

Online Appendix:

Press Coverage and Accountability in State Legislatures

American Political Science Review

Andrew C. W. Myers

Stanford University

Contents

A	Summary of Prior Published Research on Press Coverage and Accountability	A.1
B	Descriptive Statistics	A.2
C	Computing Congruence	A.3
D	Newspaper Corpus Data	A.4
E	Roll-Call and Bill Sponsorship Data	A.5
F	Newspaper Market–Legislative District Congruence Robustness Checks . . .	A.6
G	Electoral Selection Robustness Checks	A.7
H	Regression Discontinuity Balance Tests	A.9
I	Productivity Robustness Checks	A.10
J	Non-Parametric Estimates of Multiplicative Interactions	A.11

A Summary of Prior Published Research on Press Coverage and Accountability

Table A.1 – Summary of Prior Studies of Press Coverage and Accountability. The table lists prior studies of newspaper coverage and accountability in Congress (Panel A) and state legislatures and municipal governments (Panel B).

Panel A:		Congress			
Outcome	Snyder and Stromberg (2010)	Arnold (2004), Peterson (2021 <i>a</i>), Hayes and Lawless (2015)	Moskowitz (2021), Filla and Johnson (2010)	Trussler (2021, 22), Prior (2006), Schaffner (2006)	Canes-Wrone and Kistner (2023)
Voter knowledge	✓	✓			
Ballot rolloff/turnout	✓		✓		
Incumbency advantage	✓			✓	
Electoral returns to moderation					✓
Committee activity	✓				
Witness appearances	✓				
Missed roll-call votes					
Bill sponsorship					
Voting with party	✓				
Government spending	✓				
Ideological representation	✓				
Panel B:		Municipal Government		State Legislatures	
Outcome	Rubado and Jennings (2020)	Hopkins and Pettingill (2018), Schulhofer-Wohl and Garrido (2013)	Carpini, Keeter, and Kennamer (1994)	Rogers (2017,2023 <i>a</i>)	This Manuscript
Voter knowledge			✓		✓
Ballot rolloff/turnout	✓				✓
Incumbency advantage		✓			✓
Electoral returns to moderation					✓
Committee activity					✓
Legislative productivity					✓
Witness appearances					
Missed roll-call votes					✓
Bill sponsorship					✓
Voting with party					
Government spending					
Ideological representation				✓	✓

B Descriptive Statistics

Table B.1 – Summary Statistics for Control Variables.

Variable	Mean	Median	Min	Max	Std. Dev.	Data Source
Freshman	0.2	0.0	0.0	1.0	0.4	SLERs
Experience	1.7	1.0	0.0	10.0	1.9	SLERs
Chair	0.1	0.0	0.0	1.0	0.3	Fourinaies (2018)
Close Race	0.3	0.0	0.0	1.0	0.5	Author
Uncontested Race	0.4	0.0	0.0	1.0	0.5	SLERs
Open Seat	0.2	0.0	0.0	1.0	0.4	SLERs
Median Income	52,754.0	50,785.0	22,020.0	115,458.0	12,568.0	IPUMS
Population Density	1,938.0	336.0	0.9	113,772.0	5,302.0	IPUMS
% Urban	69.0	75.0	0.0	100.0	25.0	IPUMS
% Retired	15.0	15.0	5.3	46.0	3.6	IPUMS
% Veterans	4.6	3.4	0.2	26.0	2.8	IPUMS
% Foreign Born	7.7	5.2	0.2	53.0	7.6	Census Bureau

Table B.2 – Summary Statistics for Outcome Variables.

Variable	Mean	Median	Min	Max	Std. Dev.	Data Source
State Legislator Name Recall	0.0	0.0	0.0	1.0	0.0	Rogers (2018)
Rated State Legislator	0.8	1.0	0.0	1.0	0.4	CES
Knows Majority in U.S. House	0.6	1.0	0.0	1.0	0.5	CES
Knows Majority in U.S. Senate	0.5	1.0	0.0	1.0	0.5	CES
Knows Majority in State House	0.7	1.0	0.0	1.0	0.4	CES
Knows Majority in State Senate	0.7	1.0	0.0	1.0	0.4	CES
Roll-off in State Leg.	4.0	3.7	-15.0	15.0	3.8	Author
Roll-off in U.S. Senate (Placebo)	2.0	1.2	-14.0	15.0	2.6	Author
Dem. Vote Share in t	0.5	0.5	0.0	1.0	0.2	SLERs
Dem. Vote Share t+1	0.5	0.5	0.0	1.0	0.1	SLERs
Percent Floor Votes Missed	3.4	0.0	0.0	97.0	8.1	LegiScan/Fourinaies and Hall (2022)
Number of Bills Sponsored	70.0	27.0	0.0	2,484.0	123.0	LegiScan/Fourinaies and Hall (2022)
Probabilitiy on Power Committee	0.4	0.0	0.0	1.0	0.5	Bucchianeri et al. (2024)
NP-Score	0.1	0.3	-3.0	3.4	1.0	Shor and McCarty (2011)

C Computing Congruence

I compute Congruence using newspaper circulation data within each district, based on observed circulation data at the newspaper-county level. Let x_{mct} be the circulation of paper m in county c in year t . Following Snyder and Stromberg (2010), I assume that the number of copies of newspaper m sold in county c in year t is proportionate across district d . I then impute district-level circulation as $x_{mdt} = \sum_c (\frac{n_{cdt}}{\sum_{d'} n_{cd't}} x_{mct})$, where n_{cdt} is the population of the part of district d in county c in year t .

Drawing on this data, I calculate m 's market share in d as

$$MarketShare_{mdt} = \frac{x_{mdt}}{\sum_{m'} x_{m'dt}}, \quad (1)$$

and m 's share of readers in district d as

$$ReaderShare_{mdt} = \frac{x_{mdt}}{\sum_{d'} x_{md't}}. \quad (2)$$

Intuitively, Market Share represents each newspaper's share of total sales in a given district, while Reader Share captures the share of a newspaper's readership that resides in the district. To capture Congruence, I weight Reader Share by Market Share to account for the probability that coverage reaches a given reader:

$$Congruence_{dt} = \sum_{m=1}^M MarketShare_{mdt} ReaderShare_{mdt}. \quad (3)$$

D Newspaper Corpus Data

To build a comprehensive dataset of observed legislative news coverage, I identify 272 local and regional newspapers on Newspapers.com, representing approximately 20% of all newspapers included in my circulation dataset. Using this text corpus, I estimate q_{mdt} —the number of articles appearing in newspaper m about the legislator representing district d in year t —by searching for the name of the legislator, their state, and the name of their legislative chamber. In total, my sample includes nearly one million articles about state legislators. Table D.1 shows the characteristics of newspapers contained (column two) and not contained (column three) in the archive. Column four of Table D.1 reports the difference between columns two and three and column five reports the standardized mean difference. Overall, the sample of newspapers to which I have full text are highly similar to newspapers not included in the archive.

Table D.1 – Newspaper Text Data Balance Table. This table reports average values for each newspaper attribute broken down by whether I have access to the newspaper’s full text. The *Difference* column reports the difference between columns two and three. Standard deviations are reported in parenthesis.

Attribute	All Newspapers (1)	Newspapers with Full Text Data (2)	Newspapers without Full Text Data (3)	Difference (4)	Standardized Mean Difference (5)
1 Average Daily Circulation	59,024 (228,489)	65,592 (91,815)	57,478 (249,998)	-6,568	-0.04
2 Share Eastern Newspapers	0.17 (0.38)	0.16 (0.37)	0.18 (0.38)	0.01	0.04
3 Share Midwestern Newspapers	0.35 (0.48)	0.33 (0.47)	0.36 (0.48)	0.02	0.04
4 Share Southern Newspapers	0.32 (0.47)	0.32 (0.47)	0.32 (0.47)	-0.00	-0.01
5 Share Western Newspapers	0.16 (0.36)	0.18 (0.39)	0.15 (0.36)	-0.03	-0.07
6 Average Rural Share of Circ.	0.63 (0.20)	0.67 (0.17)	0.62 (0.20)	-0.03	-0.18
7 Average Dem. Share of Circ.	0.08 (0.14)	0.09 (0.14)	0.08 (0.14)	-0.00	-0.03
Number of Newspapers	1,421	272	1,149	-	-

Note: The *Difference* column may not sum to the difference between columns 1 and 2 due to rounding. Rural share of circulation is calculated using Census Bureau estimates of the share of each legislative district that is rural. Democratic share of circulation is calculated using average district two-party presidential vote share within a redistricting cycle.

E Roll-Call and Bill Sponsorship Data

State legislative roll-call and bill sponsorship data were collected by the author from the online data vendor Legiscan.com and combined with similar data from Fournaies and Hall (2022). This data includes roll-call votes and bill introductions for the near-universe of chamber-years for the years 2012-2022 and roughly half of chamber-years for the years 2000-2011. Approximately 20% of the data originate from Fournaies and Hall (2022) and the remaining 80% were collected by the author from Legiscan.com. Table E.1 reports the full coverage of the roll-call dataset. Coverage of bill-sponsorship data is identical.

Table E.1 – Roll-Call Data Coverage Matrix. This table reports the coverage of my roll-call dataset in terms of states and years. Cells contain the number of roll-call votes observed in thousands.

State	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
AK	13	13	18	13	15	21	22	11	7	17	25
AL	55	122	157	121	101	139	106	105	111	116	49	178	136
AR	.	141	.	135	.	155	.	121	.	93	.	203	39	220	41	183	40	167	36	163	26	181	33
AZ	76	67	57	46	55	59	70	51	55	36	51	68	74	60	64	67	76	65	67	61	49	91	79
CA	147	137	141	128	132	115	119	118	130	213	187	262	265	254	284	279	296	295	323	321	123	259	315
CO	17	6	29	31	28	31	37	46	52	58	49	45	105	87	119	125	90	134	120
CT	12	67	52	82	61	70	89	118	98	120	18	117	91
DE	15	18	16	16	16	19	9	9	18	20	19	5	22	22
FL	.	.	.	92	110	95	90	84	82	76	109	112	112	96	87	87	81	69	57	53	53	118	119
GA	171	42	113	168	127	123	126	127	116	123	120	107	126	191
HI	52	27	26	29	42	19	28	26	26	53	42
IA	48	71	64	37	32	100	10	60	28	73	54
ID	42	43	44	43	43	46	43	44	41	44	48	43
IL	232	165	191	134	175	149	161	123	162	158	164	10	203	117
IN	0	89	53	92	83	91	68	83	66	98	60	77	67
KS	94	62	53	46	43	44	44	45	31	18	53	38
KY	17	37	42	38	82	24	56	49	42	60	66
LA	55	222	90	208	171	107	163	96	172	112	428	220	364	212	381	246	200	130	203	135	150	153	212
MA	58	45	24	30	19	19
MD	64	202	286	154	215	183	230	254	250	236	200	241	245
ME	.	.	.	43	43	59	38	34	42	43	21	41	25	85	61	88	39	83	46	60	6	78	32
MI	.	61	89	61	83	67	100	55	100	61	48	101	147	100	149	84	125	84	167	63	100	94	66
MN	51	59	74	60	45	43	49	39	67	32	54	33	34
MO	119	118	122	129	105	104	97	102	107	124	94	105	117	150	122	122	145	104	127	109	56	100	84
MS	202	186	185	182	173	168	178	155	148	140	158	134	182
MT	.	459	.	453	.	471	.	423	.	169	.	307	.	276	.	289	.	272	.	298	.	324	.
NC	2	12	203	65	207	77	170	62	141	65	142	32	96	27
ND	150	.	146	.	128	.	149	.	141	.
NH	91	104	68	102	69	99	62	101	106	77	92	99
NJ	47	49	46	58	133	116	89	84	100	95	95	104	75
NM	20	42	28	51	29	57	29	55	19	30	13
NV	39	43	10	44	3	47	2	43	1	34	4
NY	30	122	368	82	367	37	241	14	411	342	456	223	373	393
OH	13	21	20	18	39	39	33	43	21	26	20	27	18	22	26	20
OK	128	130	149	145	159	159	158	140	141	163	169	308	142	300	134	248	121	272	105	289	101	340	157
OR	1	19	18	109	18	104	18	98	12	91	17
PA	166	152	266	247	264	7	324	257	307	260	308	216	186	171
RI	2	52	50	63	91	82	67	62	24	78	75
SC	90	58	98	111	100	97	81	95	97	54	90	118
SD	.	.	.	29	30	28	29	29	29	29	50	42	47	48	48	47	43	41	70	44	48	48	55
TN	80	73	229	254	213	239	199	243	229	265	284	254	303	333
TX	304	.	367	.	486	.	444	.	450	.
UT	22	58	58	95	93	93	90	101	103	105	105	96	98
VA	333	326	335	284	301	306	307	319	353	346	556	329	389
VT	14	29	14	13	14	13	17	11	9	8	9
WA	6	2	105	68	98	70	99	73	101	78	106	89	91	86
WI	70	25	31	23	26	28	24	20	9	12	21	17
WV	1	8	58	67	69	73	87	99	95	83	121	104	111	104
WY	10	29	37	45	46	71	55	80	48	46	52	37	34

F Newspaper Market–Legislative District Congruence Robustness Checks

Table F.1 – Newspaper Reader Share and Legislator Press Coverages. After controlling for legislator, election, and district variables, newspaper Reader Share strongly predicts observed press coverage. As a result, the Congruence between newspaper markets and districts is also highly predictive of legislative newspaper coverage.

		Count of Articles About Legislator (q_{mdt})				Sales-Weighted Count of Articles About Legislator (q_{dt})			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<div> <div>Legislator Controls</div> <div>Election Controls</div> <div>District Controls</div> <div>Additional Controls</div> </div>	Reader Share	108.57 (6.37)	119.47 (6.86)	109.39 (6.23)	119.39 (6.84)				
	Congruence					132.26 (2.58)	97.98 (2.59)	130.05 (2.57)	97.85 (2.59)
	Freshman		-1.58 (0.52)	-1.63 (0.53)	-1.59 (0.53)		-0.32 (0.29)	-0.30 (0.30)	-0.33 (0.29)
	Experience		0.65 (0.17)	0.62 (0.18)	0.64 (0.17)		0.23 (0.07)	0.30 (0.07)	0.22 (0.07)
	Chair		3.03 (0.99)	2.96 (0.99)	2.99 (0.99)		1.11 (0.39)	1.55 (0.41)	1.08 (0.39)
	Close Race		-0.72 (0.64)	-0.71 (0.65)	-0.72 (0.64)		-0.60 (0.28)	-0.52 (0.29)	-0.60 (0.28)
	Uncontested Race		-2.28 (0.57)	-2.26 (0.58)	-2.25 (0.58)		-1.46 (0.26)	-1.37 (0.27)	-1.43 (0.26)
	Open Seat		-5.01 (0.77)	-5.11 (0.79)	-5.02 (0.78)		-1.72 (0.31)	-1.50 (0.33)	-1.73 (0.31)
	Median Income		-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)		-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
	Population Density		-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)		-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
	% Urban		0.17 (0.04)	0.19 (0.04)	0.16 (0.04)		0.14 (0.01)	0.16 (0.01)	0.14 (0.01)
	% Retired		0.03 (0.23)	0.05 (0.23)	0.07 (0.22)		-0.16 (0.04)	-0.06 (0.05)	-0.13 (0.04)
	% Veterans		-0.70 (0.32)	-0.68 (0.33)	-0.71 (0.33)		-0.21 (0.06)	-0.52 (0.07)	-0.22 (0.06)
	% Foreign Born		0.40 (0.30)	0.43 (0.33)	0.44 (0.31)		0.25 (0.04)	0.15 (0.04)	0.28 (0.04)
	Total Circulation		4.11 (1.53)		4.10 (1.53)		3.37 (0.07)		3.37 (0.07)
	Distance to State Capital			-0.01 (0.02)	-0.01 (0.02)			-0.01 (0.00)	-0.01 (0.00)
	N	48,087	47,109	47,109	47,109	32,120	31,369	31,369	31,369
	Unit of Observation	District-Paper-Year				District-Year			
	State-Chamber-Year FEs	✓	✓	✓	✓	✓	✓	✓	✓

Note: Standard errors are clustered by district in parenthesis. The sales-weighted average number of articles about a legislator in district d in time t is $q_{dt} = \sum_{m=1}^M \text{MarketShare}_{mdt} \cdot q_{mdt}$. The definition of $q_{c dt}$ is analogous. Results are substantively similar after logging *ReaderShare* and *Congruence*.

G Electoral Selection Robustness Checks

In this section, I conduct two additional robustness checks on the midpoint method (Table 5). First, in Table G.1, I use CFscores from Bonica (2014) to measure Midpoint and Distance. Looking across the columns of Table G.1, I find strong evidence that Congruence increases with the Midpoint estimated using CFscores. In fact, the relative estimated effect of Congruence is substantially larger when using CFscores rather than HMH scores.

Second, while the addition of state-chamber-year fixed effects in Table 5 addresses concerns about omitted variable bias across time or between states and chambers, they do not ameliorate concerns that an observed confounder might be correlated with both Congruence and Democratic vote share across districts within a given chamber. To address this concern, Table G.2 replicates Table 5 after substituting in legislative district-regime fixed effects. This specification focuses on changes in Congruence within the same district across election cycles

Table G.1 – Press Coverage and the Advantage of Moderate Candidates in Contested General Elections Using CFscores. This table replicates 5 using CFscores from Bonica (2014) to measure Midpoint and Distance.

	Dem. Vote Share				
	(1)	(2)	(3)	(4)	(5)
Midpoint	0.35 (0.01)	0.28 (0.02)	0.28 (0.02)	0.27 (0.02)	0.29 (0.03)
Midpoint · Congruence		0.43 (0.07)	0.44 (0.07)	0.37 (0.07)	0.27 (0.09)
Congruence		-0.20 (0.03)	-0.22 (0.04)	-0.19 (0.04)	-0.16 (0.05)
Distance	-0.04 (0.01)	-0.03 (0.01)	-0.04 (0.01)	-0.04 (0.01)	-0.06 (0.02)
Distance · Congruence			0.05 (0.05)	0.05 (0.05)	0.12 (0.06)
Rep. Pres. Vote Share	-0.75 (0.01)	-0.76 (0.01)	-0.76 (0.01)	-0.75 (0.01)	-0.77 (0.01)
Rep. Primary Contributions				-0.00 (0.00)	-0.00 (0.00)
Dem. Primary Contributions				0.00 (0.00)	0.00 (0.00)
N	21,743	21,743	21,743	21,743	11,385
State-Chamber-Year FEs	✓	✓	✓	✓	✓
District, Legislator, and Election Controls	Yes	Yes	✓	✓	✓

Note: The outcome is either Democratic vote share or a Democratic win indicator. Robust standard errors are clustered by district in parentheses. Midpoint and Distance variables are scaled to run from 0 to 1. The sample is limited to contested general elections in single member districts.

Table G.2 – Press Coverage and the Advantage of Moderate Candidates in Contested General Elections Using District Fixed Effects. This table replicates 5 using district-regime fixed effects to hold the unobserved median constant.

	Dem. Vote Share				
	(1)	(2)	(3)	(4)	(5)
Midpoint	0.16 (0.02)	0.07 (0.03)	0.07 (0.03)	0.08 (0.03)	0.07 (0.05)
Midpoint · Congruence		0.27 (0.15)	0.20 (0.15)	0.19 (0.14)	0.19 (0.24)
Congruence		-0.26 (0.09)	-0.09 (0.10)	-0.10 (0.10)	-0.03 (0.16)
Distance	0.00 (0.01)	-0.01 (0.02)	0.04 (0.02)	0.03 (0.02)	0.05 (0.04)
Distance · Congruence			-0.33 (0.12)	-0.30 (0.11)	-0.42 (0.17)
Rep. Pres. Vote Share	-0.60 (0.01)	0.38 (0.03)	0.38 (0.03)	0.39 (0.03)	0.41 (0.04)
Rep. Primary Contributions				-0.00 (0.00)	-0.00 (0.00)
Dem. Primary Contributions				0.00 (0.00)	0.00 (0.00)
N	7,986	7,986	7,986	7,986	4,475
District FEs	✓	✓	✓	✓	✓
District, Legislator, and Election Controls	✓	✓	✓	✓	✓

Note: The outcome is either Democratic vote share or a Democratic win indicator. Robust standard errors are clustered by district in parentheses. Midpoint and Distance variables are scaled to run from 0 to 1. The sample is limited to contested general elections in single member districts.

(but within the same redistricting period), and further mitigates concerns about confounding from district-level characteristics. If anything, the results using this specification are larger than the baseline model, suggesting that the observed effects of Congruence on Democratic vote share are not driven by static, unobserved district-level characteristics. However, because there is less variation in Congruence within a district, these results are estimated with more noise than my baseline specification.

H Regression Discontinuity Balance Tests

Table H.1 – Regression Discontinuity Design Balance Tests. This table reports results estimates from a local linear regression of the variable in the “Outcome” column on the running variable, a treatment indicator, and the interaction of the two using the optimal bandwidth from Calonico, Cattaneo, and Titiunik (2014). No evidence of imbalance in key covariates is found.

Outcome	Estimate	Std. Error	t	p-value
Lagged Legislative Vote Share	-0.00	0.01	-0.04	0.97
Lagged Presidential Vote Share	0.00	0.00	0.09	0.93
Lagged Congruence	-0.01	0.01	-1.01	0.31
Lagged NP-Score	-0.01	0.04	-0.26	0.80
Year	0.10	0.27	0.38	0.70

Note: Robust standard errors clustered by district-regime in parentheses.

I Productivity Robustness Checks

Since the missed vote and sponsorship rate may be correlated with travel time to the capital, in Table I.1 I add a control for the distance between each district's centroid and the state capital. My results are unchanged following this inclusion.

Table I.1 – Active Newspaper Coverage Increases Legislative Productivity. Active newspaper coverage is associated with fewer missed roll-call votes, more bill sponsorships, and more-active committee membership.

	Percent of Floor Votes Missed		Number of Bills Sponsored		Probability on Power Committee	
	(1)	(2)	(3)	(4)	(5)	(6)
Congruence	-1.42 (0.33)	-1.35 (0.34)	9.91 (3.79)	7.76 (3.40)	0.06 (0.02)	0.05 (0.02)
N	37,312	37,312	37,312	37,312	47,324	47,324
Average Outcome	3.3	3.3	27	27	.45	.45
State-Chamber-Year FEs	✓		✓		✓	
State-Chamber-Year-Party FEs		✓		✓		✓
District, Legislator, and Election Controls	✓	✓	✓	✓	✓	✓
Distance to Capital Control	✓	✓	✓	✓	✓	✓

Note: Outcomes are reported in column headers. Standard errors are clustered by district in parentheses.

J Non-Parametric Estimates of Multiplicative Interactions

Hainmueller, Mummolo, and Xu (2019) show that multiplicative interaction models—including Tables 5, 6, and 8—may yield misleading results if researchers incorrectly assume linearity in effect or common support of the moderating variable (i.e., Congruence). In response, Figure J.1 reports the diagnostic measures proposed by Hainmueller, Mummolo, and Xu (2019) and implemented using the R package *Interflex* for every analysis in the main article that employs a multiplicative interaction term.

Each diagnostic figure below divides the moderator into three bins—representing low, medium, and high values—and estimates the conditional marginal effects of the key independent variable within each bin. This approach relaxes the linear interaction effect assumption, allowing the marginal effects to vary non-linearly across bins, and ensures that the estimated effects rely only on observed data, mitigating extrapolation beyond the support of the independent variable.

Looking at the figures, we observe a strong linear relationship between the binned estimates and the moderator (i.e., the red point estimates are very close to the black line). We also observe strong overlap in the moderator across values of the independent variable. In short, the assumptions of the multiplicative interaction model appear to hold, and after using an alternative setup to explore effect heterogeneity, my results are highly similar.

Figure J.1 – Marginal Effects Plots for Multiplicative Interaction Models Using *Interflex*.

