

A Average Number of Page Hits Per Newspaper for Each Item by Decade

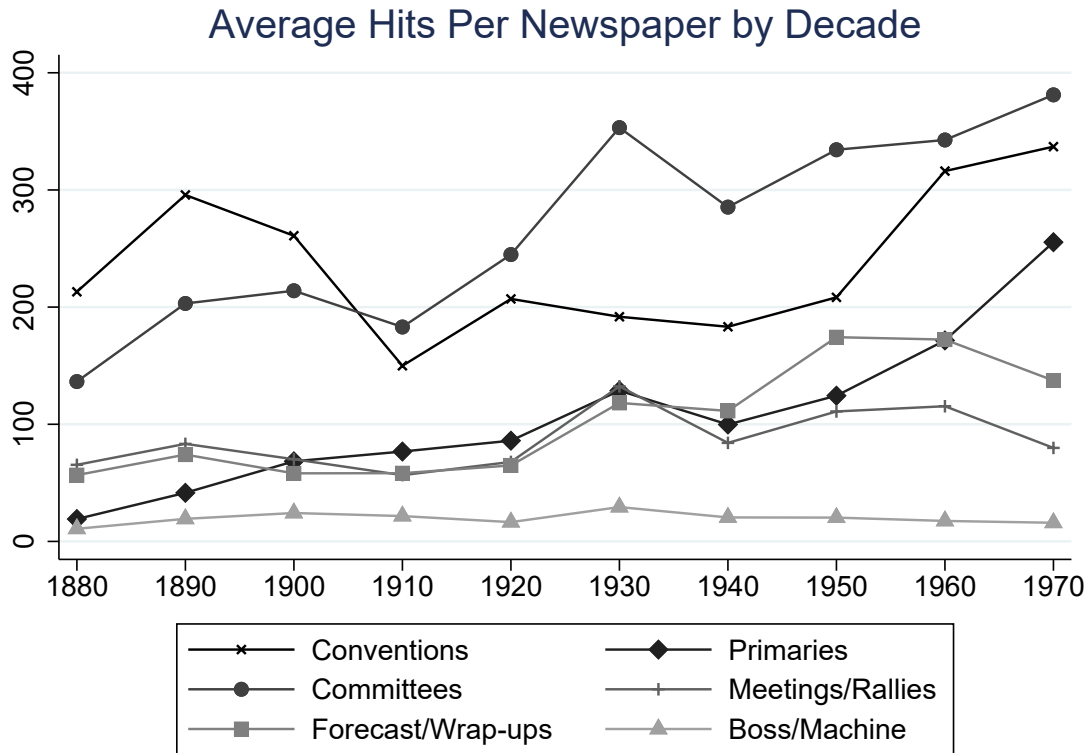


Figure A1: Average Number of Page Hits Per Newspaper for Each Item by Decade

Sources: Newspapers.com

B Robustness Checks for “Era of Partisan Press, 1880–1900” Section

Here we show tables on 1880–1900 period with the set of newspapers restricted to the multi-decade sample.

Table B1: Newspaper Partisan Content, 1880 to 1900, Multi-Decade Sample

Item	<i>RCS</i> in R Papers	<i>RCS</i> in D Papers	Partisan Gap	Number of Obs.	Number of Papers
Committees	0.59	0.39	0.20	1932	284
Meetings/Rallies	0.73	0.35	0.38	1535	255
Conventions	0.59	0.38	0.21	2089	287
Primaries	0.70	0.30	0.41	930	206
Forecasts/Wrap-ups	0.59	0.40	0.19	1539	251
Boss/Machine	0.65	0.23	0.42	1329	250
Combined Index	0.64	0.34	0.29	1889	276

Number of Observations is the number of newspaper-years used in calculating the RCS for either Democratic or Republican newspapers. Number of Papers is the number of newspapers that are used at least once.

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son’s American Newspaper Annual* (various years)

Table B2: **Newspaper Within-State Partisan Gap, 1880 to 1900, Multi-Decade Sample**

Item	Within-State Partisan Gap	Number of Obs.	Number of Papers	Number of States
Committees	0.12	159	213	19
Meetings/Rallies	0.33	139	191	17
Conventions	0.14	164	217	20
Primaries	0.23	91	184	14
Forecasts/Wrap-ups	0.16	140	203	18
Boss/Machine	0.47	138	201	17
Combined Index	0.24	157	204	18

Number of Observations is the number of state-years used in calculating the average Within-State Partisan Gap. Number of Papers is the number of newspapers that are used at least once. Number of States is the number of states that are used at least once.

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

C Robustness Checks for “Trends” Section

C.1 Trends with Fixed Samples of Newspapers

In the main analysis in the paper we restrict the sample to long-lived newspapers that existed for at least 50 years of our 100-year period of investigation. However, even in these figures, there is some entry and exit into the sample. Here we examine the trends in the partisan gap within overlapping 50-year windows, holding the sample of newspapers approximately fixed within each window. For each window, we keep a newspaper if it is present for at least 42 of the 50 years (to allow for gaps), as long as it is present in the first and last year of the window. The 50 year windows are 1880–1930, 1890–1940, 1900–1950, 1910–1960, 1920–1970 and 1930–1980.

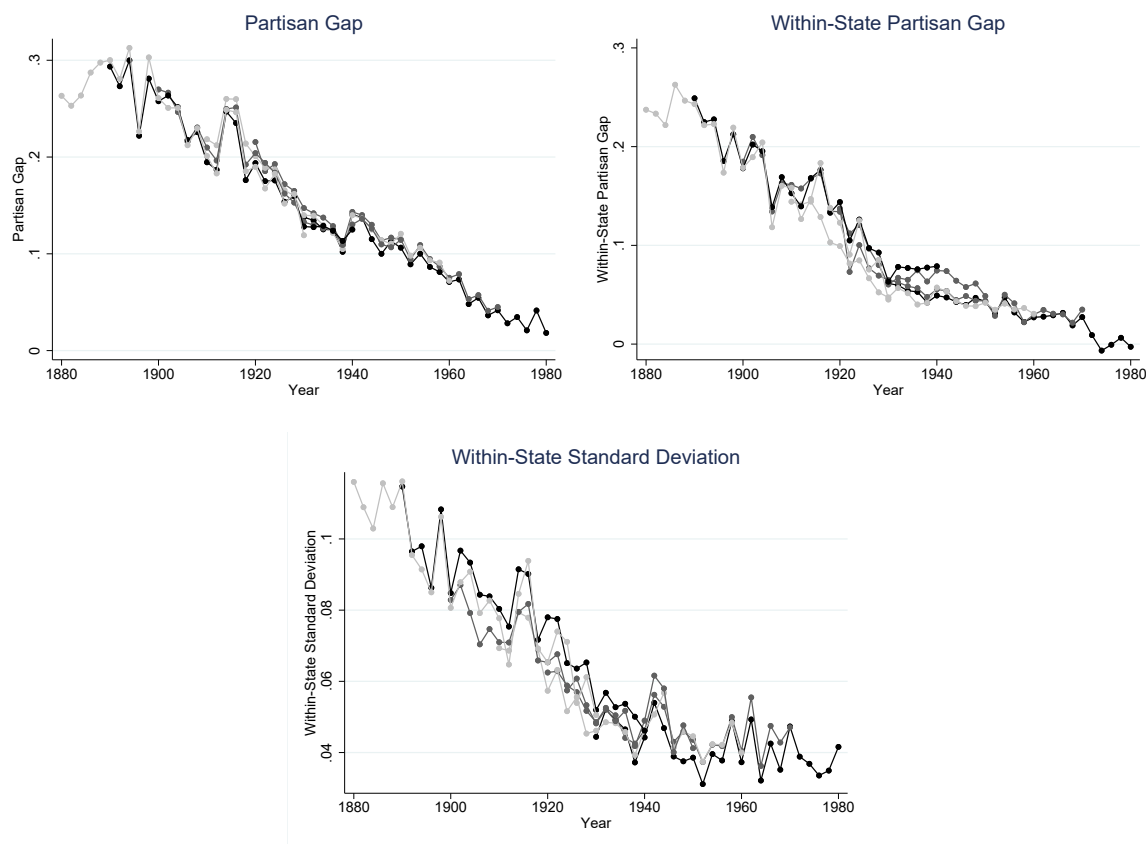


Figure C2: Fixed Overlapping 50 Year Samples

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

C.2 Different Thresholds for Multi-Decade Sample

Here we show that the main patterns documented in the paper are robust to using different thresholds for including newspapers in the multi-decade sample. Recall that the threshold is defined as the minimum number of years for which we can compute RCS for the *Combined Index*.

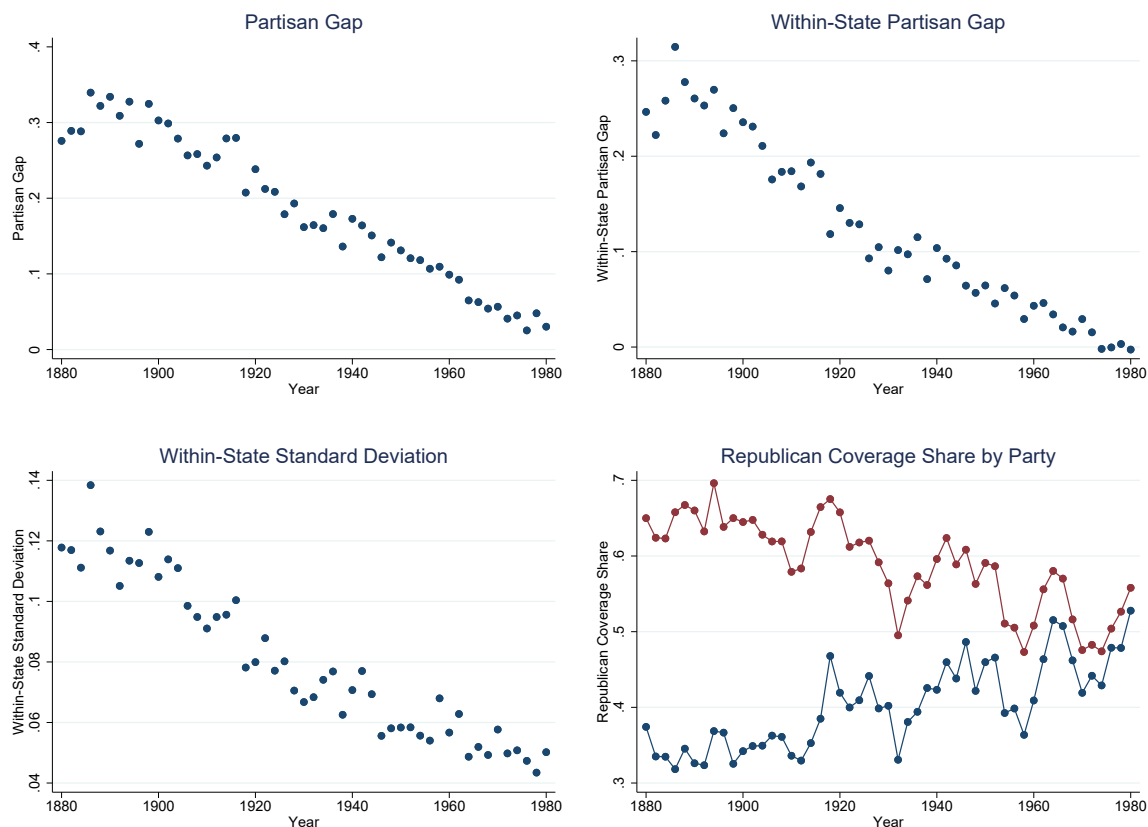


Figure C3: Combined Index, Multi-Decade Sample with Threshold=30 Years

Sources: Newspapers.com, ICPSR 30261, *N. W. Ayer & Son's American Newspaper Annual* (various years)

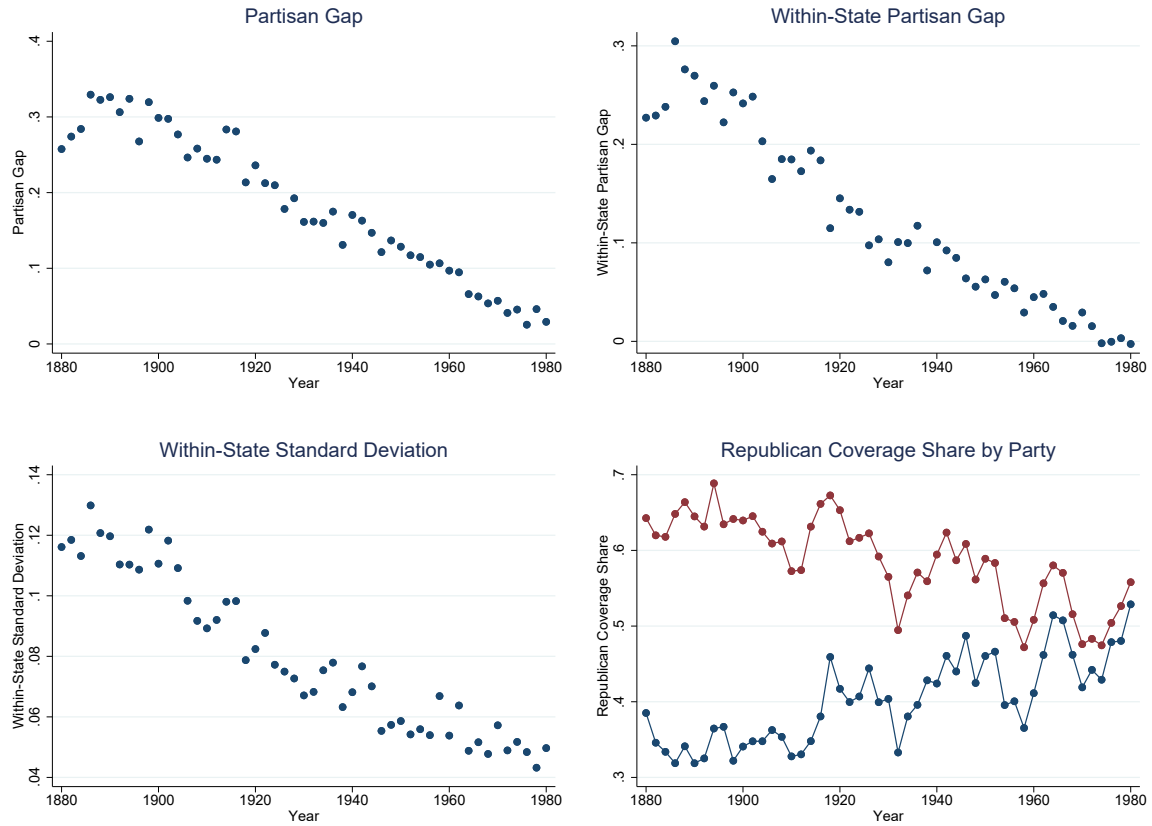


Figure C4: Combined Index, Multi-Decade Sample with Threshold=40 Years

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

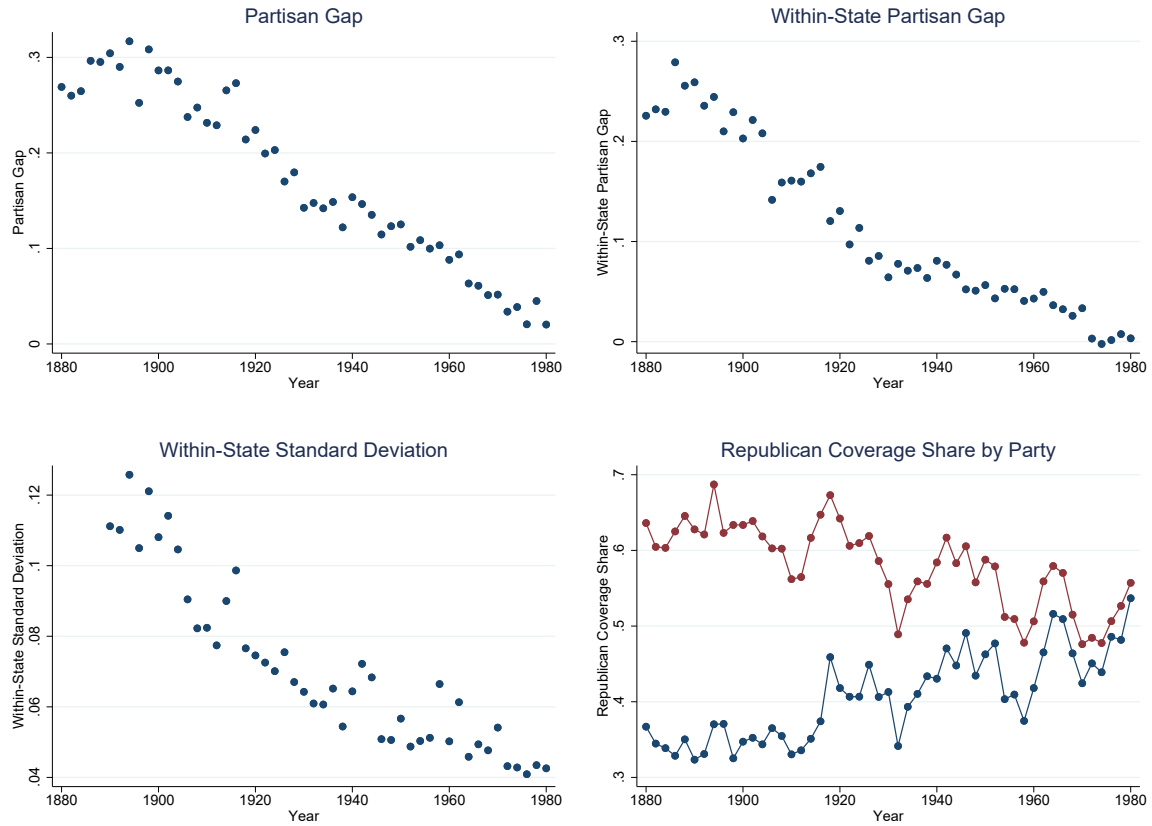


Figure C5: Combined Index, Multi-Decade Sample with Threshold=60 Years

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

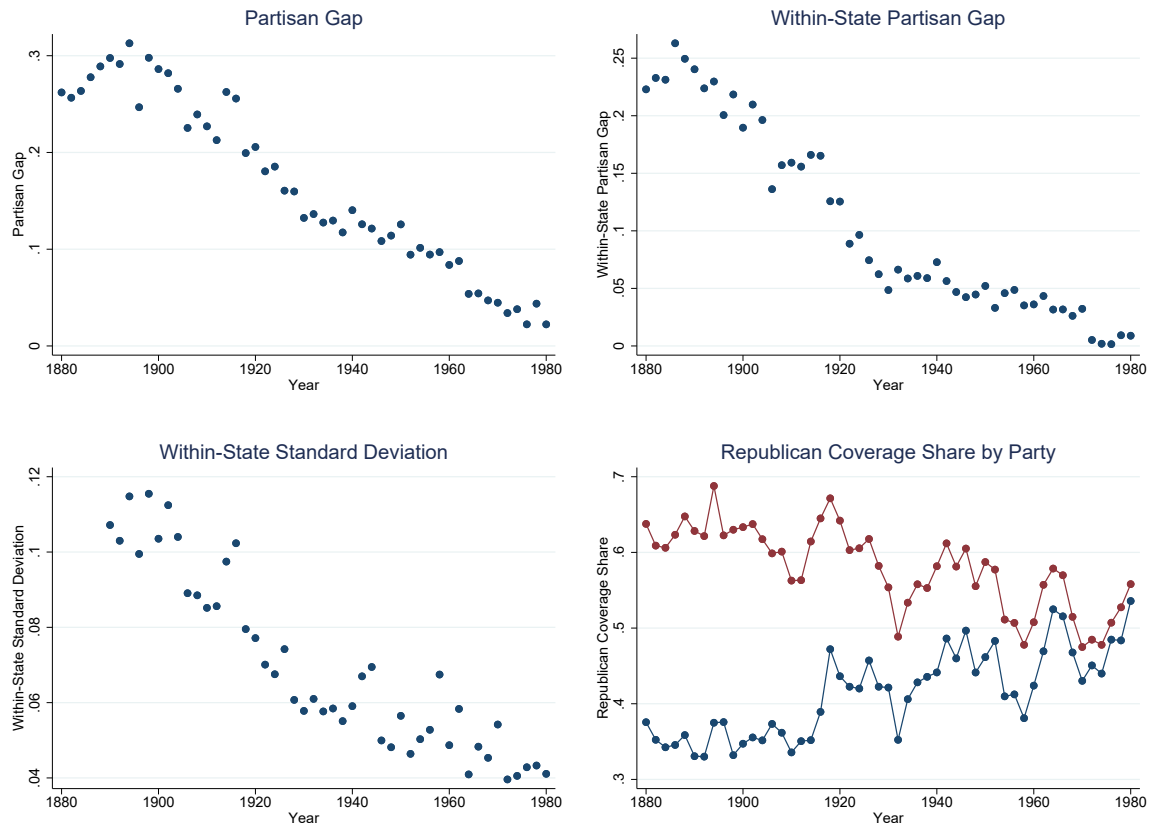


Figure C6: Combined Index, Multi-Decade Sample with Threshold=70 Years

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

C.3 Within-State Standard Deviation Using All Available Newspapers

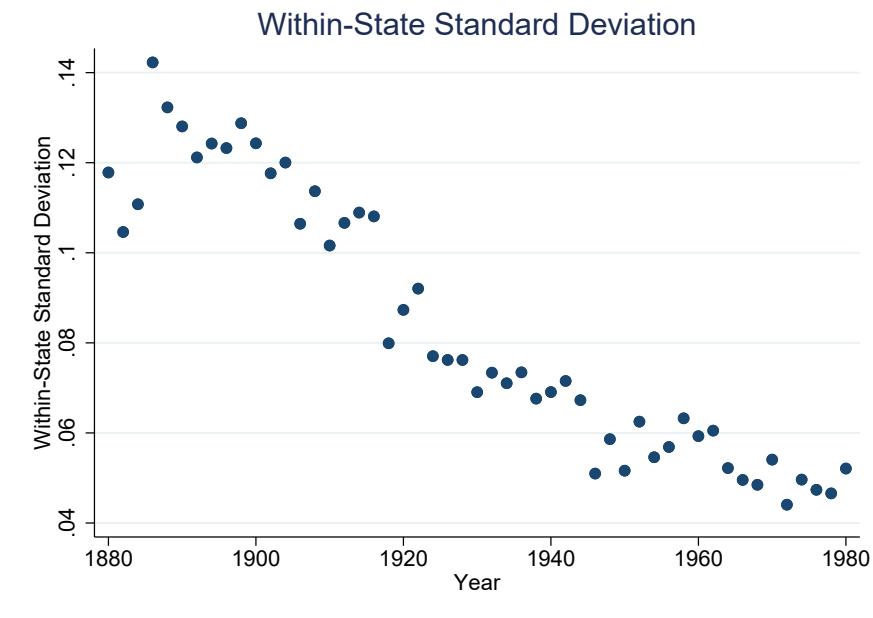


Figure C7: Average Standard Deviation, Combined Index, All Available Newspapers

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

C.4 Trends for Individual Items

Here we present figures of the four main measures analyzed in the main text, for each of the six newspaper content items separately. Figure C8 presents the *PG*. Figure C9 shows the *WSPG*. Figure C10 displays the within-state standard deviation. Finally, Figure C11 shows the *RCS* for Democratic and Republican newspapers separately.

Note that since “boss” and “machine” are terms with negative connotations, a “pro-Republican” pattern of coverage would use these terms more in conjunction with the Democratic party than the Republican party. This is why the bottom right panels of Figures C8 and C9 are negative and tend towards zero over time.

In Figure C10, the Convention item exhibits an on-year/off-year presidential cycle due at least in part to the substantial coverage of the Democratic and Republican national

conventions, which are held only in presidential years.

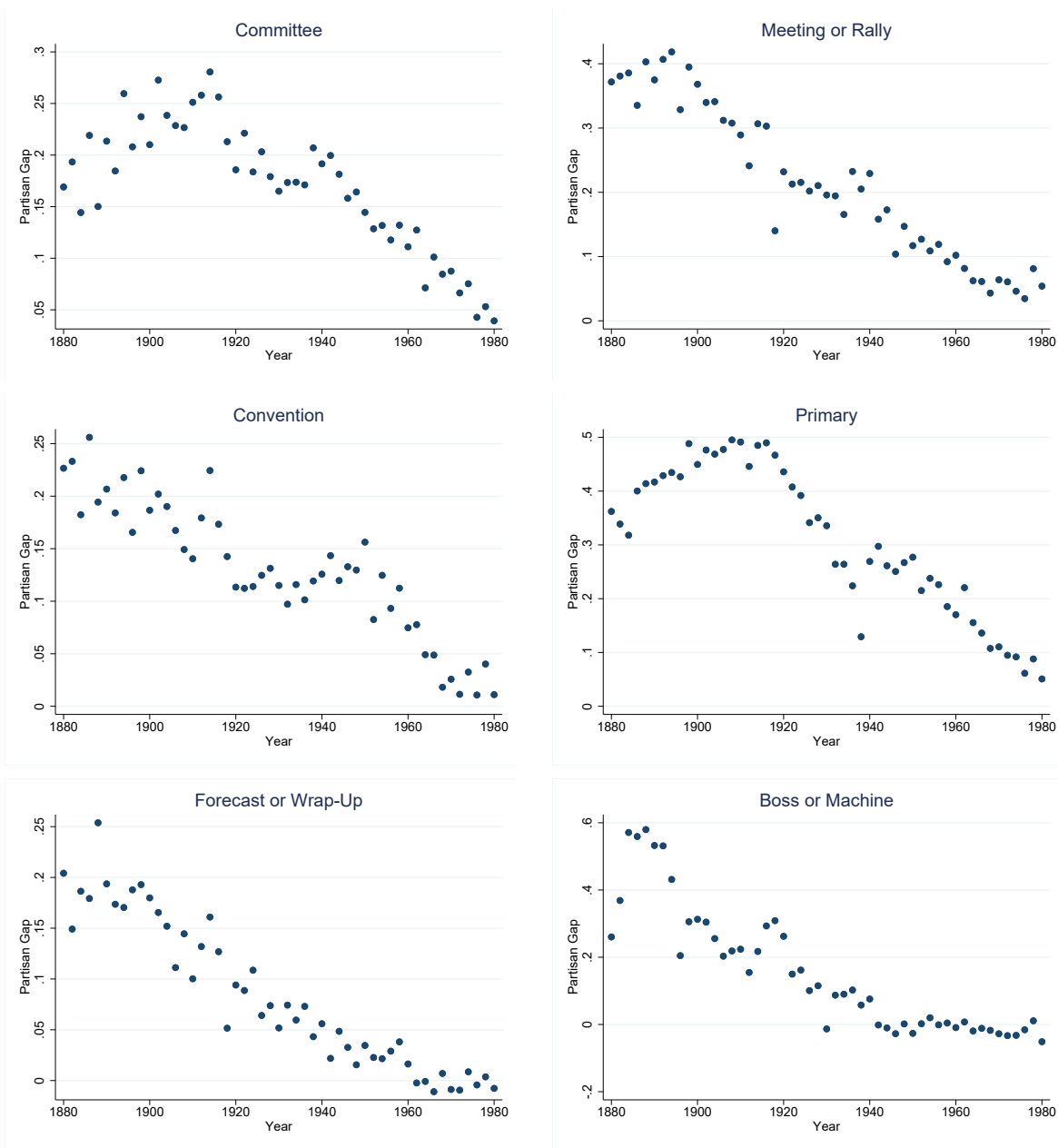


Figure C8: Partisan Gap Over Time, by Item

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

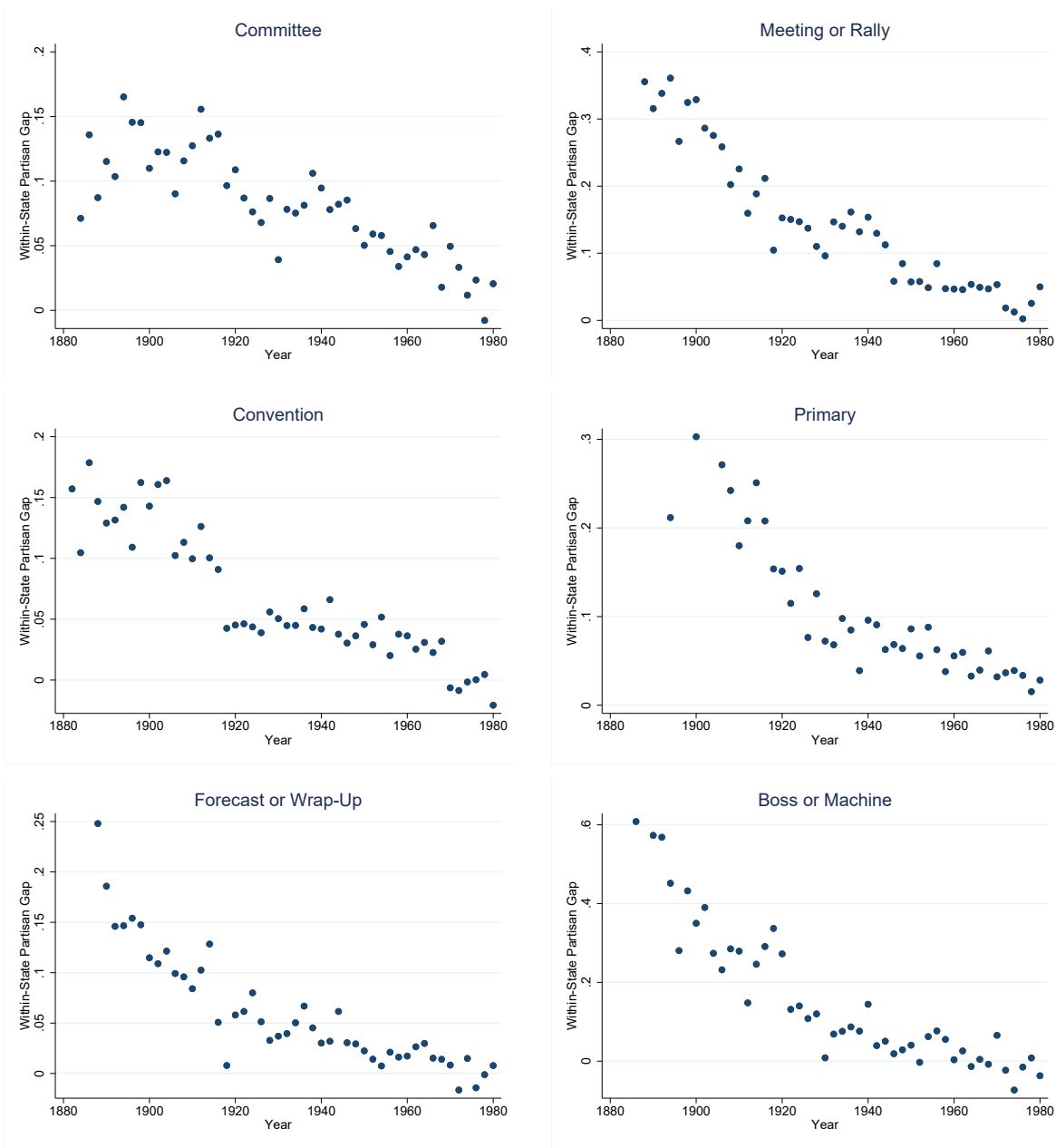


Figure C9: Within-State Partisan Gap Over Time, by Item

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

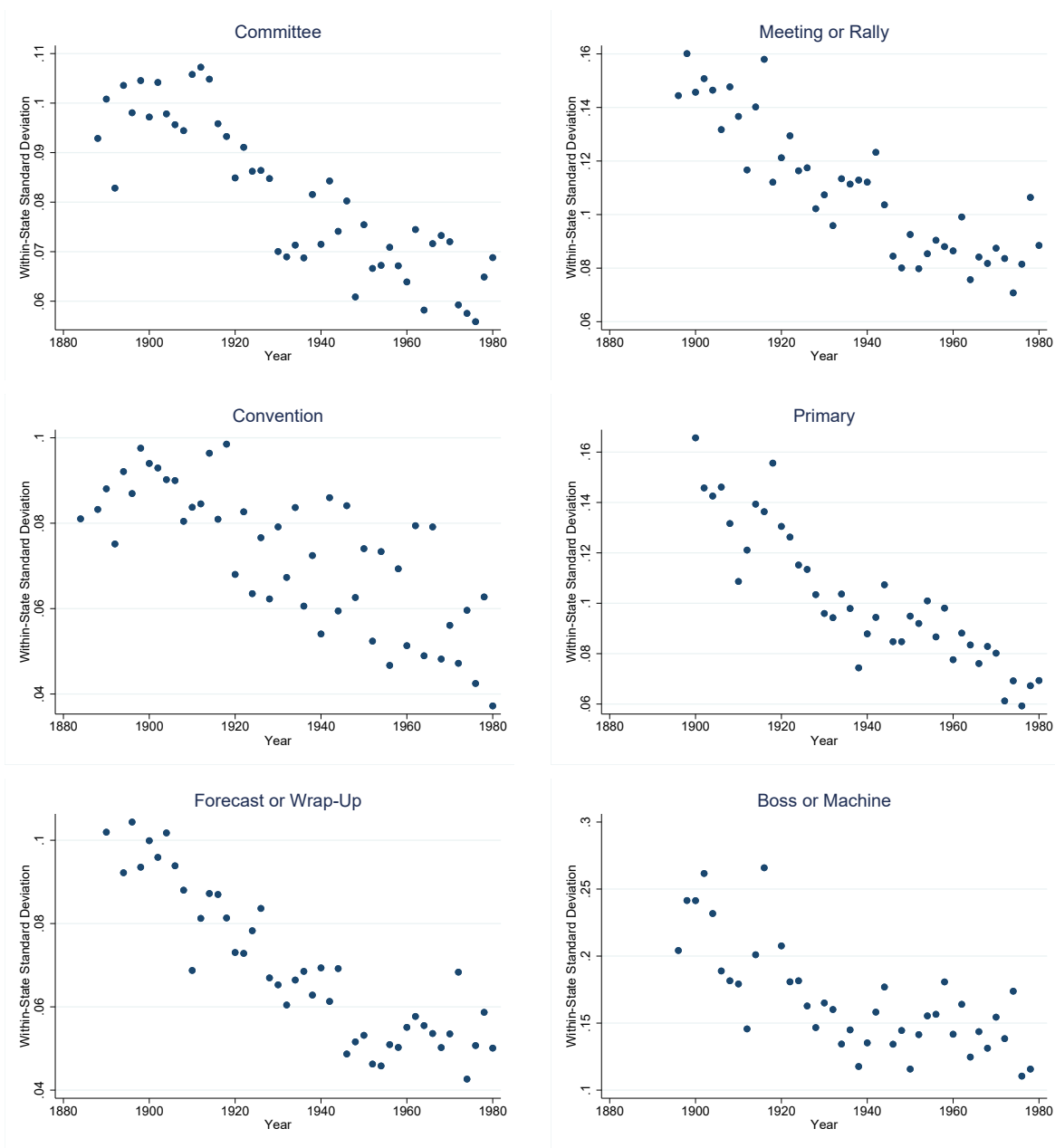


Figure C10: Within-State Standard Deviation Over Time, by Item

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

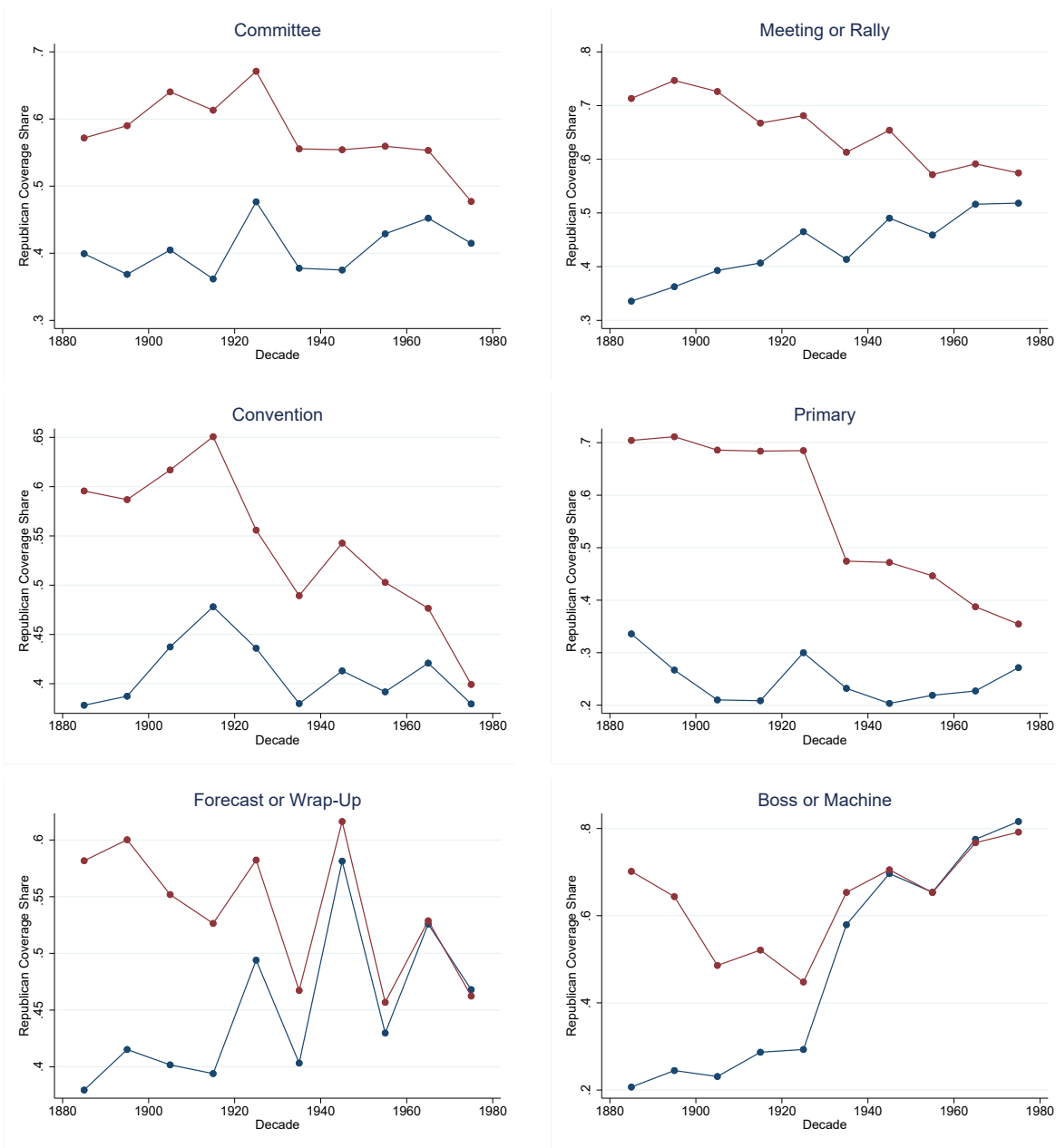


Figure C11: Republican Coverage Share by Party and Item

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

D Counting Newspaper Mentions: Pages vs. Articles

In *Newspapers.com* the searches are done by *page*. Our counts therefore provide a crude estimate of the “true” amount of coverage we hope to capture in our measures (D_{ijt} and R_{ijt} above), and might even lead to measurement error in our relative measures based on shares, *Republican Coverage Share* $_{ijt}$. Consider, for example, a newspaper that prints three stories about the upcoming Republican convention, all on the same page, and one story about the upcoming Democratic convention. Then *RCS* would be 0.5 in this case since the Republican stories would only be counted once. If the stories were of similar length and importance, then an *RCS* of 0.75 would be a more accurate measure of partisan behavior. On the other hand, if the Democratic story was three times as long and detailed, then 0.5 might be better.

Another popular newspaper archive, in which searches are done by article rather than page, is *Proquest Historical Newspapers*. Unfortunately, the *Proquest* archive only contains about two dozen general interest newspapers, versus more than 2000 in *Newspapers.com*. Also, almost all of the newspapers in *Proquest* serve major U.S. cities. Therefore, we cannot use the *Proquest* archive as our main data source, since it is not large enough or representative enough for our purposes.

Fortunately, since most of the *Proquest* newspapers are also in *Newspapers.com*, we can directly compare measures based on page counts versus those based on article counts to see whether there are large systematic differences between them. The list of papers and years of overlap is shown in Appendix Table D3.

The results are encouraging. Figure D12 presents scatter plots of *RCS*. The y-axis is based on page counts from *Newspapers.com* and the x-axis is based on article counts from *ProQuest*. The panel on the left displays the scatter plot pooling all possible newspaper-year-item observations and the panel on the right shows the plot of the newspaper-years for the Combined Index. The figure also presents 45-degree lines. In both cases, the two measures are highly correlated. Pooling all newspaper-year-item observations, the correlation is 0.95. The correlations are also high within newspaper-item. The average and median correlations are 0.91 and 0.94, respectively. The correlation for the *Combined Index* is 0.95.

Sources: *Newspapers.com*, ICPSR 30261, *N.W. Ayer & Son’s American Newspaper Annual* (various years), *ProQuest Historical Newspapers*TM

Finally, we can conduct an analysis of the over time patterns similar to those above,

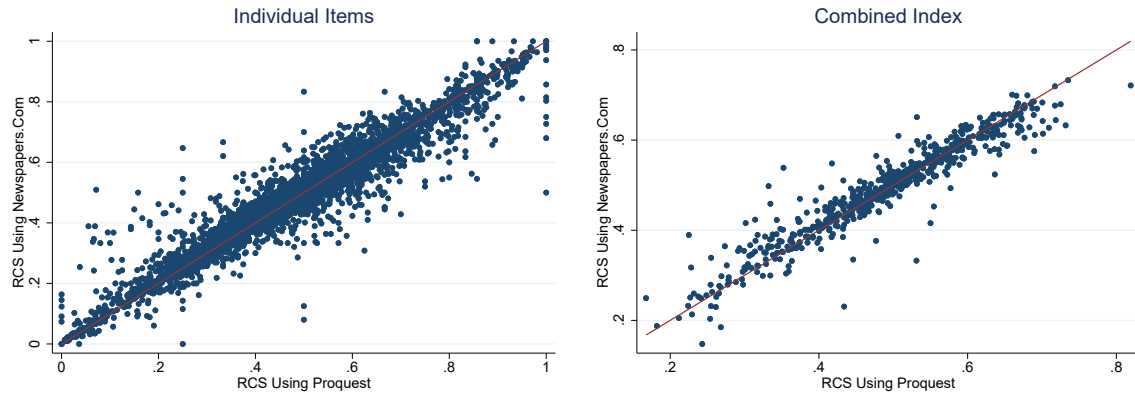


Figure D12: Newspaper-Year Article Hits versus Page Hits

comparing the *Newspapers.com* and *ProQuest* samples. This analysis is only suggestive, since there are only six Democratic newspapers and four Republican newspapers that are in both archives and in our multi-decade sample. Despite the small size of the overlapping sample, the trends in the *PG* for the *Combined Index* for both measures are similar to that in Figure 1 based on the full multi-decade sample. That is, both exhibit long downward trends between 1900 and 1980 (see Appendix Figure D13).

Table D3: **ProQuest and Newspapers.com Overlap**

Newspaper Name	Overlapping Years
The Atlanta Constitution	1880 to 1980
The Austin American Statesman	1880 to 1980
The Baltimore Sun	1880 to 1980
The Boston Globe	1880 to 1980
Chicago Tribune	1880 to 1980
Cincinnati Enquirer	1880 to 1922
Detroit Free Press	1880 to 1980
Hartford Courant	1880 to 1980
Los Angeles Times	1881 to 1980
Louisville Courier Journal	1880 to 1980
The Nashville Tennessean	1910 to 1922
The New York Times	1880 to 1980
New York Tribune	1880 to 1922
Philadelphia Inquirer	1880 to 1980
Pittsburgh Courier	1911 to 1976
San Francisco Chronicle	1880 to 1922
St. Louis Post Dispatch	1880 to 1980
Wall Street Journal	1888 to 1922
The Washington Post	1899 to 1922

Sources: Newspapers.com, ProQuest Historical Newspapers™

Figure D13 presents the *PG* for the *Combined Index* over time. The Republican papers included in this measure are the *Hartford Courant*, *Chicago Tribune*, *Los Angeles Times*, and *Philadelphia Inquirer*. The Democratic newspapers included are the *Atlanta Constitution*, *Austin American Statesman*, *The Boston Globe*, *Detroit Free Press*, *Louisville Courier Journal*, and *St. Louis Post Dispatch*. The left-hand side panel shows the gap using *Newspapers.com* page counts, while the panel on the right shows the gap based on *ProQuest* article counts. Both graphs reveal the same basic pattern, a long downward trend and a drop between 1960 and 1962.

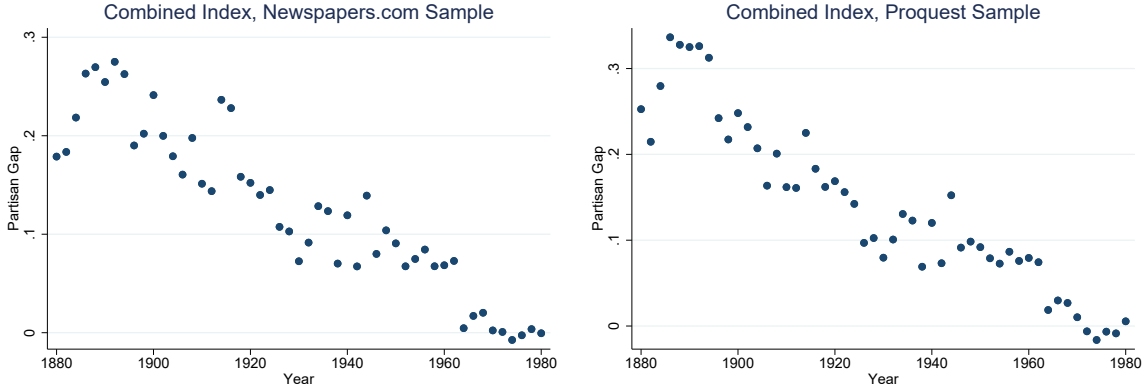


Figure D13: Partisan Gap Using Articles versus Pages to Measure Coverage

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years), ProQuest Historical NewspapersTM

E Potential Bias From OCR Errors

The quality of Optical Character Recognition (OCR) varies across newspapers and years. It is possible that systematic errors could bias our measures of newspaper partisanship. However, as we show here, to account for our main substantive findings the errors must follow particular patterns. Empirically, it appears that the actual patterns are quite different. Thus, we doubt that OCR errors can be the main driver of our results.

Recall that we measure PG as $PG_t = RCS_{Rt} - RCS_{Dt}$. Suppose that due to OCR errors the measured gap is $\tilde{G}_t = \widetilde{RCS}_{Rt} - \widetilde{RCS}_{Dt}$. Clearly the measured gap is greater than the true gap only if $\widetilde{RCS}_{Rt} > RCS_{Rt}$ or $\widetilde{RCS}_{Dt} < RCS_{Dt}$ or both.

Recall that for each newspaper i , $RCS_{it} = R_{it}/(R_{it} + D_{it})$, where R_{it} and D_{it} equal the number of pages with Republican and Democratic items, respectively. Let $(1 - \gamma_{it}^R)$ be the error rate in newspaper i on Republican items and $(1 - \gamma_{it}^D)$ be the error rate on Democratic items. Then the measured hits are: $\tilde{R}_{it} = \gamma_{it}^R R_{it}$ and $\tilde{D}_{it} = \gamma_{it}^D D_{it}$. Now, $1/RCS_{it} = 1 + D_{it}/R_{it}$ and $1/\widetilde{RCS}_{it} = 1 + \tilde{D}_{it}/\tilde{R}_{it}$. Evidently, $RCS_{it} > \widetilde{RCS}_{it}$ iff $1/RCS_{it} < 1/\widetilde{RCS}_{it}$ iff $D_{it}/R_{it} < \tilde{D}_{it}/\tilde{R}_{it} = \gamma_{it}^D D_{it}/\gamma_{it}^R R_{it}$. Thus, $\widetilde{RCS}_{it} > RCS_{it}$ iff $\gamma_{it}^D > \gamma_{it}^R$. Since \widetilde{RCS}_{Rt} is the average of \widetilde{RCS}_{it} across all Republican newspapers and RCS_{Rt} is the average of RCS_{it} across all Republican newspapers, if $\gamma_{it}^D < \gamma_{it}^R$ for all Republican newspapers i at time t , then $\widetilde{RCS}_{Rt} > RCS_{Rt}$. Similarly, if $\gamma_{it}^D < \gamma_{it}^R$ for all Democratic newspapers i at time t , then

$$\widetilde{RCS}_{Dt} > RCS_{Dt}.$$

Recall that the error rate in newspaper i on Republican items is $(1 - \gamma_{it}^R)$ and on Democratic items is $(1 - \gamma_{it}^D)$. Thus, if $(1 - \gamma_{it}^D) > (1 - \gamma_{it}^R)$ for all Republican newspapers i at time t and $(1 - \gamma_{it}^D) < (1 - \gamma_{it}^R)$ for all Democratic newspapers, then $\widetilde{G}_t > G_t$. More generally, \widetilde{G}_t will tend to be larger than PG_t when Republican newspapers systematically make more errors on Democratic items than on Republican items, and/or Democratic newspapers do the opposite.

Empirically, we cannot observe the true OCR error rates. However, we can observe the frequencies of many common errors. Some of the most common OCR errors involve confusion between “i” and “l”, “c” and “e”, “m” and “nn” or “rn”, “cl” and “d”, “vv” and “w”, “d” and “ol”, and “B” and “R”. Other errors involve hyphenation or spacing. In fact in the newspapers.com OCR text the hyphenation/spacing errors appear to be more common than errors due to commonly mistaken letters.

To check whether Democratic and Republican newspapers appear to have different error rates for Democratic and Republican terms, we searched for each of the following “mistakes” in the words Democratic or Republican in our search strings: Dcmocratic, Dennocratic, Demoeratic, Democratlc, Democratie, De-mocratic, Dem-ocratic, Demo-cratic, Democ-ratic, Democr-atic, Democra-tic, Democrat-ic, and Bepublican, Rpublican, Repubiican, Republlican, Republician, Re-publican, Rep-ublican, Republican, Repub-lican, Republ-ican, Republi-can.⁵⁵ For example, for the phrases involving “committees” we ran one search on the string [Democratic committee] OR [Dcmocratic committee] OR [Dennocratic committee] ... OR [Democrat-ic committee], to generate *Hits Including OCR Errors*, and a separate search on the string [Democratic committee] to generate *Hits On Desired String*.

We then calculate the fraction of pages missed by not incorporating the various OCR errors as $1 - \text{Hits On Desired String} / \text{Hits Including OCR Errors}$. Let E_{Dt}^D be the fraction of missed pages for Democratic newspapers on the searches involving the Democratic party and let E_{Dt}^R be the fraction of missed pages for Democratic newspapers on the searches involving the Republican party, and define E_{Rt}^D and E_{Rt}^R analogously for Republican newspapers. The four E 's provide estimates of the $(1 - \gamma)$'s above.

There are two important points. First, the error rates are relatively small – between 0.02 and 0.03 for the 1880 to 1910 period and between 0.01 and 0.02 for the 1960 to 1980 period – and are therefore unlikely to account for much of the observed change in PG . Second, the error rates in Democratic and Republican newspapers are similar to one

another. In both types of papers, the error rates are higher for Republican hits, and by roughly the same amount. Thus, the bias in the measured PG is even smaller, in fact it is miniscule. The PG based on the target strings alone in the early period is 0.1592 and in the later period is 0.0616.⁵⁶ If instead we calculate the PG using the target strings along with the OCR errors listed above, we find 0.1586 and 0.0621, respectively.

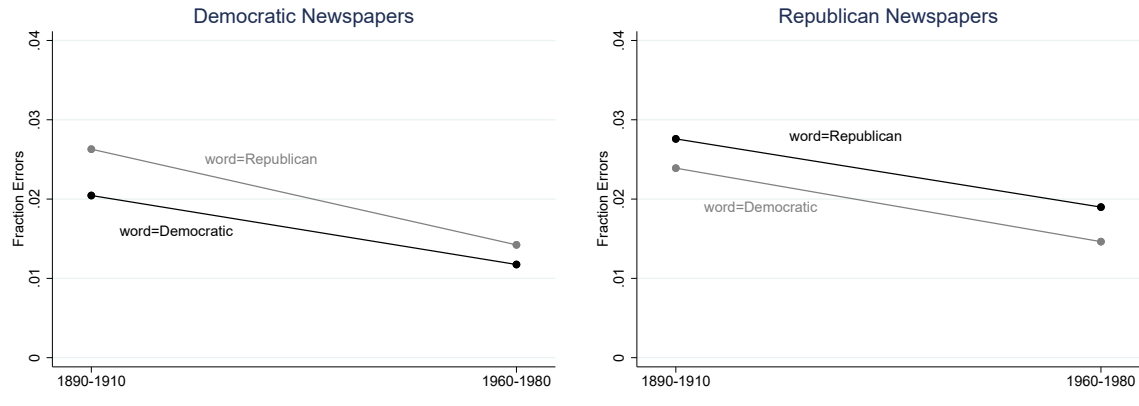


Figure E14: OCR Error Rates by Newspaper Type

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years)

F Robustness Checks for “Tone-Based Partisan Differences” Section

Here we show that partisan gap on specific issues restricting attention to the multi-decade sample.

Table F4: Newspaper Partisan Content on Selected Issues, Multi-Decade Sample

Item	<i>RCS</i> in R Papers	<i>RCS</i> in D Papers	Partisan Gap	Number of Obs.
<i>Multi-Decade Sample</i>				
Elections Bill, 1890–1892	0.60	0.31	0.29	200
Court Reform Plan, 1937	0.50	0.29	0.21	265
Health Insurance, 1949–1950	0.66	0.64	0.02	295
Bay of Pigs, 1962–1964	0.14	0.13	0.01	271

Number of Observations is the number of newspaper-years used in calculating the *RCS* for either Democratic or Republican newspapers.

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son’s American Newspaper Annual* (various years)

G Robustness Checks for “Possible Explanations” Section

G.1 Using 5th-Order Polynomials

This shows the analog to Figure 6 in the text but using 5th-order polynomials of the year trend and controls, rather 3rd-order polynomials.

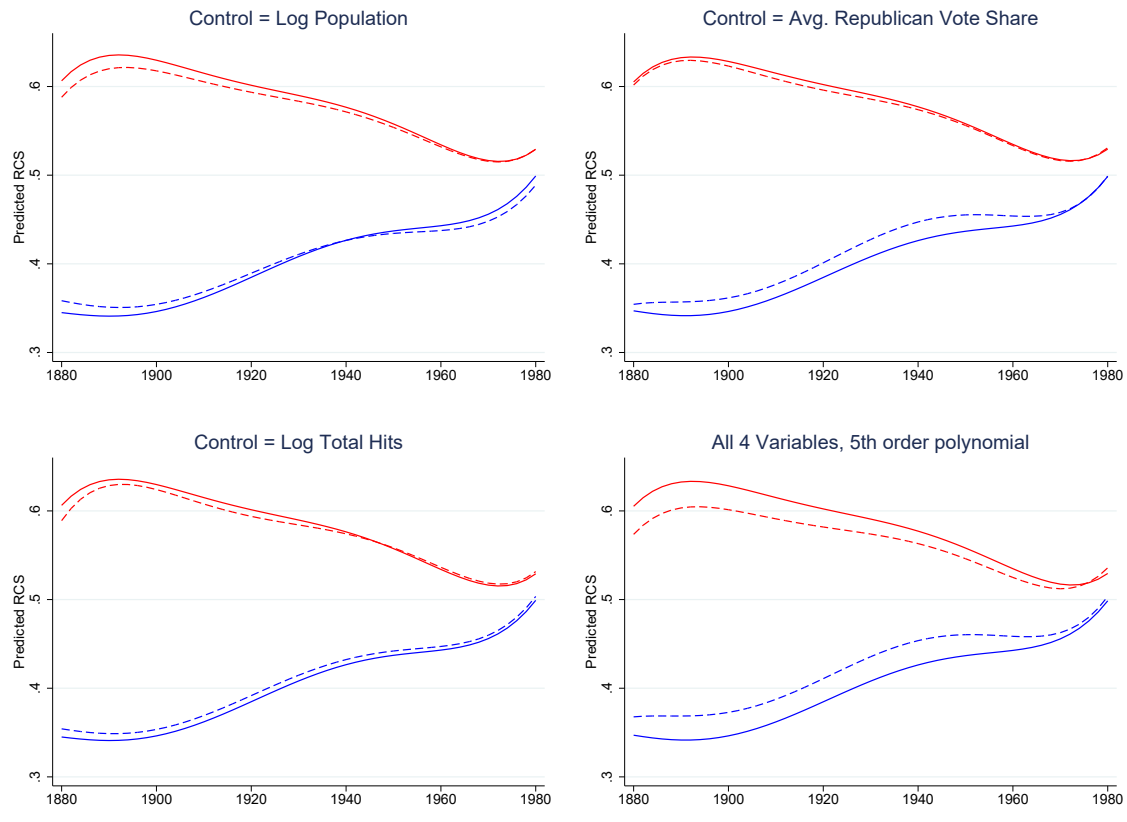


Figure G15: Predicted *RCS* of Combined Index, With and Without Controls for Log of Population, Republican Vote Share, Standard Deviation of Republican Vote Share, or Log of Total Hits

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years), U.S. Census of Population (various years), ICPSR 1, ICPSR 13, Hirano and Snyder (2019)

G.2 Using Wealth Instead of Population

As noted above, population is a crude proxy for potential advertising revenue, and county level income data are not available until 1950. Here we use wealth. The measure of wealth that is available for the greatest number of years – 1880, 1890, 1900, 1904 and 1912 – is the total assessed value of real estate. The data are from the U.S. Census, as well as the *Report on Wealth, Debt and Taxation* (various years) produced by the Census Bureau. Measures that include personal wealth are also available in 1880 and 1912. The correlation between total wealth and real estate wealth is above 0.99 in both years. We deflate wealth by the consumer price index values from the Federal Reserve Bank of Minneapolis.

Figure G16 is analogous to Figure 6 above but restricted to the period 1880 to 1912.⁵⁷ The panel on the left shows the predicted *RCS* values, with and without log of population as a control, and the panel on the right shows them with and without log of wealth as a control. Comparing the pairs of dashed lines in the two panels, it appears that the predictions using wealth as a control are no better than those using population.

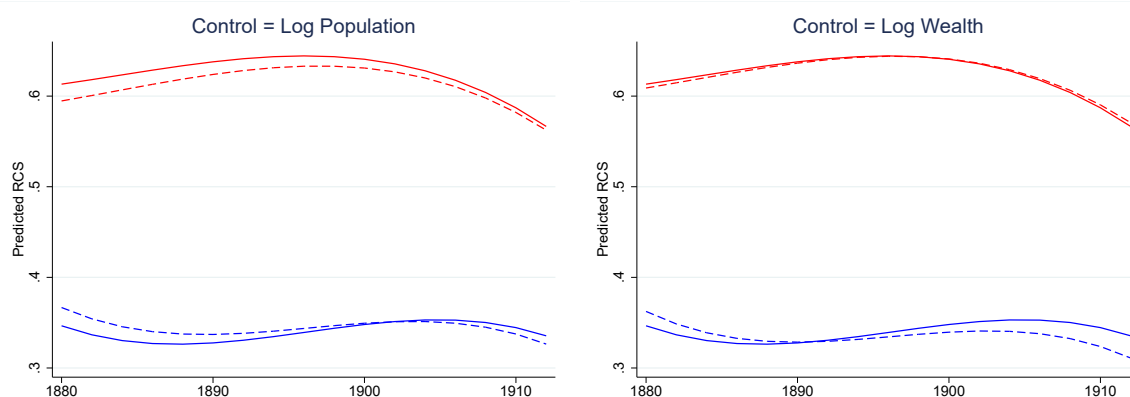


Figure G16: Predicted *RCS* of Combined Index, With and Without Controls for the Log of Population or Log of Wealth

Sources: Newspapers.com, ICPSR 30261, *N. W. Ayer & Son's American Newspaper Annual* (various years), U.S. Census of Population (various years), *Report on Wealth, Debt and Taxation* (various years), ICPSR 1, ICPSR 13, Hirano and Snyder (2019)

G.3 Radio and Television

Here we show the analyses regarding the introduction and expansion of radio and television. There are four separate regressions, two for radio and two for television. For each, one regression covers Republican newspapers and one covers Democratic newspapers. In all cases, the dependent variable is the change in *RCS* at the newspaper level. The two independent variables are the change in the share of households with radios (columns 1 and 2) or televisions (columns 3 and 4), and the change in the log of population as an additional control. In each column of Table G5, the main coefficient of interest – i.e., the change in the share of households with radios or TVs – is small and not statistically significant.

Table G5: **Estimated Effect of Radio and Television on RCS**

Item	Radio		Television	
	Republican Newspapers	Democratic Newspapers	Republican Newspapers	Democratic Newspapers
Radio Households	-0.000 (0.001)	-0.001 (0.001)	–	–
TV Households	–	–	0.009 (0.048)	-0.012 (0.067)
Log population	0.036 (0.070)	-0.038 (0.055)	0.037 (0.025)	-0.018 (-0.018)
Observations	156	128	144	119

Point estimates of regression of change in RCS on change in radio or tv penetration.
Standard errors in parenthesis.

Sources: Newspapers.com, ICPSR 30261, *N.W. Ayer & Son's American Newspaper Annual* (various years), U.S. Census (1930, 1940), ICPSR 22720 *Introduction of Television to the United States Media Market, 1946-1960*