased since 2008, it has done so to a lesser degree than newspapers, so the gap between groups’ and newspapers’ average APRE has diminished over time.

Endorsements by California newspapers – and, to a lesser degree, groups – are better explained by a one-dimensional model over time, as evidenced by their increasing APRE. But have newspapers become *more extreme* in their endorsement behavior relative to groups over time? As Figure 4 shows, this is not the case. We define each endorser’s “extremism” in each period as the absolute distance between the endorser’s estimated position and the estimated position of California voters.¹³ Figure 4 shows how the extremism of newspapers (in black) and groups (in gray) has evolved over time using the optimal classification model

¹³Since the total number of endorsers varies by period, the optimal classification rankings are normalized each year to run from 1 to 100.

in the left panel and the W-NOMINATE model in the right panel. For each overlapping time period, newspapers are less extreme than groups. The gap between the two, if anything, widens in the middle periods, before shrinking again somewhat in more recent years. There is no evidence that newspapers' increasing APREs are due to a move toward the ideological extremes, relative to state voters.

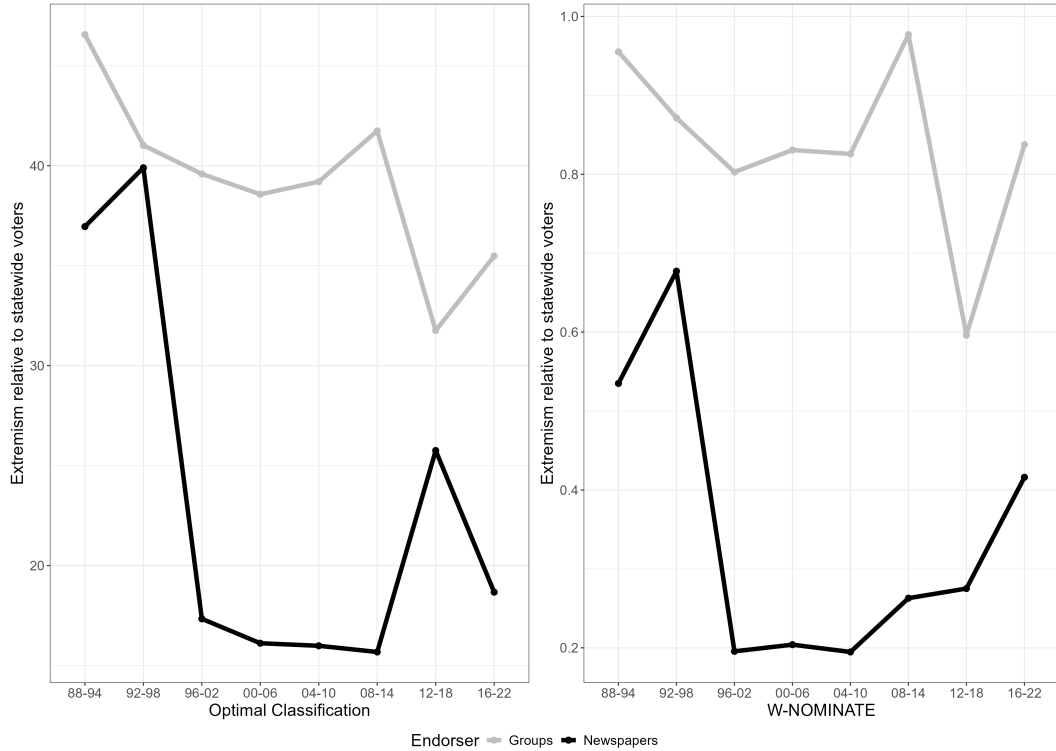


Figure 4: Extremism by period for California, optimal classification model on left, W-NOMINATE model on right (newspapers in black, groups in gray)

3.3 Supplementary Analyses

In this subsection, we briefly detail the results of several supplementary analyses. To begin, we consider the possibility that newspapers' endorsement behavior merely reflects the preferences of voters in their circulation market. Using their text-based measure of media slant, Gentzkow and Shapiro (2010) find that newspapers' observed slants are close to their profit-maximizing slant, suggesting a demand-side effect of *news* (rather than editorial) coverage. To examine this relationship, we include all 58 counties' dichotomized vote choice in our optimal classification and W-NOMINATE analyses to place California voters on the same unidimensional scale as newspapers and groups, and then examine the correlations between

newspapers' ideological positions and the positions of voters in those newspapers' home counties.¹⁴ There is a positive but modest correlation – 0.27 in the optimal classification model and 0.32 in the W-NOMINATE model. The correlation with home-county voters is somewhat higher for the alt press than for mainstream newspapers. These correlations suggest that newspapers are responsive to the demands of their target audience, but that these demand-side effects cannot completely explain newspaper editorial board behavior.

Next, we consider whether newspapers have easily measured preferences over a *multidimensional* issue space. This does not seem to be the case. Neither the second dimension of a two-dimensional NOMINATE model nor a many-dimensional principal components analysis reveal easily recognizable issue dimensions. We also examine bundles of related ballot measures separately. Across both the optimal classification and W-NOMINATE models and for eight issue bundles considered, we find that the average APRE for newspapers is always noticeably smaller than the average APRE for groups. Appendix Appendix A.2 goes into more detail.

Finally, we consider the possibility that newspapers behave *strategically* in making endorsements – either by abstaining on particularly cross-cutting or controversial issues or by trying to signal moderation or bipartisanship to voters to make their endorsements appear more credible. As a hypothetical example of the latter, a liberal-leaning newspaper might endorse the conservative alternative on a relatively insignificant ballot measure or on a ballot measure that is expected to pass or fail by a large margin. We check for strategic endorsing behavior in two ways – first, by examining endorsements on close votes; and second, by examining endorsing behavior as a function of total spending on each ballot proposition campaign. Appendix A.3 provides the details. In neither case do we find evidence of strategic endorsing.

It is more difficult for us to examine strategic *abstention* given the limitations of our data, because we cannot be completely confident that the absence of an endorsement in our

¹⁴County is an imperfect proxy for a newspaper's market. There are some papers for which the overlap is likely high, either because the paper purports to cover the county (e.g. the *Orange County Register*) or the paper is headquartered in a city that contains a significant portion of the county's population (e.g. the *Chico News and Review*). But there are others, like the many smaller papers in massive Los Angeles County, for which the market is a fraction of the county, or major dailies like the *Los Angeles Times* whose market extends well into neighboring counties.

dataset constitutes neutrality by a newspaper rather than a missing entry on our part. But there are reasons to believe this type of strategic abstention – insofar as it exists – would likely make newspapers appear *more* extreme than they really are. Since strategic abstention would most likely occur on the types of issues for which a potentially reliably right- or left-of-center newspaper may be “conflicted” and would frequently endorse in the direction opposite to its normal slant, by declining to make such an endorsement, it would make the newspaper appear more rather than less ideologically consistent. In this case, a unidimensional ideological model would better explain newspaper behavior.

4 Text Analysis

The results of our optimal classification and W-NOMINATE models indicate that newspapers – as judged by their endorsements on statewide ballot propositions – are less ideological than previous research has generally assumed. But what, if not ideological considerations, drives newspaper endorsement behavior? We have largely exhausted the insights we can glean from the zeros and ones of the endorsement dataset itself. To better understand newspaper “ideology,” we next turn to analyzing the *text* of editorial endorsements. To do so, we employ two natural language processing methods: stance detection and document similarity.

4.1 Stance Detection

When editorial boards publish articles endorsing a vote on a ballot proposition, how much space do they allocate in the article to each side of the debate? Do endorsement articles present only arguments in favor of their preferred position, or do they counterbalance their position by offering opponents’ perspectives? And does the ratio of the two vary based on the ideological scores and APREs of the newspapers we estimated above?

To answer these questions, we employ stance detection methods to our corpus of endorsement articles from NewsLibrary. Stance detection is a sentiment analysis task that seeks to classify an author’s position on a stance object, or target (Mohammad et al. 2016; Best-

vater and Monroe 2023). A stance is a type of “public communication in which the speaker simultaneously evaluates an object, takes a position in relation to the sociocultural context, and aligns oneself with like-minded actors” (Du Bois 2007). A target is any entity or issue that can be expressed as a short phrase (e.g., `Donald Trump, states’ rights`). Formally, given a document d and a target t , we classify the stance taken by the author in d as either `PRO` towards t , `CON` towards t , or `NEUTRAL` towards t .¹⁵ In the present case, the documents d are editorial board endorsement articles and the targets t are statewide ballot propositions extracted from the articles using an algorithm described below.

We implement an information extraction and sentiment analysis pipeline to (a) automatically detect which ballot propositions are discussed in the articles and (b) quantify the proportion of each article that is `PRO`, `CON`, or `NEUTRAL` towards each detected proposition. First, we implement a rule-based algorithm that uses regular expression pattern matching and part of speech tagging to extract every statewide ballot proposition referenced in a given article. Second, we fine-tune a neural network to perform stance detection on each article, with respect to each ballot proposition detected (many endorsement articles reference multiple ballot propositions).

To simulate a human-annotated stance detection dataset, we first apply our information-extraction algorithm to discover each ballot proposition mentioned in each article. We then subset to all articles for which the newspaper’s known endorsement for a ballot proposition matches the label generated by instructing `GPT-4o` to classify the article’s stance on that ballot proposition.¹⁶ We then compare `GPT-4o` to a human annotator on a small subset of the articles. We find human and machine labels match on nearly all of the articles.¹⁷ We repeat the process at the paragraph level for that subset of articles, prompting `GPT-4o` to generate a stance label and measuring agreement with human annotators.¹⁸ We find human and machine labels match on just over three quarters of the paragraphs.¹⁹ Note that we removed neutral articles for the article-level dataset.²⁰

¹⁵In the natural language processing literature, `FAVOR`, `AGAINST`, `NONE` are also used to refer to the class set.

¹⁶We use the following prompt: `Please classify the article as endorsing the PRO, CON, or NEUTRAL position on the ballot initiative: x`, where we substitute x with the ballot proposition.

¹⁷96% agreement, Cohen’s κ : .92

¹⁸We use the following prompt: `Please classify the paragraph as endorsing the PRO, CON, or NEUTRAL position on the ballot initiative: x`, where we substitute x with the ballot proposition.

¹⁹76% agreement, Cohen’s κ : .53

²⁰Some editorial boards write an “endorsement” on propositions on which the board could not agree, or

We next fine-tune a BERT model, a pre-trained transformer-based language model that represents text in a low-dimensional vector space that encodes semantic and world knowledge acquired from initial training on a massive corpus (Devlin 2018). Each example in our training data consists of two strings, a document and a target, and our stance label, PRO, CON, or NEUTRAL. Appendix A.4 provides more details about the dataset and methodology used to train the model. We find that our best-performing model achieves near-human performance on the unseen *test* dataset, with accuracy and F1 scores above 90% (see Table 5), exceeding state-of-the-art performance on similar stance detection tasks and datasets (Allaway and McKeown 2020; Liu et al. 2022). We then use our model to predict stance for every paragraph-ballot proposition pair in our dataset and to calculate several sets of outcomes.

Metric	Train	Development	Test
Accuracy	0.93	0.92	0.92
F1 Macro	0.92	0.91	0.90
F1 Weighted	0.93	0.92	0.92

Table 5: Stance detection model performance

Our outcomes at the newspaper i by proposition j level are: (1) the proportion of words or paragraphs classified as supporting, opposing, or exhibiting no stance by newspaper i in its associated article towards the extracted ballot proposition j ; (2) those same proportions where paragraphs are inverse-weighted²¹ to account for reader drop-off. Since we know the official recommendation of the editorial boards on any given ballot proposition, we also calculate (3) the proportion of words or paragraphs classified as supporting the same side, opposite side, and neither side of the official recommendation. Using (3), we calculate (4) normalized one-sidedness, the extent to which an article allocates more text supporting the newspaper’s official recommendation compared to opposing views. This measure takes values between 0 and 1 as is given by equation 1

$$\text{Normalized One-Sidedness (Words)} = \frac{\left(\frac{\text{prop}_{\text{same}} - \text{prop}_{\text{opposite}}}{\text{prop}_{\text{same}} + \text{prop}_{\text{opposite}}}\right) + 1}{2} \quad (1)$$

conclude some articles with calls to vote one’s conscience.

²¹Inverse-weighting downweights paragraphs farther from the top of the article to simulate readers failing to read the entire article.

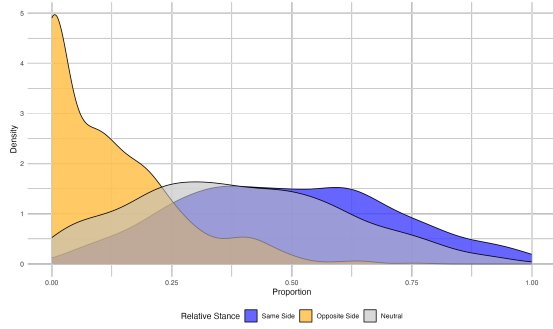
where $\text{prop}_{\text{same}}$ is the proportion of words in the article text supporting the same side as the actual recommendation and $\text{prop}_{\text{opposite}}$ is the proportion of words in the article text supporting the opposite side as the actual recommendation.

In the analyses that follow, we limit our sample of editorial endorsement articles to those articles that mention only one proposition ($n_{\text{all}} = 1,035$, $n_{\text{CA}} = 814$). We do so because including articles that endorse on multiple propositions artificially inflates the proportion of neutral text, since any text that pertains to an endorsement on a given proposition is typically neutral toward other propositions. Beyond relatively small changes to the proportion of text classified as neutral, the interpretation of the results is largely identical when analysis is run on the full corpus instead.

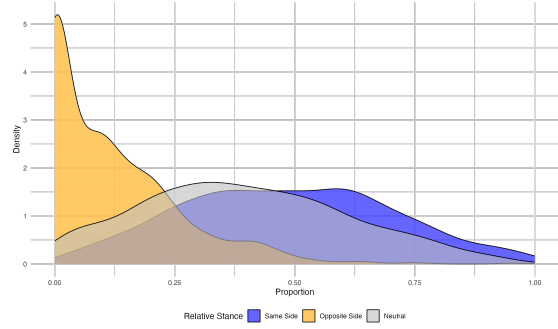
We now turn to the results. Figure 5 displays the kernel density estimate of our outcomes (3) for our sample, the proportion of words or paragraphs classified as supporting the same side (blue), opposite side (yellow), and neither side (gray) of the official recommendation.²² The left two panels are for California; the right two panels are for all states pooled together. The top two panels use words as the unit of analysis; the bottom two panels use paragraphs. We do not display the inverse-weighted figures here because they are so similar to the non-weighted figures. The distributions show that the bulk of endorsement articles dedicate scant attention to counterbalancing their own endorsed positions by airing opponents’ perspectives. Moreover, almost no articles dedicate their text completely to advocacy of their chosen side; the distributions of same side and neutral text are similar. These figures suggest that editorial endorsement articles, taken together, dedicate equal parts to background information and advocacy for a chosen position, with perhaps a sentence or two to share opponents’ perspectives.

Do newspapers at different points along the “ideological” spectrum, as estimated from our optimal classification and W-NOMINATE models, dedicate different amounts of space to the same side or opposite side of a debate in their editorial board endorsement articles? Do newspapers we classify as “moderate” based on their endorsement behavior also appear more balanced in their writing? If newspaper ideology were easily captured on a unidimensional ideological scale, we would expect that newspapers at the ideological extremes would

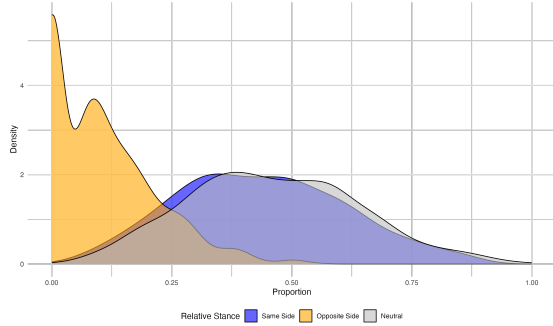
²²A kernel density estimate is akin to a smoothed histogram for continuous variables.



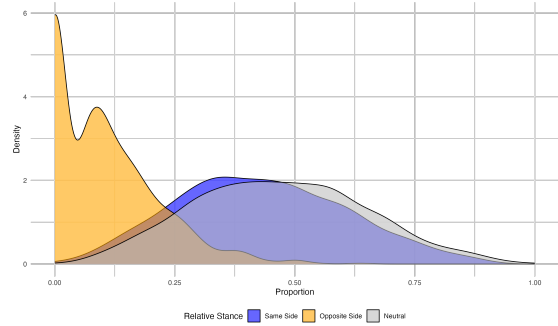
(a) California by words



(b) All states by words



(c) California by paragraphs



(d) All states by paragraphs

Figure 5: Relative distribution of stance in editorial board endorsement articles towards recommended ballot propositions

dedicate more space to advocating their chosen viewpoint and less space to acknowledging the opposing viewpoint than would newspapers placed closer to the state median voter.

To test this, we fit separate regression models for each combination of three dependent variables from our stance detection measures and five independent variables from our ideal point estimations. Our independent variables are first-dimension W-NOMINATE ideal point; optimal classification rank; W-NOMINATE ideal point extremism (absolute distance between the endorser’s estimated position and the estimated position of the endorser’s state’s voters); optimal classification ideal point extremism; and estimated probability of recommending the observed position, estimated from the W-NOMINATE model. Our dependent variables are same-side article proportion, opposite-side article proportion, and normalized one-sidedness (see equation 1), all at the word level, as described by stance detection outcome measures (3) and (4). We estimate two models for each combination – a linear probability model using Huber-White (HC1) robust standard errors and a beta regression model. We then consider the R^2 to understand to what extent the stance detection and ideal point variables are correlated. If the stance detection and ideal point variables are uncorrelated, the estimated

ideology of a newspaper has no linear predictive power for the one-sidedness of editorial board endorsement articles that newspaper publishes, and we would observe a small R^2 . Conversely, if these variables are highly correlated, the newspaper’s ideology would be a strong linear predictor of article one-sidedness, and the R^2 will be high.

We report our findings from the linear probability model. Our findings from the Beta regression model are qualitatively similar and reported in Appendix A.5. Figure 6 summarizes the results. We find that newspapers’ ideal points – across the five ways we estimate them – are minimally predictive of the stance proportions of editorial articles. The highest R^2 value from the 30 regression models is 0.008. This is another strike against the ideal points and further evidence that newspaper endorsement behavior is poorly captured on a unidimensional ideological scale.

4.2 Group Similarity

If newspaper editorial boards are not behaving ideologically when publishing proposition endorsements, what else might explain their behavior? One possibility is that newspapers, much as voters, follow cues from experts. If editorial boards follow the lead of issue-specific experts when making endorsements, we would expect that newspapers’ endorsements would resemble public communications from experts about those same propositions.

To examine this possibility, we are in the process of gathering texts from interest groups and other issue-relevant experts on the same ballot propositions. Because there are no centralized databases of such text (to our knowledge) – as there are for newspapers – this process is manual and time-intensive, so the results here are best seen as exploratory as we attempt to gather a larger corpus of group text. Thus far, we have sufficient data to match “expert” text from two California groups – the California Labor Federation (CLF) and the California Federation of Teachers (CFT) – to 245 editorial board endorsement articles on 21 ballot propositions across 28 newspapers. Insofar as newspapers follow experts’ lead, we would expect the greater similarity between newspapers’ text and groups’ text on propositions that fall within those groups’ realm of expertise – namely, labor propositions for CLF and education propositions for CFT.

Same Side Article Proportion (Words)	0.006	0.004	0.000	0.006	0.002
Opposite Side Article Proportion (Words)	0.003	0.005	0.002	0.001	0.006
Normalized One-Sidedness (Words)	0.005	0.000	0.000	0.005	0.004

(a) California

Same Side Article Proportion (Words)	0.002	0.001	0.000	0.006	0.005
Opposite Side Article Proportion (Words)	0.003	0.001	0.001	0.002	0.003
Normalized One-Sidedness (Words)	0.002	0.002	0.000	0.008	0.006
	Est. Prob. of Recommendation (NOMINATE)	Ideal Point Extremism (NOMINATE)	Ideal Point Extremism (OC)	NOMINATE 1st Dimension	OC Rank

(b) All States

Figure 6: Heatmaps of R-squared values for linear probability models. Outcomes are estimated from stance detection, predictors from ideal points.

We use two measures of document similarity: `tfidf`, which measures actual word and phrase overlap, and `sbert`, which measures semantic similarity in a low-dimensional latent space. `tfidf` calculates the averaged cosine similarity between TF-IDF-weighted n -gram vectors of the text from the articles written by the editorial board of a newspaper and the group’s endorsement for a given ballot proposition, as extracted by our algorithm.²³ This method emphasizes relevant terms, effectively capturing importance-weighted word and phrase overlap. `sbert` calculates the averaged cosine similarity between low-dimensional representations of the text from the articles written by the editorial board of a newspaper and the group’s endorsement for a given ballot proposition, as extracted by our algorithm. These embeddings are generated using `Sentence-BERT`, a model designed to create semantically meaningful document representations that can be directly compared using cosine similarity, allowing us to quantify the semantic relatedness of the documents (Reimers 2019). This method captures nuanced semantic relationships without relying on direct word overlap, potentially identifying similarities that might be missed by `tfidf`.

Figure 7 displays the results of the `tfidf` model in the top panel and the `sbert` model in the bottom panel. Higher values on the x-axis signify greater semantic similarity between newspaper and group text. These values (especially for `sbert`) appear high, though we caution against drawing too many implications from this; because the group and newspaper text are on the same propositions, the language used is bound to have a significant amount of overlap. Of more interest is the relative distributions of the semantic similar on all propositions (blue) and on domain-specific propositions (education-related propositions in green, labor-related propositions in green). If newspapers were following the expertise of these two groups, we would expect the red/green boxes to be to the right of the blue boxes – that is, we would expect newspapers to exhibit greater semantic similarity to group text on issues in which those groups have credible claims to expertise. In three of four cases (all but CLF on the `tfidf` model), we find the opposite. Overall, there is little evidence from this exploratory analysis that California newspapers follow expert-group leadership when crafting their endorsements. It is possible that they follow those groups in deciding *what position* to endorse, but we do not find evidence that they follow those groups in *how to*

²³`tfidf` uses unigrams, bigrams, and trigrams. `tfidf` also removes common English words, ignores very frequent and very rare terms, and limits the vocabulary to the 10,000 features.

write their endorsement article.

5 Conclusion

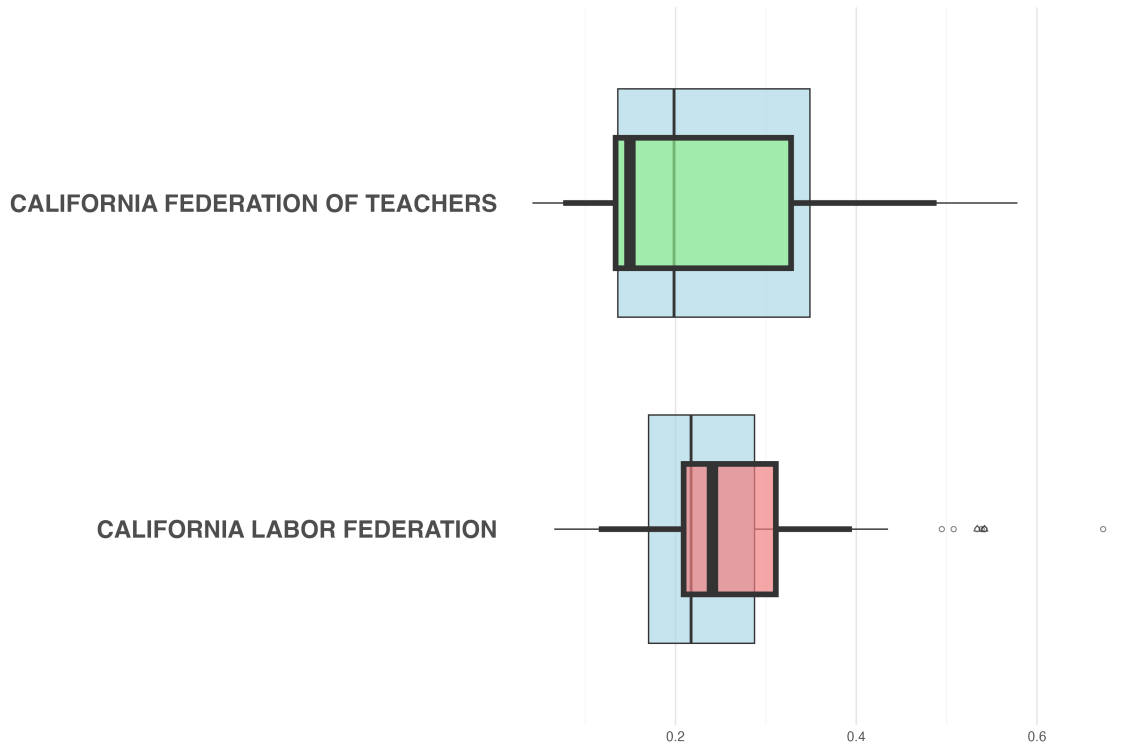
That newspaper ideology can be readily collapsed onto a unidimensional left-right ideological spectrum is largely taken for granted. In this paper, we study newspapers' and groups' endorsements on statewide ballot propositions – and voters' electoral choices on those same propositions – and find that this assumption is a poor one. Newspapers are not well-explained by a simple one-dimensional spatial model.

We leave the reader with two possible implications. The first is that we should expect the influence newspapers have on voting behavior, if any, to be more nuanced than what we would predict when modeling newspaper content as a flow of liberal/conservative messages. The same newspaper might shift voters to the left on some issues, but to the right on others. Insofar as our findings about (the lack of) newspaper ideology would extrapolate from the editorial page to the news page, we might go one step further and speculate that one reason the typical American voter is not ideological is because the news they read every day is not ideological in the narrow liberal-conservative sense of the word.²⁴

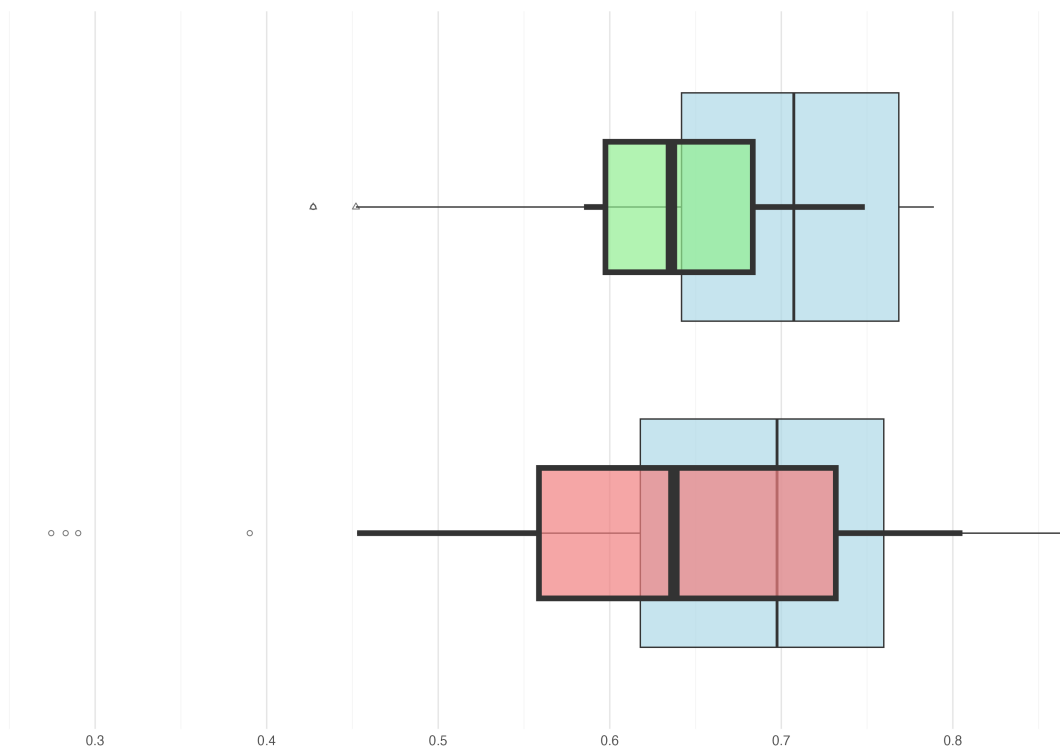
A second, perhaps even more sweeping, interpretation stems from considering newspaper editors as sophisticated American voters. Since Converse (1964), it has become commonplace to assert that most American voters are “unsophisticated” and lack “constraint” in their policy attitudes.²⁵ Constraint in ideological terms – high correlations between attitudes across diverse policy areas – implies that a one-dimensional spatial model would fit well the attitudes of the politically sophisticated. But the policy attitudes of newspaper editorial boards – who we would almost certainly say are made up of the highly politically sophisticated – are *not* well-explained by a simple, one-dimensional, ideological spatial model (at least as revealed by their endorsements). Thus, we cannot conclude that the lack of “constraint” is equivalent to a lack of thought, as much of the literature seems to imply. The study of newspaper editors reminds us that ideological constraint is not necessarily a good

²⁴Gans 1979, among other scholars, discusses the news as ideological in the sense that it represents a set of “enduring values” including ethnocentrism, small-town pastoralism, and individualism.

²⁵In fact, Converse and others argue that many American voters even lack policy attitudes.



(a) TF-IDF



(b) SBERT

Figure 7: Distribution of editorial board endorsement article average similarity scores with group. Blue boxes are all propositions, green boxes are education-related propositions, and red boxes are labor-related propositions.

thing in and of itself. Ideology is at best a shortcut for deep, careful analysis and reasoning. Of course, we cannot conclude that the lack of constraint in the policy attitudes of ordinary citizens means that they are as sophisticated as newspaper editors. It does imply that scholars should emphasize other evidence – such as incoherent arguments in open-ended survey questions – when they want to argue that citizens are poorly equipped to engage with politics.

References

- Allaway, Emily and Kathleen McKeown (2020). “Zero-shot stance detection: A dataset and model using generalized topic representations”. In: *arXiv preprint arXiv:2010.03640*.
- Anderson, Simon P. and John McLaren (2012). “Media Mergers and Media Bias With Rational Consumers”. In: *Journal of the European Economic Association* 10.4, pp. 831–859.
- Ansolabehere, Stephen, Rebecca Lessem, and James M. Snyder Jr. (2006). “The Orientation of Newspaper Endorsements in U.S. Elections, 1940–2002”. In: *Quarterly Journal of Political Science* 1.4, pp. 393–404.
- Arceneaux, Kevin, Johanna Dunaway, et al. (2020). “Strategic Candidate Entry and Congressional Elections in the Era of Fox News”. In: *American Journal of Political Science* 64.2, pp. 398–415.
- Arceneaux, Kevin, Martin Johnson, and Chad Murphy (2012). “Polarized Political Communication, Oppositional Media Hostility, and Selective Exposure”. In: *The Journal of Politics* 74.1, pp. 174–186.
- Baker, C. Edwin (2007). *Media Concentration and Democracy: Why Ownership Matters*. Cambridge University Press.
- Baron, David P. (2006). “Persistent Media Bias”. In: *Journal of Public Economics* 90, pp. 1–36.
- Baum, Matthew A. and Tim Groeling (2008). “New Media and the Polarization of American Political Discourse”. In: *Political Communication* 25.4, pp. 345–365.
- Bernhardt, Dan, Stefan Krasa, and Mattias Polborn (2008). “Political Polarization and the Electoral Effects of Media Bias”. In: *Journal of Public Economics* 92.5-6.

- Bernhardt, Lea, Ralf Dewenter, and Tobias Thomas (Jan. 2023). “Measuring Partisan Media Bias in US Newscasts from 2001 to 2012”. In: *European Journal of Political Economy*, p. 102360.
- Besley, Timothy and Andrea Prat (2006). “Handcuffs For the Grabbing Hand? Media Capture and Government Accountability”. In: *American Economic Review* 96.3, pp. 720–736.
- Bestvater, Samuel E and Burt L Monroe (2023). “Sentiment is not stance: Target-aware opinion classification for political text analysis”. In: *Political Analysis* 31.2, pp. 235–256.
- Chan, Jimmy and Wing Suen (2008). “A Spatial Theory of News Consumption and Electoral Competition”. In: *The Review of Economic Studies* 75.3.
- (2009). “Media As Watchdogs: The Role of News Media in Electoral Competition”. In: *European Economic Review* 53.7, pp. 799–814.
- Chiang, Chun-Fang and Brian Knight (2011). “Media Bias and Influence: Evidence from Newspaper Endorsements”. In: *The Review of Economic Studies* 78.3, pp. 795–820.
- Clinton, Joshua D. and Ted Enamorado (2014). “The National News Media’s Effect on Congress: How Fox News Affected Elites in Congress”. In: *The Journal of Politics* 76.4, pp. 928–943.
- Converse, Philip (1964). “The Nature of Belief Systems in Mass Publics”. In: *Ideology and Discontent*. Ed. by David E. Apter. New York: Free Press.
- Dalton, Russel J., Paul A. Beck, and Robert Huckfeldt (1998). “Partisan Cues and the Media: Information Flows in the 1992 Presidential Election”. In: *American Political Science Review* 92.1, pp. 111–126.
- Darr, Joshua P., Matthew P. Hitt, and Johanna L. Dunaway (Dec. 2018). “Newspaper Closures Polarize Voting Behavior”. In: *Journal of Communication* 68 (6), pp. 1007–1028. ISSN: 14602466. DOI: 10.1093/joc/jqy051.

- Dellavigna, Stefano and Matthew Gentzkow (2010). “Persuasion: Empirical Evidence”. In: *Annual Review of Economics* 2, pp. 643–669.
- DellaVigna, Stefano and Ethan Kaplan (2007). “The Fox News Effect: Media Bias and Voting”. In: *The Quarterly Journal of Economics* 122.3, pp. 1187–1234. (Visited on 05/09/2023).
- Devlin, Jacob (2018). “Bert: Pre-training of deep bidirectional transformers for language understanding”. In: *arXiv preprint arXiv:1810.04805*.
- Druckman, James N. and Michael Parkin (2005). “The Impact of Media Bias: How Editorial Slant Affects Voters”. In: *The Journal of Politics* 67.4, pp. 1030–1049.
- Du Bois, John W. (2007). “The Stance Triangle”. In: *tancetaking in Discourse: Subjectivity, Evaluation, Interaction*. Amsterdam: John Benjamins Publishing, pp. 139–182.
- Duggan, John and Cesar Martinelli (2011). “A Spatial Theory of Media Slant and Voter Choice”. In: *The Review of Economic Studies* 78.2, pp. 640–666.
- Eberl, Jakob-Moritz, Hajo G. Boomgaarden, and Markus Wagner (2017). “One Bias Fits All? Three Types of Media Bias and Their Effects on Party Preferences”. In: *Communication Research* 44.8, pp. 1125–1148.
- Gans, Herbert J. (1979). *Deciding What’s News: A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. Evanston, IL: Northwestern University Press.
- Gans, Joshua S. and Andrew Leigh (2012). “How Partisan is the Press? Multiple Measures of Media Slant”. In: *Economic Record* 88.280, pp. 127–147.
- Garz, Marcel and Jonna Rickardsson (2023). “Ownership and Media Slant: Evidence from Swedish Newspapers”. In: *Kyklos* 76.1, pp. 18–40.
- Garz, Marcel, Gaurav Sood, et al. (2020). “The Supply of Media Slant Across Outlets and Demand for Slant within Outlets: Evidence from US Presidential Campaign News”. In: *European Journal of Political Economy* 63, p. 101877.

- Gentzkow, Matthew, Edward L. Glaeser, and Claudia Goldin (2006). “The Rise of the Fourth Estate. How Newspapers Became Informative and Why It Mattered”. In: *Corruption and Reform: Lessons From America’s Economic History*. Ed. by Edward L. Glaeser and Claudia Goldin. Chicago, IL: University of Chicago Press, pp. 187–230.
- Gentzkow, Matthew and Jesse M. Shapiro (2006). “Media Bias and Reputation”. In: *Journal of Political Economy* 114.2.
- (2010). “What Drives Media Slant? Evidence From U.S. Daily Newspapers”. In: *Econometrica* 78.1, pp. 35–71.
- Gentzkow, Matthew, Jesse M. Shapiro, and Michael Sinkinson (2011). “The Effect of Newspaper Entry and Exit on Electoral Politics”. In: *American Economic Review* 101.7, pp. 2980–3018.
- Gentzkow, Matthew, Jesse M. Shapiro, and Daniel F. Stone (2015). “Media Bias in the Marketplace: Theory”. In: *Handbook of Media Economics*. Vol. 1. Elsevier, pp. 623–645.
- Gerber, Alan S., Dean Karlan, and Daniel Bergan (2009). “Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions”. In: *American Economic Journal: Applied Economics* 1.2, pp. 35–52.
- Groeling, Tim (2013). “Media Bias By the Numbers: Challenges and Opportunities in the Empirical Study of Partisan News”. In: *Annual Review of Political Science* 16, pp. 129–51.
- Groseclose, Tim and Jeffrey Milyo (2005). “A Measure of Media Bias”. In: *The Quarterly Journal of Economics* 120.4, pp. 1191–1237.
- Hamborg, Felix, Karsten Donnay, and Bela Gipp (2019). “Automated Identification of Media Bias in News Articles: An Interdisciplinary Literature Review”. In: *International Journal on Digital Libraries* 20, pp. 391–415.
- Ho, Daniel E. and Kevin M. Quinn (2008). “Measuring Explicit Political Positions of Media”. In: *Quarterly Journal of Political Science* 3, pp. 353–377.

- Kim, Eunji, Yphtach Lelkes, and Joshua McCrain (2022). “Measuring Dynamic Media Bias”. In: *Proceedings of the National Academy of Sciences* 119.32, e2202197119.
- Ladd, Jonathan McDonald and Gabriel S. Lenz (2009). “Exploiting a Rare Communication Shift to Document the Persuasive Power of the News Media”. In: *American Journal of Political Science* 53.2, pp. 394–410.
- Larcinese, Valentino, Riccardo Puglisi, and James M. Snyder Jr. (2011). “Partisan Bias in Economic News: Evidence on the Agenda-Setting Behavior of US Newspapers”. In: *Journal of Public Economics* 95.9-10, pp. 1178–1189.
- Levendusky, Matthew S. (2013). “Why Do Partisan Media Polarize Viewers?” In: *American Journal of Political Science* 57.3, pp. 611–623.
- (2022). “How Does Local TV News Change Viewers’ Attitudes? The Case of Sinclair Broadcasting”. In: *Political Communication* 39.1, pp. 23–38.
- Liu, Yujian et al. (2022). “POLITICS: Pretraining with same-story article comparison for ideology prediction and stance detection”. In: *arXiv preprint arXiv:2205.00619*.
- Lott, John R. and Kevin A. Hassett (2014). “Is Newspaper Coverage of Economic Events Politically Biased?” In: *Public Choice* 160.1/2, pp. 65–108.
- Martin, Gregory J. and Ali Yurukoglu (2017). “Bias in Cable News: Persuasion and Polarization”. In: *American Economic Review* 107.9, pp. 2565–99.
- Mohammad, Saif et al. (2016). “Semeval-2016 task 6: Detecting stance in tweets”. In: *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, pp. 31–41.
- Mullainathan, Sendhil and Andrei Shleifer (2005). “The Market For News”. In: *American Economic Review* 95.4, pp. 1031–1053.

- Peterson, Erik (2021). “Paper Cuts: How Reporting Resources Affect Political News Coverage”. In: *American Journal of Political Science* 65 (2), pp. 443–459. ISSN: 15405907. DOI: 10.1111/ajps.12560.
- Peterson, Erik and Johanna Dunaway (2023). “The New News Barons: Investment Ownership Reduces Newspaper Reporting Capacity”. In: *Annals of the American Academy of Political and Social Science* 707 (1), pp. 74–89. DOI: 10.1177/00027162231211426.
- Petrova, Maria (2011). “Newspapers and Parties: How Advertising Revenues Created An Independent Press”. In: *American Political Science Review* 105.4, pp. 790–808.
- Poole, Keith T. (2005). *Spatial Models of Parliamentary Voting*. Cambridge: Cambridge University Press.
- Poole, Keith T. and Howard Rosenthal (1997). *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Puglisi, Riccardo and James M. Snyder Jr. (2011). “Newspaper Coverage of Political Scandals”. In: *The Journal of Politics* 73.3, pp. 931–950.
- (2015a). “Empirical Studies of Media”. In: *Handbook of Media Economics*. Ed. by Simon P. Anderson, Joel Waldfogel, and David Strömberg. Oxford, UK: Elsevier, pp. 647–667.
- (2015b). “The Balanced US Press”. In: *Journal of the European Economic Association* 13.2, pp. 240–264.
- Reimers, N (2019). “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks”. In: *arXiv preprint arXiv:1908.10084*.
- Roberts, Kevin W. S. (1977). “Voting over income tax schedules”. In: *Journal of Public Economics* 8.3, pp. 329–340.
- Schulz, Norbert and Joachim Weimann (1989). “Competition of Newspapers and the Location of Political Parties”. In: *Public Choice* 63 (2), pp. 125–147.

- Seamans, Robert and Feng Zhu (2014). “Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers”. In: *Management Science* 60 (2), pp. 476–493. ISSN: 00251909. DOI: 10.1287/mnsc.2013.1785.
- Silcock, Emily et al. (2024). “Newswire: A Large-Scale Structured Database of a Century of Historical News”. In: *arXiv:2406.09490*, pp. 679–691.
- Song, BK (2020). “Did Television Reduce the Effect of Partisan Press on Electoral Politics?” In: *American Politics Research* 48.1, pp. 78–98.
- Stone, Daniel F. (2011). “Ideological Media Bias”. In: *Journal of Economic Behavior and Organization* 78 (3), pp. 256–271.
- Strömberg, David (2004). “Mass Media Competition, Political Competition, and Public Policy”. In: *The Review of Economic Studies* 71.1, pp. 265–284.

A Appendix

A.1 Newspapers.com Data Collection

Using simple searches for ballot propositions in given states, we first scraped the optical character recognition (OCR) output for 432,364 Newspapers.com newspaper pages. However, we found it infeasible to extract individual article text from the OCR output because the OCR was riddled with transcription errors and the output concatenated all text on the page into a single string, arbitrarily splitting articles. Since many of the pages we scraped could not plausibly contain editorial board endorsement articles since they were not on the opinion page, we trained an n -gram multinomial naive Bayes classifier using a dataset of 1,800 human-coded OCR output newspaper pages to identify each page as an opinion page or not. Table A.1 displays the performance metrics of this model. We note that we set our objective function to maximize recall because we are interested in discovering as many opinion pages as possible. After running inference on the full set of 432,364 newspaper pages, we then scraped the 73,125 high-quality newspaper page scans inferred to be opinion pages.

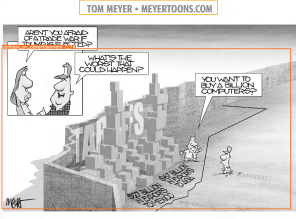
Metric	Train	Test
Precision	.96	.88
Recall	.98	.91
F1 Score	.97	.89

Table A.1: Performance metrics for the training and testing sets with the opinion dataset of 1,800 human-coded examples.

We are currently in the process of extracting article text from these scans using methods in computer vision. First, we use the Google Cloud Vision API to extract high-quality OCR for discovered bounded rectangles throughout the scan. Second, we develop a custom algorithm that uses hierarchical agglomerative clustering (HAC-GCV) to automatically detect which rectangles contain article text (as opposed to titles, advertisements, etc.), group rectangles by article, and then order those grouped rectangles in the reading order of the article. We are also experimenting with the Mask R-CNN deep learning model introduced in Silcock et al. (2024) to discover bounded rectangles throughout the scan. Figure 8 visualizes the bounded rectangle computer vision process.

Gun measure too high-stakes to fail

California has among the lowest firearm-related crime rates in the nation, at least according to the FBI. But the state's gun laws are so strict that many of those who would seek out guns there are from out of state. The Sacramento Bee editorial board generally believes reducing gun violence is a top priority for the state, and we support the measure.



LETTERS TO THE EDITOR

Farmer must understand
The author of the article "Farmer must understand" (Page 1A) is wrong. The author is a farmer and should know that the author is wrong. The author is a farmer and should know that the author is wrong.

California's initiative system is broken

The initiative system in California is broken. It is a system that allows citizens to bypass the legislature and pass laws directly. This system is broken because it is too easy to pass laws and too difficult to amend them.

Opinion

Rich people will pay
The author of the article "Rich people will pay" (Page 1A) is wrong. The author is a rich person and should know that the author is wrong. The author is a rich person and should know that the author is wrong.

Smoking continues

The author of the article "Smoking continues" (Page 1A) is wrong. The author is a smoker and should know that the author is wrong. The author is a smoker and should know that the author is wrong.

Democracy will add

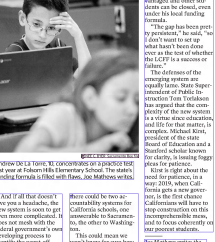
The author of the article "Democracy will add" (Page 1A) is wrong. The author is a democrat and should know that the author is wrong. The author is a democrat and should know that the author is wrong.

The fight against

The author of the article "The fight against" (Page 1A) is wrong. The author is a fighter and should know that the author is wrong. The author is a fighter and should know that the author is wrong.

New accountability system could leave poor students behind

California's education system is in a state of crisis. The new accountability system could leave poor students behind. The system is flawed and needs to be reformed.



Will Prop. 54 make Legislature work better for Californians?

Proposition 54 is a measure that would change the way the California Legislature works. It would require the Legislature to hold public hearings on all bills before they are passed.

Yes! It will help voters participate in lawmaking

Proposition 54 will help voters participate in lawmaking. It will give voters a voice in the legislative process and make the government more accountable.

No! It plays into the hands of special interests

Proposition 54 plays into the hands of special interests. It will give special interests a voice in the legislative process and make the government less accountable.

Mask R-CNN model HAC-GCV algorithm

Figure 8: Newspapers.com digitization

A.2 Issue Subsets

One possibility is that newspapers have easily measured preferences over a multidimensional issue space, e.g. with issue dimensions representing bundles such as healthcare, taxes, or education. To test this, we hand-code each proposition for California, 1988-2022, for its relevant issue or issues. Each proposition can be coded with multiple issues – for instance, a proposed increase in the gasoline tax to fund the construction of new K-12 schools would be coded as related to taxes, transportation, and education. We identify 8 issue bundles that have a sufficient number of propositions to analyze separately: bonds, business, crime, education, elections, healthcare, moral issues, and taxes. For each analysis, we restrict analysis to propositions with at least 6 endorsements and endorsers with at least 15 endorsements.

	Optimal Classification		W-NOMINATE		
	Newspapers	Groups	Newspapers	Groups	# Props
Bonds	43.9	75.7	34.3	70.5	86
Business	48.2	80.0	16.0	76.7	28
Crime	43.0	71.1	23.8	71.0	43
Education	46.0	59.1	22.0	55.1	31
Elections	37.1	48.1	14.9	36.4	40
Healthcare	34.5	63.6	18.0	48.7	31
Moral issues	45.4	68.9	26.1	46.3	26
Taxes	30.4	62.4	18.8	51.8	63
Avg	41.1	66.1	21.7	57.1	43.1

Table A.2: APRE of issue-specific bundles

Table A.2 displays the results. For each issue bundle, and using both the optimal classification and W-NOMINATE models, the average APRE of newspapers is significantly lower than the average APRE of groups – by a 25-35 percentage point margin across issues. There is little evidence that newspapers behavior can be cleanly measured in a multidimensional ideological space.

A.3 Strategic Endorsing Behavior

We check for strategic endorsing behavior in two ways – first, by examining endorsements on close votes; and second, by examining endorsing behavior as a function of total spending on each ballot proposition campaign. First, we examine endorsements on close votes, defined as those for which the margin of victory or defeat is under 7.5 percentage points.²⁶ We define moderate endorsers as those within 33 optimal-classification ranks of the center (approximately the 50 percent closest to the center).²⁷ We run a series of linear regressions of endorsements – more precisely, whether a newspaper took a position consistent with their average ideological slant – on indicator variables for a close election, a moderate endorser, and the interaction between the two. As the top panel of Table A.3 shows, the point estimates of the coefficient on the interaction term are small and statistically insignificant. This is true whether or not we control for proposition fixed effects or proposition and endorser fixed effects.

²⁶The results are robust to using 5 percent or 10 percent thresholds instead.

²⁷The results are robust to using thresholds of 25, 30, 35, and 40 as well.

	Model 1	Model 2	Model 3
Moderate endorser × close vote	-0.009 (0.021)	-0.009 (0.021)	-0.007 (0.020)
Fixed effects	None	Prop	Prop & Endorser
Moderate endorser × log spending	-0.041*** (0.005)	-0.042*** (0.004)	-0.040*** (0.004)
Fixed effects	None	Prop	Prop & Endorser
<i>Note:</i>			***p<0.001

Table A.3: Strategic endorsing linear regressions

Second, we examine endorsements as a function of total campaign spending, under the assumption that higher spending is a proxy for “revealed importance” of a ballot measure. The regressions are identical to the above, except we replace the indicator variable for close election with a continuous variable for logged spending. The bottom panel of Table A.3 shows that, regarding the interaction between moderate endorser and total (log) spending, the estimated coefficients are statistically significant but wrongly signed from the point of view of the strategic endorsing hypothesis. In other words, moderate endorsers are more likely to endorse contrary to their typical slant on ballot measures where spending is high.

A.4 BERT Model

We fine-tune a BERT model – a pre-trained neural network that represents text in a low-dimensional vector space that encodes semantic and world knowledge acquired from initial training on a massive corpus – on training data that consists of two strings, a document and a target, and our stance label, PRO, CON, or NEUTRAL. We provide the model with two strings: a document and a target, separated by a special token. These strings are then embedded by BERT as a sentence pair into a single 768-dimensional vector. We then fit a two-layer feed-forward neural network on the classification token embedding from the last hidden state of the model while simultaneously fine-tuning additional layers of the pre-trained model to predict the stance label. We run a Bayesian optimization search over hyperparameters of the model, training $n = 16$ variations of our model on random splits of our data, selecting our best performing model on the unseen *development* dataset. Table A.4 summarizes the

class distributions from the BERT model.

Dataset	Examples	NEUTRAL	CON	PRO
Articles	4,025	N/A	2,247	1,778
Paragraphs	63,090	36,559	17,708	8,823

Table A.4: Summary of dataset with class distribution at the article and paragraph level

A.5 Beta Regression Model

Figure 9 displays R^2 values from beta regression models on the various combinations of stance detection measurements (DVs) and ideal point estimates (IVs).

Same Side Article Proportion (Words)	0.006	0.004	0.000	0.006	0.002
Opposite Side Article Proportion (Words)	0.003	0.005	0.002	0.001	0.006
Normalized One-Sidedness (Words)	0.005	0.000	0.000	0.005	0.004

(a) California

Same Side Article Proportion (Words)	0.006	0.004	0.000	0.006	0.002
Opposite Side Article Proportion (Words)	0.003	0.005	0.002	0.001	0.006
Normalized One-Sidedness (Words)	0.005	0.000	0.000	0.005	0.004
	Est. Prob. of Recommendation (NOMINATE)	Ideal Point Extremism (NOMINATE)	Ideal Point Extremism (OC)	NOMINATE 1st Dimension	OC Rank

(b) All States

Figure 9: Heatmaps of R-squared values for beta regression models. Outcomes are estimated from stance detection, predictors from ideal points.