Hayes, Danny and Jennifer L. Lawless. "The Decline of Local News and Its Effects: New Evidence from Longitudinal Data."

Online Appendix A: Conducting the Local News Content Analysis

There is no accepted method of identifying or analyzing the local news outlets that serve a particular House contest. We developed an approach that allows for a comprehensive assessment of the volume and content of local news coverage of congressional campaigns.

We began by selecting an appropriate newspaper for each House race in 2010. We consulted maps of every congressional district, identified the largest city in each, and then determined whether that city had a daily newspaper we could access through an electronic databases or the newspaper's online archives. In the few cases for which we could not gain access to newspaper coverage from the district's largest-circulation daily paper, we relied on coverage from the next largest paper. Because of redistricting, as well as shifts in newspaper circulation within districts, we repeated this exercise for every congressional district in 2014. Across election cycles, the average newspaper circulation size in our districts varied greatly; the smallest paper had a readership of only about 5,000, whereas the largest circulation size approximated 1.8 million. The average circulation of the papers in our data set is roughly 178,000.

For every district, after choosing a newspaper, we identified in the 30 days leading up to the election every news story that mentioned at least one of the two major party candidates. We collected straight news reports, news analyses, editorials, and op-ed columns. We did not restrict the data collection strictly to "campaign" stories because we assume that any information about the House candidates is potentially relevant for voters. Our method produced 10,375 stories about 1,550 candidates in 815 districts who received at least some local news coverage in either the 2010 or 2014 midterms.

Coders read the full text of each article and recorded several pieces of information.[1] In addition to whether both candidates were mentioned, the two key variables for the purposes of this paper are

[1] Before undertaking the content analysis, two coders participated in several hours of practice coding, using news stories from House elections in previous years.

references to issues and traits associated with a candidate. More specifically, we tracked every time an issue was mentioned, beginning with a list of issues commonly included in previous studies and then recording references to additional issues as they emerged in the coverage. Overall, we tracked coverage of 218 separate issues. We also recorded the number of explicit references to candidate traits, both positive and negative. These references could come from the candidates themselves, their opponents, or reporters. Although we began the coding with a list of traits commonly included in previous studies, we recorded references to every additional trait we encountered in the coverage. In all, we coded for 207 separate traits.

In contrast to many studies of campaign media coverage, which tend to conduct the analysis at the story or paragraph level, we carried out our coding at the level of the individual reference. In other words, we account for *every* time an issue or trait is mentioned. Inter-coder reliability for our variables was high. We achieved 100% agreement between two coders on the number of relevant stories in the district (which is expected, since this simply required searching relevant databases). We then double-coded 5% of the stories in our sample (N=518). Cohen's kappa was 0.98 for whether a story mentioned both candidates; 0.92 for the number of issues mentioned; and 0.86 for the number of traits mentioned. These scores are well above those generally accepted as indicators of a reliable coding scheme. The detail and depth of our coding allows for an unusually nuanced analysis of the local news environment in both 2010 and 2014.

For all of the attributes of our media data, it is important to recognize a limitation when using it to draw direct comparisons between House race coverage in 2010 and 2014. In some cases, district boundaries shifted as a result of redistricting, so the largest circulating newspaper in 2010 was not always the largest paper for the district in 2014. In other cases, the paper remained the same but the composition of the district changed. Although these are important considerations, the evidence we have suggests that they do not compromise our analysis.

First, our findings are virtually the same regardless of whether the relevant local paper was different in the two elections, or whether the paper stayed the same. For instance, in districts where the paper changed, the decrease in the number of stories from 2010 to 2014 was 2.4. In same-paper districts, it was 2.9. The patterns for the other measures in Figure 1 – percentage of stories mentioning both candidates, issue mentions, and trait mentions – are also similar.

Second, changes in the composition of districts do not account for the decline in coverage. Because previous research (e.g., Arnold 2004) has found that district demographics can influence local news coverage, we control in our models for district-level race, income, and education. The fact that the same factors, like electoral competitiveness and newspaper circulation size, predict coverage across our two elections in a very similar way suggests that redistricting does not explain the over-time changes in the volume and substance of news we observe.

Third, in separate models predicting the number of stories about a race (controlling for candidate spending and quality, as well as district demographics), the Cook Rating coefficient – which gauges electoral competitiveness – is 5.8 in the "same paper" model and 4.4 in the "different paper" model. In a pooled model that includes an interaction between the Cook Rating and whether the paper changed or stayed the same, the interaction term is not significant. In other words, the same/different paper status does not moderate the effect of our key explanatory variable or compromise the basic process that creates the news.

Even with the measurement challenges of redistricting, there is little doubt that the total amount of congressional campaign news coverage reported in the largest circulating newspapers in districts across the country was less, and less substantive, in 2014 than in 2010.

Online Appendix B. Predicting the Volume and Substance of Newspaper Coverage in House Elections

	2010				2014			
	Number of Stories	**Both Candidates Mentioned**	**Number of Issue Mentions**	**Number of Trait Mentions**	**Number of Stories**	**Both Candidates Mentioned**	**Number of Issue Mentions**	**Number of Trait Mentions**
Cook Rating (4-point scale)	3.013 *	8.133 *	15.958 *	2.049 *	5.238 *	9.485 *	27.711 *	4.406 *
	(0.723)	(1.620)	(3.361)	(0.633)	(0.756)	(2.190)	(3.913)	(0.600)
Newspaper Circulation	-0.163 *	-0.309 *	-0.867 *	-0.088 *	-0.050 *	-0.309 *	-0.442 *	0.026
	(0.032)	(0.078)	(0.157)	(0.030)	(0.016)	(0.077)	(0.131)	(0.020)
Open Seat	-2.196	14.186 *	-3.370	-0.299	0.522	20.919 *	5.476	1.234
	(1.907)	(4.444)	(8.854)	(1.669)	(1.413)	(4.427)	(7.295)	(1.118)
Uncontested	-2.851	--	-15.209	-2.451	-2.689 *	--	-17.904 *	-2.598 *
	(2.155)		(10.918)	(2.057)	(1.175)		(6.429)	(0.986)
Quality Candidate	4.213 *	1.089	10.265	2.578 *	1.904 *	-0.697	9.141	0.437
	(1.479)	(3.367)	(6.962)	(1.312)	(1.134)	(3.365)	(5.858)	(0.898)
Candidate Spending	0.160 *	0.195 *	0.581 *	0.097 *	0.051 *	0.120	0.376 *	0.045 *
	(0.039)	(0.087)	(0.179)	(0.034)	(0.029)	(0.084)	(0.148)	(0.023)
District % White	-0.016	0.228 *	0.032	-0.032	0.024	0.342 *	0.312 *	0.014
	(0.034)	(0.084)	(0.169)	(0.032)	(0.026)	(0.090)	(0.142)	(0.022)
District Median Income	-0.092	0.148	1.152	-0.144	0.196	3.151 *	4.179 *	0.322
	(0.598)	(1.430)	(2.887)	(0.544)	(0.479)	(1.497)	(2.520)	(0.386)
District % College Educated	0.222 *	-0.250	0.999 *	0.206 *	0.017	-0.426 *	-0.100	0.001
	(0.099)	(0.238)	(0.474)	(0.089)	(0.067)	(0.220)	(0.362)	(0.056)
Constant	7.710 *	28.524 *	11.209	3.014	6.016 *	12.147	-8.229	-0.903
	(2.800)	(6.875)	(14.026)	(2.643)	(2.471)	(8.344)	(13.441)	(2.061)
N	435	380	405	405	435	353	416	416
Adjusted R^2	0.306	0.332	0.313	0.180	0.304	0.306	0.364	0.298

Notes: Cell entries are OLS regression coefficients with robust standard errors clustered on congressional district in parentheses. Levels of significance: * p < .05, one-tailed. The "Both Candidates Mentioned" models are restricted to races that included two major-party candidates and at least one story about the race. The "Issue Mentions" and "Trait Mentions" models include only districts with at least one story. Poisson regressions produce the same results. Our variable coding follows Hayes and Lawless (2015).

Online Appendix C: The CCES Panel Data and Analysis

The Cooperative Congressional Election Study is a nationally representative survey conducted in two stages by YouGov and in collaboration with dozens of academic institutions. Respondents complete one survey in the days leading up to the midterm elections and another shortly after Election Day. All CCES respondents answer basic demographic and political questions, including who they plan to vote for in their House elections.

The 2010 – 2014 panel component of the CCES makes it especially valuable for assessing changes in political engagement over time. The 2010 CCES interviewed more than 55,000 adults who answered a common battery of questions before and after the midterm elections. For the 2014 wave, YouGov attempted to re-interview 22,346 respondents who had completed both the 2010 and 2012 waves. They successfully completed interviews with 15,252 panelists (68% of those they attempted to re-contact). The 2010-2014 Panel Study includes a subset of 9,500 of these respondents, all of whom were interviewed before and after the 2014 elections.

Although panel attrition is always a concern, on the most relevant measures for our analysis, differences between the base sample and the 2014 panel respondents are small. Moreover, there are no partisan differences in attrition. Importantly, voters were much more likely to be successfully re-contacted compared to non-voters, but that biases our estimates downward and makes for a more difficult case to uncover media effects.[2]

In addition to the central difference-in-difference analysis presented in the manuscript, we also conducted a cross-sectional analysis, in which we used the 2014 CCES and our 2014 news data to determine whether the volume of House coverage correlates with political engagement. Unlike other cross-sectional analyses, however, we can account for respondents' past knowledge and participation, which are very strong predictors of future engagement. To the extent that we find news coverage matters

[2] For a detailed description of the panel student, re-interview rates, and attrition, see: https://dataverse.harvard.edu/file.xhtml?fileId=2864258&version=6.0.

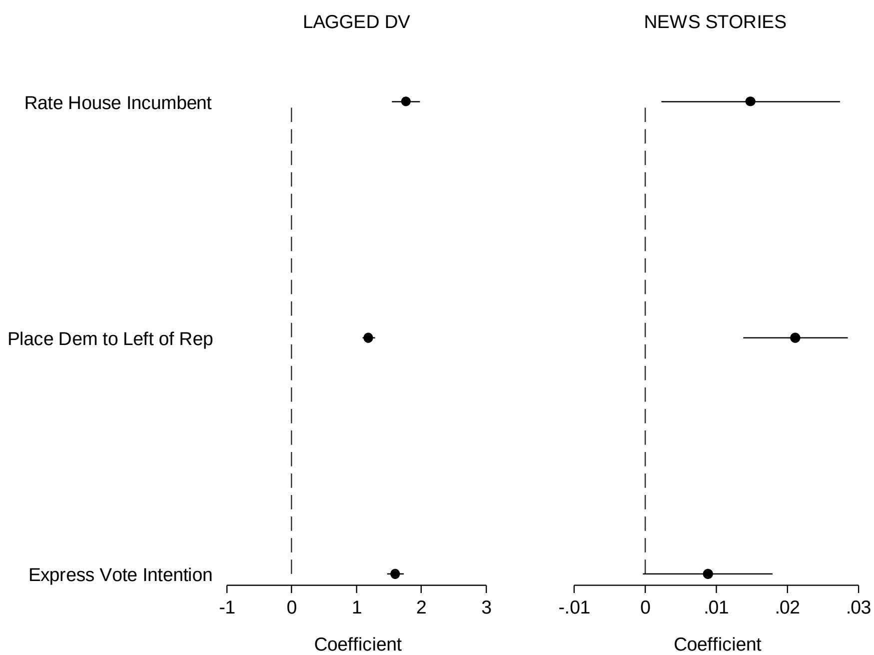
above and beyond an individual's habitual patterns of engagement, the more confidence we can have that it truly matters.

We specified three logistic regression equations. Our dependent variables indicated whether respondents in 2014 could (1) offer a rating of their House incumbent, (2) place the Democratic House candidate to the left of the Republican on an ideological scale, and (3) offer a vote intention in the pre-election survey. The key independent variables are the number of stories about the 2014 House race in a respondent's congressional district and a lagged dependent variable indicating the respondent's knowledge or participation in 2010. We include standard controls for demographics, the competitiveness of the race, candidate spending, and district characteristics. If local media affects engagement, individuals living in districts with little news coverage will be less informed about their House race than people living in districts with more coverage.

The regression results reveal significant effects for news coverage, and also the value of panel data. Consider first the effect of the lagged dependent variables, whose regression coefficients and 90% confidence intervals are displayed on the left side of Figure A1. People who in 2010 could rate their House member, place the Democratic House candidate to the left of the Republican, and express a pre-election vote intention were much more likely than people who couldn't to be able to do those things in 2014 as well. For instance, the probability that a respondent who did not express a vote choice in 2010 would express one in 2014 was 0.52. But for an otherwise identical respondent who did express a vote choice in 2010, the probability in 2014 was 0.84.

Even accounting for the substantial explanatory power of previous knowledge and participation, we still find significant news effects. The coefficients presented on the right side of Figure A1 indicate that as the number of stories about a congressional race increased, respondents were more likely to rate the incumbent and place the Democratic candidate to the left of the Republican on the ideology scale. If we relax the threshold for statistical significance, then they were also more likely to express a vote intention ($p < 0.11$). The magnitude of the effects is modest. A one standard deviation (about 9 stories) decrease in news coverage, for example, decreases by about 2.0 percentage points the probability that a respondent who did not vote in 2010 expressed a vote intention in 2014. (The magnitude of the effect is about half that for a respondent who did express a vote intention in 2010). But given that we are controlling for standard individual-level factors, other features of the electoral environment, and respondents' history of knowledge and participation, it's remarkable that we find any measurable relationship at all.

Figure A1. Predicting Political Knowledge and Participation in House Elections in 2014: A Cross-Sectional Analysis with Lagged Dependent Variables



Notes: Dots represent logistic regression coefficients, with 90% confidence intervals. The lagged dependent variable for each equation is whether the respondent could offer an answer to the knowledge or participation question in 2010. See Appendix E for the full regression equations.

Critically, placebo tests (see Appendix E) find that general political knowledge – knowing which party controls the House and Senate – is not affected by campaign coverage, showing that the volume of news is not simply a *reflection* of district-level political engagement. This is perhaps the best evidence that our measure of news coverage provides information voters rely on when thinking specifically about local elections.

Online Appendix D.
Predicting Changes in Political Knowledge and Participation in House Elections:
A Difference-in-Difference Analysis

	Rate House Incumbent	**Place Democrat Left of Republican**	**House Vote Intention**
Difference in Number of Stories	0.005 *	0.011 *	0.007 *
	(0.003)	(0.002)	(0.002)
Difference in Total Spending	0.024	0.165 *	0.042 *
	(0.018)	(0.011)	(0.012)
Difference in Strength of Partisanship	-0.030	0.031	0.179 *
	(0.054)	(0.032)	(0.038)
Constant (cut point 1)	-2.987 *	-1.939 *	-2.530 *
	(0.049)	(0.034)	(0.041)
Constant (cut point 2)	3.540 *	1.437 *	2.181 *
	(0.063)	(0.029)	(0.035)
Pseudo R^2	0.001	0.030	0.006
Log Likelihood	-2951.463	-6976.396	-5310.856
Chi-square	8.067	424.396	61.345
N	9,218	8,061	9,013

Notes: Cell entries are ordered logistic regression coefficients. Robust standard errors clustered on congressional district are in parentheses. Levels of significance: * $p < .05$, one-tailed. Independent variables represent the differences in the number of stories about the House race, the total spending in the House race, and the respondent's strength of partisanship in 2014 compared to 2010.

Online Appendix E. Predicting Political Knowledge and Participation in House Elections in 2014: A Cross-Sectional Analysis with Lagged Dependent Variables

	PLACEBO TESTS				
	Know House Majority	**Know Senate Majority**	**Rate House Incumbent**	**Place Democrat Left of Republican**	**House Vote Intention**
Lagged Dependent Variable	1.967 *	2.380 *	1.765 *	1.192 *	1.604 *
	(0.075)	(0.082)	(0.133)	(0.058)	(0.077)
Number of Stories	0.004	0.002	0.015 *	0.021 *	0.009
	(0.003)	(0.004)	(0.008)	(0.004)	(0.006)
Competitiveness	-0.007	0.034	-0.101	0.150 *	-0.019
	(0.060)	(0.054)	(0.145)	(0.067)	(0.071)
Uncontested	-0.089	-0.129	0.239	--	-1.058 *
	(0.087)	(0.092)	(0.185)		(0.108)
Open Seat	-0.259 *	-0.085	-0.490 *	-0.462 *	-0.475 *
	(0.115)	(0.102)	(0.227)	(0.127)	(0.141)
Quality Candidate	0.015	0.003	0.097	0.407 *	0.097
	(0.080)	(0.084)	(0.178)	(0.102)	(0.109)
Republican Spending	-0.001	-0.001	0.069	0.090 *	0.001
	(0.021)	(0.021)	(0.065)	(0.035)	(0.023)
Democratic Spending	-0.079 *	-0.047	0.129	0.233 *	0.051
	(0.035)	(0.041)	(0.095)	(0.044)	(0.045)
Education	0.260 *	0.235 *	0.142 *	0.153 *	0.059 *
	(0.021)	(0.022)	(0.037)	(0.018)	(0.026)
Respondent Sex (female)	-1.264 *	-1.291*	-0.608 *	-0.614 *	-0.684 *
	(0.068)	(0.069)	(0.111)	(0.049)	(0.071)
Strength of Partisanship	0.183 *	0.158*	0.198 *	0.158 *	0.463 *
	(0.029)	(0.028)	(0.047)	(0.023)	(0.030)
Constant	-0.751 *	-0.920 *	0.767*	-1.843 *	-0.028 *
	(0.130)	(0.133)	(0.244)	(0.111)	(0.146)
Pseudo R^2	0.208	0.241	0.083	0.149	0.176
Log Likelihood	-4593.092	-4563.618	-1531.179	-5556.531	-3752.916
Chi-square	1431.311	1618.530	254.465	1217.6975	1006.005
N	9,364	9,347	9,171	8,017	8,951

Notes: Cell entries are logistic regression coefficients. Robust standard errors clustered on congressional district are in parentheses. The "Place Democrat Left of Republican" model is restricted to contested races. Levels of significance: * p < .05, one-tailed.