Populist Persuasion in Electoral Campaigns: Evidence from Bryan's Unique Whistle-Stop Tour

Online Appendix

Johannes C. Buggle* Stephanos Vlachos†

*University of Vienna. e-mail: johannes.buggle@univie.ac.at.

†University of Vienna. e-mail: stephanos.vlachos@univie.ac.at.

Contents

A Data description **4**

A.1 Counties 4

A.2 Campaign speeches 5

A.3 Electoral outcomes 6

A.4 County characteristics 8

A.5 Crop value change 9

A.6 Historical newspapers 9

B Cross-sectional specification: Additional results **12**

B.1 Determinants of speeches 12

B.2 Details on estimating methods 14

B.3 Sensitivity to sample selection 21

B.4 General itinerary plotted 22

B.5 Effects on turnout and Republican vote 26

B.6 Congress election 28

C Difference-in-differences specification: Additional results **29**

C.1 Difference-in-differences evidence 29

C.2 Falsification in no-speech states 30

C.3 Turnout 33

C.4 Populist vote in Congress 33

C.5 Sensitivity analysis 36

C.6 Sensitivity to sample selection 39

D Where do the votes come from? Additional results **40**

D.1 Salience of debate on silver 40

D.2 Heterogeneous impacts of Bryan's speeches 45

D.3 Rival campaign information 47

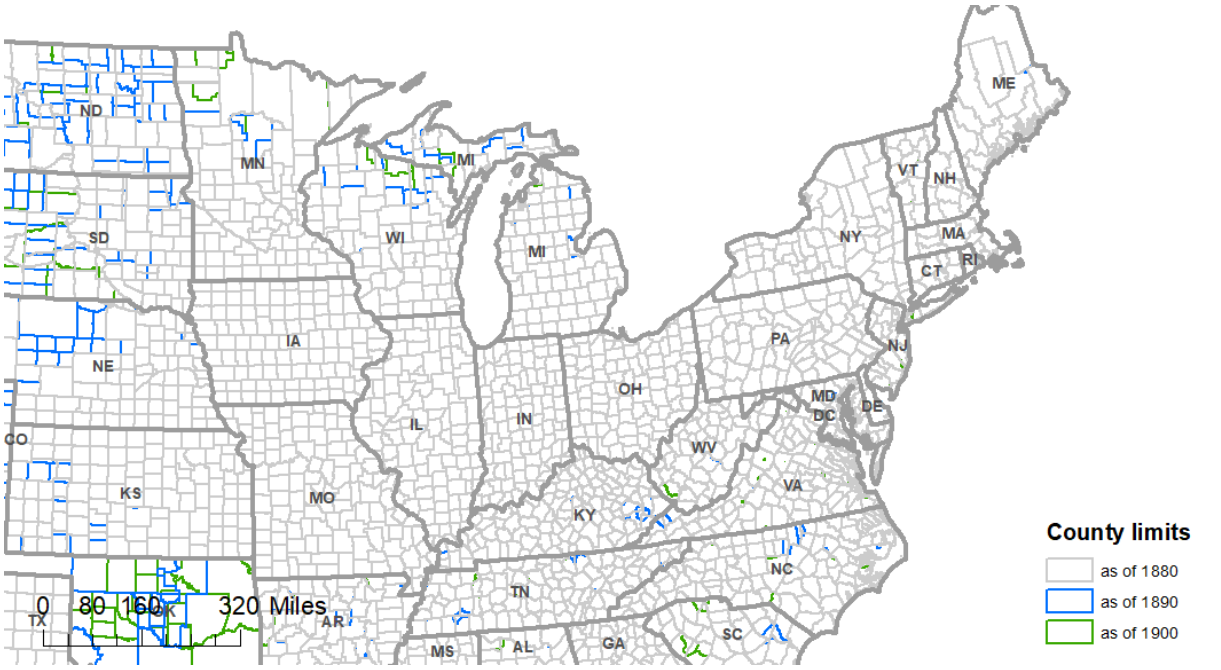
D.4 Quantification 48

A Data description

A.1 Counties

Source data. The unit of observation in this paper is a US County. Throughout the 1880 to 1900 period, which is the period of interest for this study, the geography of US counties changed. To harmonize the counties and keep a constant unit of analysis we exploit the historical county GIS Shapefiles that are available on the *National Historical Geographic Information System* (NHGIS). The 1880, 1890, and 1900 county shapefiles contains 2,614, 2,799, and 2,848 counties, respectively (including territories). The county limits are presented in Figure A.1. There are few changes in the period of interest in the states in which Bryan gave speeches. The largest changes take place in North Dakota and South Dakota (that became states in 1889), and Nebraska. Some changes can also be observed in northern Wisconsin and Michigan.

Figure A.1: Counties: County limits in 1880, 1890, and 1900



Note: Map of US counties and territories in 1880, 1890, and 1900 in the states in which William J. Bryan gave at least one speech. Source: *National Historical Geographic Information System* (NHGIS).

Harmonization. We harmonize the counties by keeping the largest common area. In doing so, we disregard minor changes. We define minor changes as those referring to less than 5% of

a county's surface. This avoids merging counties that only concede small parts of their surface to another county. This procedure leads to a set of 2,507 counties, 1,536 of which are located in speech states.

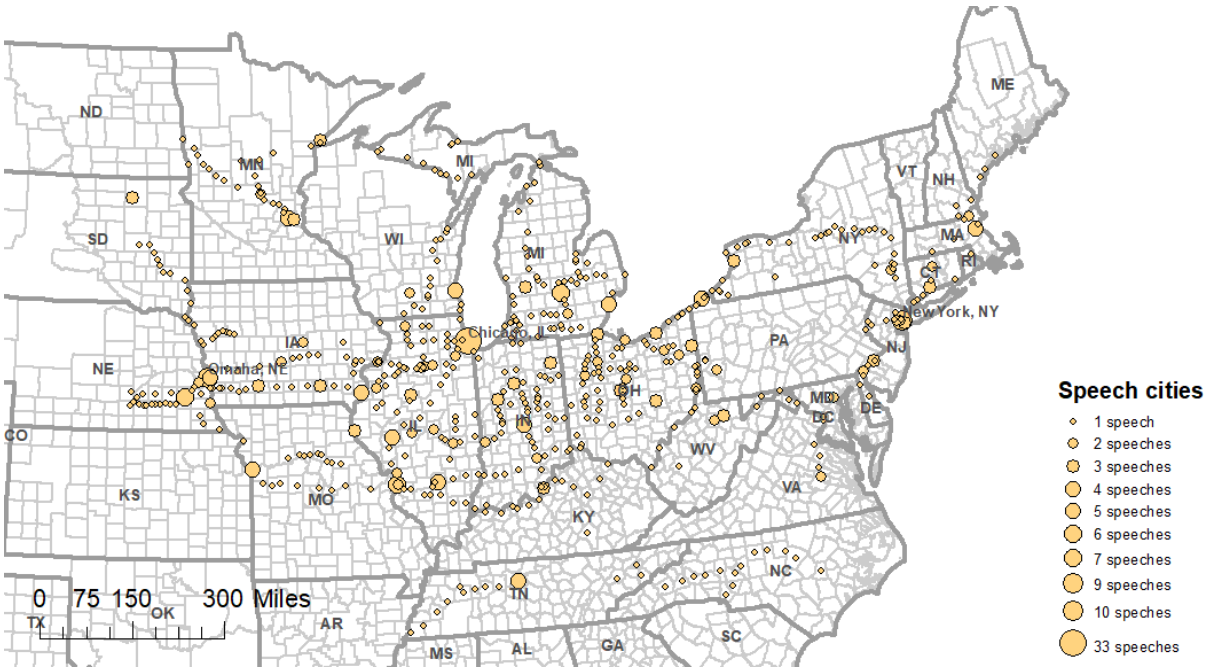
A.2 Campaign speeches

Source data. Information on Bryan's railroad trips and the location of his speeches comes from the University of Nebraska "Railroads and the Making of Modern America - A Digital History Project"', which can be accessed via http://railroads.unl.edu/topics/bryan.php (last accessed on Jan 19, 2022). The data contains for each speech Bryan gave, the city in which the speech was given, the date (day/year) and - for a limited number of observations - the time of the day. It also contains the detailed location where Bryan addressed the public, which ranges from town halls, to hotels, and railways depots. Information on the crowd size is only available for a small number of speech places. The data also contains information on the railroad that Bryan used to go from one place to another.

The speeches are separated into 4 distinct trips: (i) From Chicago, IL, after the Democratic National Convention (July 09, 1896) to Council Bluffs, IA (July 22, 1896), (ii) from Lincoln, NE (August 07, 1896) to New York, NY, for the acceptance of his nomination (August 12, 1896), and back to Lincoln, NE (September 08, 1896), (iii) the "Whistle-stop tour" from Lincoln, NE (September 11, 1896), and back to Council Bluffs, IA (October 31, 1896), and (iv) the election-eve tour in Nebraska, from Lincoln, NE, to Omaha, NE (November 02, 1896). The geography of the speeches is presented in Figure A.2.

Descriptives. Bryan gave 745 speeches, in 559 cities. The city in which he gave most speeches is Chicago, IL (33); he also gave 10 speeches in New York, NY, and 9 in Omaha, NE, while there are 481 cities in which he only gave a single speech. The speech cities are located in 382 counties, in 27 states (including DC). The state in which Bryan gave most speeches is Illinois, where he gave 135 speeches in 51 counties. Other states in which he heavily campaigned are Ohio (85 and 43), Michigan (78 and 39), Indiana (68 and 43), Iowa (55 and 28), New York (49 and 19), and Nebraska (43 and 17). In contrast, he only gave 1 speech in Rhode Island, 2 in North Dakota, and 3 in New Hampshire and Maine.

Figure A.2: Campaign speeches: Spatial distribution of Bryan speech cities



Note: Map of cities in which William J. Bryan gave at least one speech. Bigger circles correspond to more speeches. Harmonized US counties in the background.

A.3 Electoral outcomes

Source data. Electoral results come from *United States Historical Election Returns, 1824-1968*. The data is openly available on the Inter-university Consortium for Political and Social Research (ICPSR) platform (Study no. 1). The *United States Historical Election Returns, 1824-1968* dataset contains the state ICPSR code, the county ICPSR code, the name of the county, the number of the congressional district, and the vote tally of each party. We focus on outcomes for presidential, congressional, and midterm elections during the 1880-1900 period. The data is very rich and contains information on both major parties (the Republican and Democratic party) and minor parties for all elections. In total, 60 parties are present in the 1880 to 1900 elections. These party labels are state-specific; as such, in several cases, the same party has multiple labels, as for example, "Populist" and "People's", or "National Democrat" and "Gold Democrat".

The *United States Historical Election Returns, 1824-1968* does not contain information on turnout. We therefore retrieve this information from Clubb et al. (1987). The data is openly available on the Inter-university Consortium for Political and Social Research (ICPSR) plat-

form (Study no. 8611). The Clubb et al. (1987) dataset contains the state ICPSR code, the county ICPSR code, the name of the county, and information on turnout, the vote shares of major parties, and the legally eligible electorate for all elections during the 1840 to 1972 period. Turnout in the Clubb et al. (1987) data is given as a fraction over the legally eligible electorate. Gentzkow (2006) argues this measure is superior to the number over registered voters.

Data verification. The data is very rich and contains information on most parties and counties in the United States around the period. We nonetheless crosscheck the data using *Dave Leip's Atlas of the US Presidential Elections* (Leip, 1999) and Wikipedia. The data verification procedure follows several steps. We first harmonize the party labels. We then proceed to corrections that refer to data that is missing in the *United States Historical Election Returns, 1824-1968*, namely the 1892 Virginia presidential election, 9 counties in the 1892 Pennsylvania presidential election, the Prohibition party in the 1888 Tennessee presidential election, the Labor party in the 1888 Vermont presidential election, and Write-ins in the 1892 Georgia presidential election. We complement this information using Wikipedia.

At a second step, we proceed to state-level verification. To do so, we collapse our data at the state level, and cross the state-level data with Leip (1999). The differences are minor; the largest differences are found in the 1884 Pennsylvania presidential vote, which coincides with Wikipedia but not Leip (1999). We then verify the state-level party vote shares. The average difference between the two data sources in the states in which Bryan gave a speech is of 0.01 percentage points, with a standard deviation of 0.05 pp. The maximum difference is of 0.42 pp.

Our main outcome of interest is the vote share of Democrats over valid ballots in the 1896 presidential election. We construct similar variables for every Presidential election during the 1880 to 1900 period. In 1896, Bryan represented both the Democratic and the Populist party. The vote shares of these two parties are therefore aggregated for previous Congress and Presidential elections, but not for the 1900 election in which the two parties presented their own candidates. We also construct the vote share for all other parties in the relevant period. All data is harmonized to the county definitions described in Section A.1.

Congressional districts. The *United States Historical Election Returns, 1824-1968* contains information on the congressional district of every county. Counties that contain several districts (typically large urban centers) are given a generic district number in the original data. As a re-

sult of our county harmonization, some counties are associated to more than one district. This is typically the case when two counties that belong to different districts merge. In the states in which Bryan gave a speech, this is the case for 28 out of the 1,536 counties; 5 counties are associated to three districts: New York - Westchester, NY, Ashland - Chippewa - Florence - Forest - Iron - Langlade - Lincoln - Vilas - Marinette - Oconto - Oneida - Sawyer, WI, Beltrami - Kittson - Roseau, MN, Granville - Vance - Warren, NC, and Durham - Orange - Wake, NC. When this is the case, we attribute the district number of the county with the largest area.

A.4 County characteristics

Demographic variables. We calculate demographic characteristics of counties based on the decennial censuses available from the "National Historical Geographic Information System" (nhgis.org) and the "Integrated Public Use Microdata Series" (usa.ipums.org). In particular, we compute the (log) population per county, the share of the district population that lives in the county, the population share of whites, male above 21 years, individuals below 21, natives, the literacy rate for individuals aged 21 and above, and whether a county is urban. These censuses are available for 1880, 1890, and 1900; following Gentzkow (2006) we linearly interpolate the data for inter-census years.

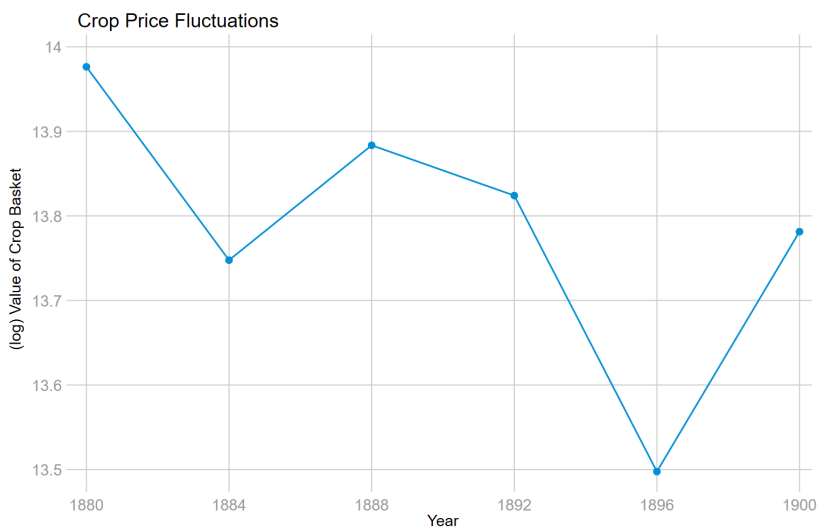
Geographic variables. We compute the (log) distances of each county to the state capital, and to the largest city in the state, as well as its geographic coordinates using ArcGIS. In addition, since Bryan used the railway network to travel from one city to another, our empirical analysis should rule out the possibility that differential access to railways confounds the results. While our sample already restricts to counties with railroad access, we additionally control for the density of the railroad network using data from Atack (2013), computed by dividing the length of all railroads by the county's area.

Economic variables. The economic variables come from several sources. The share of farmer households as a fraction of total households, and the share of manufacturing workers as a share of total population come from the "National Historical Geographic Information System". Newspaper circulation per eligible voter comes from Gentzkow et al. (2011), and the location of silver and gold mines is taken from Couttenier et al. (2017).

A.5 Crop value change

County-level crop value. We construct a county specific value of crop baskets, following the work of Eichengreen et al. (2019). We use information about the crop production by county in 1890 from the *US Agricultural Census* (Haines et al., 2018). The basket of crops contains corn, barley, oats, wheat, Irish potatoes, sweet potatoes, hay, rye, buckwheat, cotton, and tobacco. We apply to this basket yearly crop prices measured at the national level, taken from the Historical Statistics of the United States Millennial Edition Online (Tables Da667 to Da773). This gives a yearly measure of the value of all the crops produced in a given county. We construct two main variables that we control for in the different specifications: in the cross-section, we include the change in the (log) value of the crop basket between the election year and the previous election year, while we add to the difference-in-differences specification the (log) of the basket value in the election year. Figure A.3 displays the average (log) value of the crop baskets by election year in the period 1880 to 1900.

Figure A.3: Crop value by election year

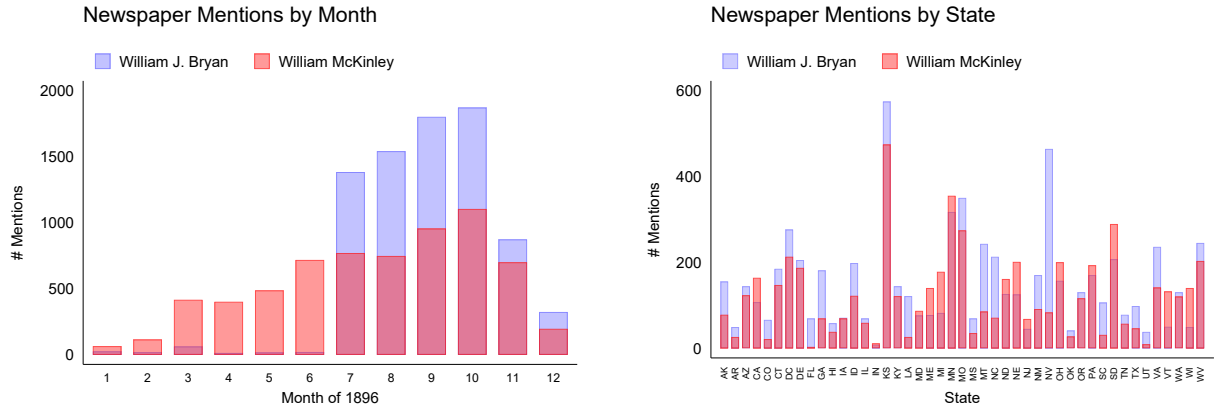


A.6 Historical newspapers

Data description. To study the coverage of the 1896 campaigns in the media, we have collected data on newspaper coverage from Chronicling America, a collection of digitized news-

papers provided by the *The Library of Congress*. We searched the database for mentions of "William J. Bryan", as well as "William McKinley" and found a total of 7,940 mentions for Bryan and 6,666 for McKinley in the year 1896. Figure A.4 displays the number of mentions by state and by month. Looking at the mentions by month, we observe that Bryan did not figure in newspapers before the start of his campaign in July 1896, but was then covered numerously and is mentioned more often than McKinley during the period from July to November 1896. The mentions of McKinley indicate that he was already present in newspapers in the first half of 1896, and that he was mentioned more frequently during the months preceding the election.

Figure A.4: Historical newspapers: Data description

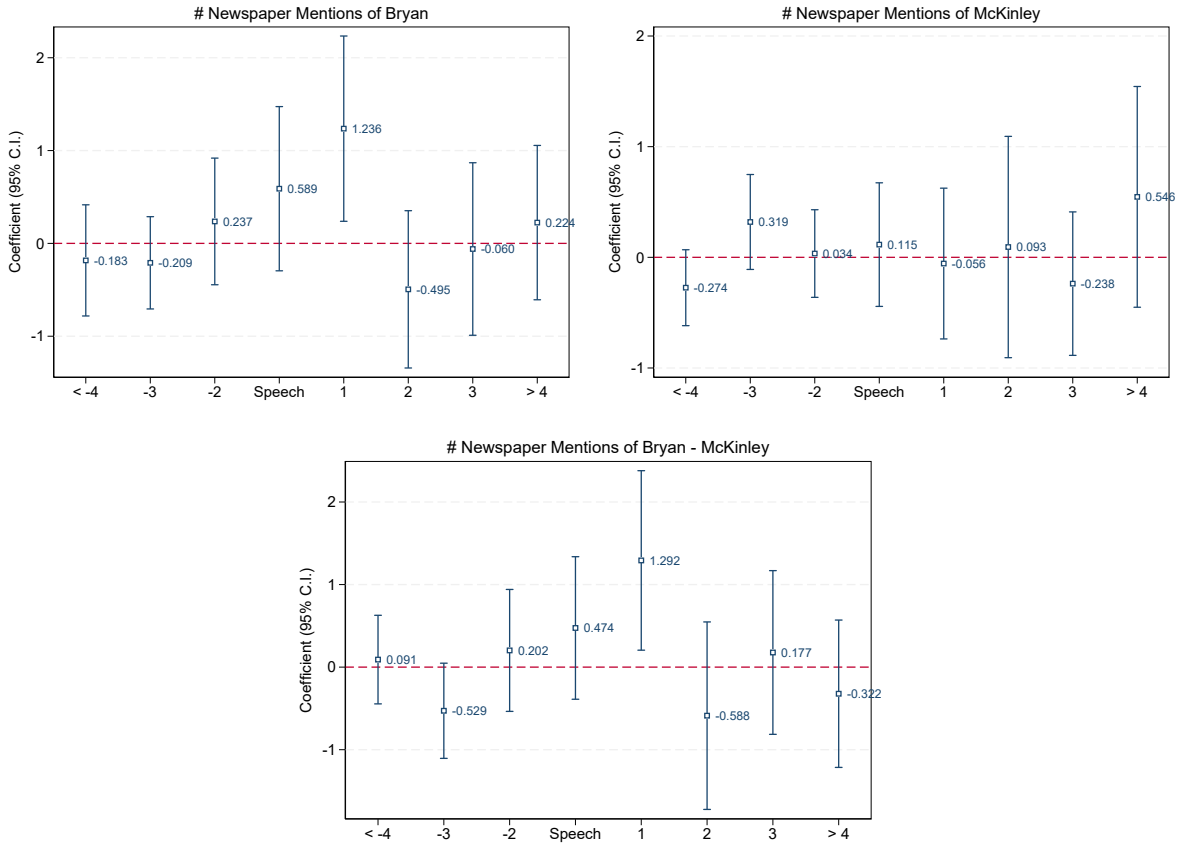


Note: This Figure shows the number of mentions of Bryan and McKinley by US State and month in the year of 1896 based on local newspaper data from Chronicling America, a collection of digitized newspapers provided by the The Library of Congress.

Newspaper coverage in the vicinity of speeches. We also examine whether the visits by Bryan had an effect on him being mentioned in a local newspaper. To this end, we link newspapers to counties and estimate an event-study at the county $\times$ week level. The results of this exercise are presented graphically in Figure A.5. We find suggestive evidence that Bryan was indeed covered more frequently in the week after he gave a speech in a county. This is evident by looking at the number of mentions of Bryan alone, and by the difference between the newspaper mentions of Bryan and McKinley. The effect is, however, short-lived, and not present from week 2 onwards. Two other points are important to note: first, there is no indication that Bryan was covered more frequently in the weeks before he gave a speech, which could indicate that he did not target places where his support base was already strong. Second, the newspaper mentions of McKinley do not react to the appearance of Bryan in a county. Overall, the

evidence presented suggests that Bryan's speeches could have been amplified by local newspaper coverage. This finding is also present in Shaw and Gimpel (2012) that show that Governor Perry's visits also increase his television and newspaper coverage. However, in the analysis of voting (Table D.4) we do not find that number of mentions of Bryan strengthens the effect of a speech in the months before the election.

Figure A.5: Historical newspapers: Newspaper coverage in the vicinity of speeches



Note: This Figure shows the coefficients and 95% confidence intervals of the effect of a speech by Bryan on the the number of newspaper mentions of Bryan and McKinley for the fours weeks before and after. The omitted category is the week before the speech. The first lead and last lag are binned, i.e., they take on the value 1 for all weeks before or after, respectively.

B Cross-sectional specification: Additional results

B.1 Determinants of speeches

In Table B.1 we document the determinants of speeches. We first estimate a state-level regression (Columns [1] and [2]). These regressions indicate that one of the most important predictor of speeches at the state level is the 1892 Democratic-Republican margin. This comes as no surprise, since the US presidential election is a state-level election, in the sense that the Electoral College is elected at the state-level. They also indicate that the choice of states to visit was also driven by the 1892 vote share of the Democratic and Populist party. A joint significance test confirms that past vote variables were important determinants when choosing the states to campaign (F-stat=11.75).

We then turn to lower levels of administration by estimating the same specification at the district and county level. The district-level regressions indicate that the main driver of Bryan speeches was the presence of manufacturing workers, see Columns (3) and (4); this findings is confirmed by the county-level regressions that also reveal that population and urban status were important drivers (Columns [4] to [7]). Importantly, these regressions indicate that, previous voting outcomes (vote shares, the vote margin, and turnout) were not factored in when choosing districts/counties within a state (F-stat=0.57, p-value=0.69). Moreover, county-level regressions indicate that accessibility (as captured by railroad density) was an important determinant of the visits.

We also examine the extend to which local demand for media, literacy, and demographics, predict the locality of speeches. Interestingly, while demographics are individually insignificant, they appear to be marginally jointly significant at the district level, as shown by the F-statistic at the bottom of Table B.1. Nonetheless, they do not seem to matter when focusing at the county level. To sum up, these findings imply that the logic of visits followed an intuitive, nested strategy: (i) Bryan first chose the states to visit based on past state-level voting patterns, and (ii) Bryan chose the counties within these states where there were potentially more voters to persuade, under the constraints of railroad accessibility. This is in line with a campaign strategy that focuses on locations in which there are still voters left to be convinced. Recall that the theoretical definition of persuasion in DellaVigna and Kaplan (2007), where persuasion f is defined as $f = 100 \times (y_T - y_C)/(e_T - e_C) \times /(1 - y_0)$, accounts for this by adjusting for size of

population left to be convinced through the $(1 - y_0)$ term.

Table B.1: Determinants of speeches

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Variable	Speech (Binary)							(log) Crowd	
Observation unit	State		District		County			County	
Controls	Baseline	Vote	Baseline	Vote	Baseline	Vote	Demo	Baseline	State FE
Democratic 1892 (%)	0.011	0.036**	–0.004	–0.003	0.000	0.001	0.001	–0.004	–0.008
	(0.014)	(0.016)	(0.005)	(0.005)	(0.001)	(0.001)	(0.001)	(0.012)	(0.012)
Populist 1892 (%)	0.009	0.030**	0.000	0.003	0.002	0.003	0.003*	–0.007	–0.007
	(0.013)	(0.011)	(0.015)	(0.019)	(0.002)	(0.002)	(0.002)	(0.017)	(0.021)
Margin 1892 (%)		–0.017***		–0.002		–0.000	–0.000		
		(0.004)		(0.004)		(0.001)	(0.001)		
Turnout 1892 (%)		–0.002		–0.000		0.000	0.000		
		(0.004)		(0.001)		(0.001)	(0.001)		
Pop (% of c/s/d)	–0.139	–0.075	0.008*	0.009*	0.001	0.001		–0.002	–0.000
	(0.087)	(0.087)	(0.005)	(0.005)	(0.001)	(0.001)		(0.006)	(0.006)
Urban counties	0.004	0.002	0.056	0.058	0.110***	0.110***	0.107***	0.407*	0.330
	(0.008)	(0.008)	(0.035)	(0.036)	(0.030)	(0.030)	(0.030)	(0.231)	(0.238)
Population (log)	–0.018	–0.031	–0.032	–0.040	0.103***	0.102***	0.124***	0.137	0.194
	(0.197)	(0.177)	(0.109)	(0.111)	(0.028)	(0.028)	(0.029)	(0.205)	(0.206)
Railroad density	0.029	0.020	0.009	0.009	0.006***	0.006***	0.006**	–0.015	–0.017
	(0.018)	(0.017)	(0.007)	(0.009)	(0.002)	(0.002)	(0.003)	(0.012)	(0.013)
State capital (log dist)	0.349	0.071	–0.056	–0.053	–0.038**	–0.038**	–0.036**	–0.086*	–0.065
	(0.438)	(0.447)	(0.037)	(0.036)	(0.015)	(0.015)	(0.015)	(0.051)	(0.056)
Largest city (log dist)	0.144	0.156	–0.006	–0.010	–0.037**	–0.037**	–0.043***	–0.004	0.012
	(0.328)	(0.345)	(0.033)	(0.034)	(0.016)	(0.016)	(0.016)	(0.046)	(0.062)
Latitude	0.120**	0.078**	0.001	0.002	0.031*	0.031*	0.032*	–0.006	–0.064
	(0.050)	(0.036)	(0.051)	(0.051)	(0.018)	(0.018)	(0.019)	(0.079)	(0.111)
Longitude	0.025*	0.031**	–0.056**	–0.055**	–0.015*	–0.015*	–0.016*	–0.023	–0.077
	(0.014)	(0.012)	(0.026)	(0.027)	(0.008)	(0.008)	(0.008)	(0.026)	(0.065)
Farmer households (%)	–0.006	–0.001	0.007	0.007	–0.001	–0.001	–0.001	–0.020*	–0.019
	(0.011)	(0.011)	(0.005)	(0.005)	(0.001)	(0.001)	(0.001)	(0.012)	(0.012)
Manuf workers (%)	0.021	–0.021	0.036***	0.035***	0.013***	0.013***	0.012**	0.036	0.028
	(0.056)	(0.044)	(0.008)	(0.008)	(0.005)	(0.005)	(0.005)	(0.027)	(0.026)
Crop value change (log)	–1.671	–0.728	–0.000	–0.001	–0.235	–0.238	–0.199	0.821	3.219
	(1.558)	(1.635)	(0.029)	(0.031)	(0.262)	(0.262)	(0.263)	(2.199)	(3.324)
Silver mine (log dist)	–0.202	–0.080	0.004	0.007	0.014	0.013	0.004	0.352*	0.605**
	(0.128)	(0.119)	(0.114)	(0.115)	(0.031)	(0.031)	(0.033)	(0.194)	(0.292)
Gold mine (log dist)	0.077	–0.005	–0.092	–0.089	–0.021	–0.020	–0.023	–0.066	0.065
	(0.111)	(0.137)	(0.125)	(0.128)	(0.022)	(0.022)	(0.022)	(0.174)	(0.246)
White population (%)	–0.008	0.006	0.006	0.007	0.005*	0.005*	–0.001	–0.004	–0.032
	(0.033)	(0.028)	(0.014)	(0.015)	(0.003)	(0.003)	(0.007)	(0.029)	(0.037)
Male above 21 (%)	0.056	0.046	–0.010	–0.011	–0.010	–0.010		–0.006	0.087
	(0.109)	(0.105)	(0.047)	(0.049)	(0.009)	(0.009)		(0.082)	(0.096)
Aged below 21 (%)	0.148	0.083	–0.047	–0.049	–0.014	–0.014	–0.003	0.028	0.123
	(0.103)	(0.084)	(0.042)	(0.044)	(0.011)	(0.011)	(0.008)	(0.086)	(0.103)
Native population (%)	0.008	0.001	–0.008	–0.008	0.003	0.003	0.002	0.013	0.011
	(0.011)	(0.007)	(0.010)	(0.010)	(0.003)	(0.003)	(0.003)	(0.016)	(0.017)

Notes: This table continues on the next page.

Table B.2: Determinants of speeches (cont'd)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Variable	Speech (Binary)							(log) Crowd	
Observation unit	State		District		County			County	
Controls	Baseline	Vote	Baseline	Vote	Baseline	Vote	Demo	Baseline	State FE
Literacy (%)	–0.029	–0.024	0.001	0.001	0.001	0.001	0.001	–0.034	0.019
	(0.022)	(0.022)	(0.014)	(0.014)	(0.003)	(0.003)	(0.003)	(0.037)	(0.044)
Male aged 21-44 (%)							0.000		
							(0.000)		
Male aged 45-64 (%)							–0.000		
							(0.000)		
Male age above 64 (%)							0.000		
							(0.000)		
Catholic population (%)							–0.001		
							(0.001)		
Black population (%)							–0.003		
							(0.006)		
State FE			Yes	Yes	Yes	Yes	Yes		Yes
Mean dep. variable	0.65	0.65	0.67	0.67	0.27	0.27	0.27	8.51	8.51
Observations	40	40	217	217	1,410	1,410	1,410	190	188
Clusters			23	23	228	228	228	115	113
Adjusted R^2	0.61	0.73	0.35	0.35	0.32	0.31	0.31	0.30	0.36
Past vote significance									
F-statistic	0.30	11.75	0.35	0.23	0.91	0.57	0.77	0.08	0.27
p-value	0.75	0.00	0.71	0.92	0.40	0.69	0.54	0.92	0.77
Demographics significance									
F-statistic	2.16	1.85	1.71	1.77	1.79	1.48	1.22	0.94	0.64
p-value	0.10	0.15	0.17	0.15	0.10	0.18	0.28	0.47	0.70

Notes: OLS regressions. The unit of observation is a county. Past vote significance is a joint test for the coefficients of the 1892 Democratic vote, the 1892 Populist vote, and when present, the 1892 joint Democrat-Populist and Republican margin and 1892 turnout. Demographics significance is a joint test for the coefficients of the white population, male population above 21, population aged below 21, native population, literacy, newspaper circulation, and when present, male population aged 21-44, male population aged 45-64, male population above 65, catholic population, and black population. Standard errors clustered at the Congressional district level in parentheses. * significant at 10%; ** at 5%; *** at 1%.

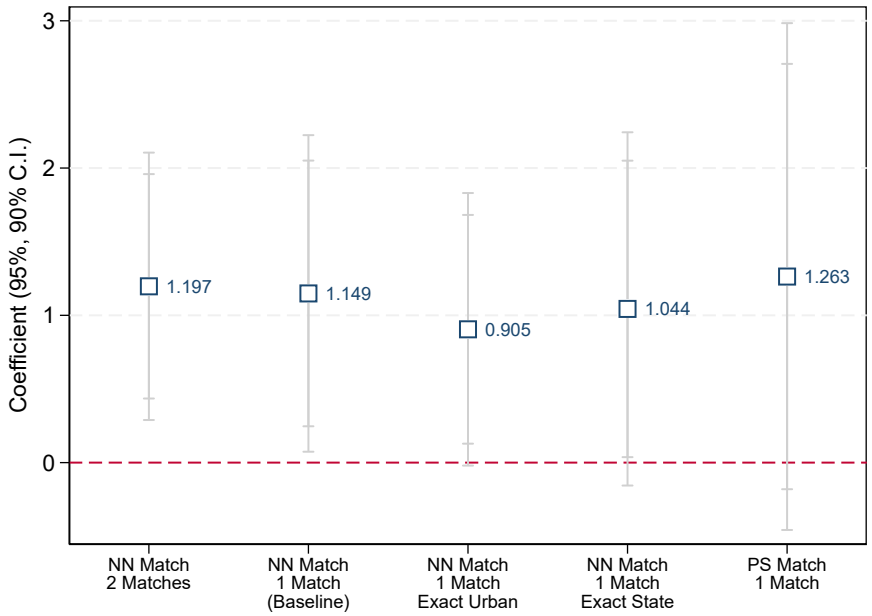
B.2 Details on estimating methods

Matching estimators. The nearest neighbor matching estimator estimates the average treatment effect of speeches by comparing counties where Bryan gave a speech to counties where his did not, but have similar observable characteristics. To match treatment and control observations, the nearest neighbor matching estimator weighs the difference between observable characteristics by the inverse of their variance-covariance matrix. In other words, the estimator, when matching counties, puts more weights to variables that vary little. The key advantage of the nearest neighbor matching estimator with respect to propensity score matching is that it is

non-parametric, in other words, it does not require any functional assumptions for the treatment model. As such, it does not depend on the quality of the prediction of the treatment, as is the case with propensity score matching.

We match counties based on the baseline covariates used in the cross-sectional estimation. Since most of our covariates are continuous variables, we use the Euclidean distance instead of the default Mahalanobis one, as suggested by Stata. Moreover, we request that the estimator includes a bias-correction for the fact that we match on continuous variable. In our baseline matching estimator we match each county to its nearest neighbor. Results for alternative specifications and matching algorithms are also presented.

Figure B.1: Details on estimating methods: Matching estimators



Notes: Sensitivity to the use of alternative matching estimator. From left to right: results when performing nearest neighbor matching on observable characteristics using the two closest matched counties; using the nearest match (baseline matching, displayed in the paper); using a single match and matching exactly on urban status; using a single match and matching exactly on the state; propensity score matching using a single match.

As shown in Figure B.3, the nearest neighbor matching estimator (displayed in the fourth panel) achieves a better balance of covariates than the within state comparison (displayed in the second panel). Some covariates remain unbalanced, namely distances to the largest city and state capital, crop value changes, and newspaper circulation. The fact that the estimator does not match these variables more precisely indicates that they exhibit a large variance.

The results from different matching estimators are presented in Figure B.1. The first four estimators present the nearest neighbor matching average treatment effects. The first one presents the nearest neighbor matching that uses the two closest neighbors; the second one presents the average treatment effect when only using the closest neighbor (our baseline matching estimator); the third presents the nearest neighbor matching when only using one neighbor and matching exactly on the urban/rural status of the county; the fourth when only using the closest neighbor and matching exactly the state in which the county is located. The coefficients in all nearest neighbor matching estimators are of comparable magnitude and statistically significant. The last estimator presents the propensity score matching. The coefficient is larger in magnitude though less precisely estimated.

Neighbor-pair fixed effects. The neighbor pair fixed effects estimator compares the vote shares of a county where Bryan gave a speech to the vote shares of all its adjacent neighbors that did not experience any speech. The key advantage of the estimator is that it compares close-by counties that share similar (geographic) characteristics, and it allows controlling for these common characteristics by including fixed effects for pairs of neighboring counties. To perform the neighbor-pair fixed effects estimator, we first identify the the subset of counties with a speech denoted by S, and all the adjacent municipalities that did not experience a speech denoted by $N(s)$. We then construct county-pairs (s, n) that each consist of one county s with a speech, and a neighbor n without speech. Each county can appear in several pairs, either as treated unit or as control. Counties that experienced a speech never act as control.

For each pair s and n that index counties with a speech and without, we estimate the following cross-sectional model:

$$\begin{aligned} dem_s &= \beta_0 + \beta_1 speech_s + \beta_2 dem_{s,t-1} + \beta_3 pop_{s,t-1} + Z_s'\gamma + \zeta_{s,n} + \varepsilon_s, \quad s \in S \\ dem_n &= \beta_0 + \beta_1 speech_n + \beta_2 dem_{n,t-1} + \beta_3 pop_{n,t-1} + Z_n'\gamma + \zeta_{s,n} + \varepsilon_n, \quad n \in N(s) \end{aligned} \tag{1}$$

where *dem* indicates the Democratic vote share in a county in 1896, *speech* is a measure of the speeches Bryan gave in county, Z is a vector of all the controls included in the cross-sectional model, and $\zeta_{s,n}$ are fixed effects for the pair (s, n).

As shown in Figure B.3, the neighbor pair fixed effects estimator (displayed in the fifth

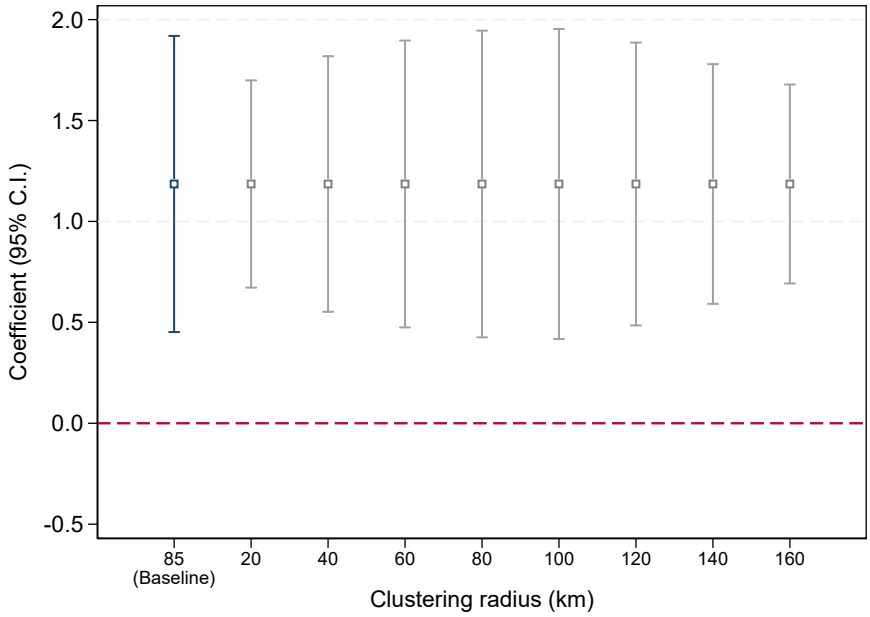
panel) achieves a better balance of covariates than the within state comparison (displayed in the second panel). The results presented in Table B.4 display a positive and significant effect of speeches on the Democratic vote share in 1896 across neighboring counties that is somewhat smaller than in the cross-sectional sample shown above. The smaller magnitude of the effect is expected given the likely possibility of spillover effects from the county that experienced a speech to its neighbors.

Entropy balancing. As an additional method to achieve pre-treatment balancedness between treated speech counties and un-treated control counties, we employ the entropy balancing procedure by Hainmueller (2012). Entropy balancing is a preprocessing procedure that allows calculating balancing weights so that covariates have equal means in the treatment and control group. The sample is therefore perfectly balanced. The treatment effect is then estimated on the balanced sample in a weighted regression where the weights are the balancing weights.

We implement entropy balancing and calculate balancing weights using the Stata package *ebalance*, see Hainmueller and Xu (2013). We consider different sets of pre- covariates on which we balance. In the case of the cross-section, we balance on the full set of covariates. By definition, differences in means among covariates between treated and untreated counties are close to zero as shown in Figure B.3. In the differences-in-difference specification, we balance on all pre-treatment characteristics without electoral variables, only pre-treatment electoral variables, and all pre-treatment covariates. The results using entropy weights are displayed for the cross-section in Table B.4, and for the differences-in-difference specification in Table C.3. Overall, the estimated treatment effects are positive and highly significant, albeit the coefficients are slightly smaller in magnitude than in the unweighted estimation.

Spatial clustering. This subsection estimates the baseline cross-sectional specification under different assumptions about the cutoff distances for spatial clustering. The results are displayed in Table B.3 and illustrated graphically in Figure B.2. We observe that - compared to our baseline distance of 85km - standard errors become smaller when we increase the distance up to 160km, and when we decrease the cutoff distance until 20km. Overall, the exercise implies that our choice of the cutoff distance of 85km represents a conservative approach: it is very close to the distance that maximizes standard error in the range between 20 to 160km.

Figure B.2: Details on estimating methods: Spatial clustering



Note: This Figure shows the coefficients and 95% confidence intervals of the effect of a speech given by Bryan on the Democratic vote share obtained from nince separate cross-sectional regressions, conditional on the full set of controls and state fixed effects. The corresponding regression results are provided in Table B.3.

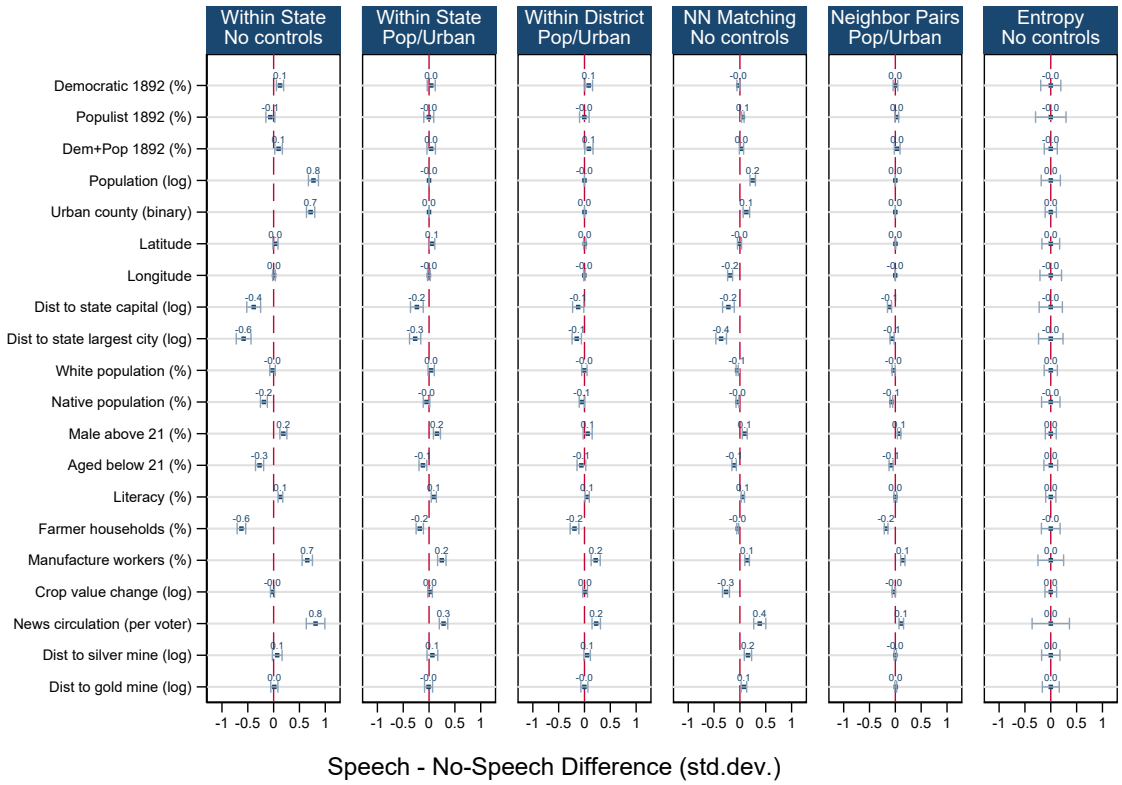
Table B.3: Details on estimating methods: Spatial clustering

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Distance Cutoff (in km's)	85	20	40	60	80	100	120	140	160
Speech (binary)	1.186***	1.186***	1.186***	1.186***	1.186***	1.186***	1.186***	1.186***	1.186***
	(0.374)	(0.262)	(0.323)	(0.363)	(0.388)	(0.392)	(0.357)	(0.303)	(0.252)
Dem (%) in 1892	0.825***	0.825***	0.825***	0.825***	0.825***	0.825***	0.825***	0.825***	0.825***
	(0.030)	(0.020)	(0.024)	(0.027)	(0.027)	(0.032)	(0.032)	(0.034)	(0.036)
Pop (%) in 1892	0.844***	0.844***	0.844***	0.844***	0.844***	0.844***	0.844***	0.844***	0.844***
	(0.038)	(0.029)	(0.031)	(0.035)	(0.037)	(0.042)	(0.038)	(0.036)	(0.039)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	46.20	46.20	46.20	46.20	46.20	46.20	46.20	46.20	46.20
Mean exp. variable	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27
Observations	1,410	1,410	1,410	1,410	1,410	1,410	1,410	1,410	1,410

Notes: OLS regressions. The unit of observation is a county. Demographic controls: Urban (binary), District population (%), Log population, White (%), Male above 21 (%), Native (%). Geographic controls: Railroad density (km per sq. km), Distance to state capital (log), Distance to state largest city (log), Latitude, Longitude. Economic controls: Farmers (%), Manufacture workers (%), Number of newspapers over eligible voters, Crop value change (log), (log) Distances to silver and gold mines. All regressions include state fixed-effects. Standard errors adjusted for spatial correlation in parentheses. * significant at 10%; ** at 5%; *** at 1%.

Balancedness. Figure B.3 displays the results of several balancedness tests for different estimation strategies. Panel 1 investigates differences controlling for state fixed effects only. As expected, the sample is particularly unbalanced along the population and urbanization dimension. Panel 2 takes therefore population and urbanization intro account by controlling for the two variables which improves balancedness substantially overall. The subsequent panels display the results for the within-district estimator, the nearest neighbor matching, the neighbor-pair fixed estimator, and for the entropy weighting procedure which forces variables to be balanced.

Figure B.3: Details on estimating methods: Balancedness



Note: Each coefficient is obtained from a separate regression that regresses a dummy taking on the value of 1 if Bryan has given a speech in the county on the covariate, controlling for state fixed effects, as well as population and urbanisation if indicated. Standard errors clustered at the Congressional district level.

Cross-sectional falsifications. Table B.4 present the results from the different aforementioned estimation methods, as well as the falsification exercises, also displayed in Figure 2 of the paper.

Table B.4: Details on estimating methods: Cross-sectional falsifications

	(1)	(2)	(3)	(4)	(5)
Dep. Variable	Democratic vote (%) in 1896				
Election year	1884	1888	1892	1896	1900
	(None)	(None)	(None)	(Bryan)	(D and R)
Panel A. Baseline (within state) estimation					
Speech (binary)	0.374	–0.018	–0.047	1.186***	0.214
	(0.374)	(0.324)	(0.295)	(0.292)	(0.314)
State FE	Yes	Yes	Yes	Yes	Yes
Observations	1,194	1,196	1,305	1,410	1,357
Clusters	223	226	232	228	229
Adjusted R^2	0.84	0.90	0.85	0.86	0.85
Panel B. Within neighbor-pair estimation					
Speech (binary)	–0.334	0.396**	0.116	0.616***	0.436**
	(0.279)	(0.188)	(0.170)	(0.172)	(0.179)
Neighbor-pair FE	Yes	Yes	Yes	Yes	Yes
Observations	2,168	2,172	2,306	2,450	2,364
Clusters	1,084	1,086	1,153	1,225	1,182
Adjusted R^2	0.80	0.90	0.90	0.89	0.90
Panel C. Within district estimation					
Speech (binary)	0.138	0.169	–0.438	0.755***	0.035
	(0.361)	(0.284)	(0.283)	(0.286)	(0.305)
Congressional district FE	Yes	Yes	Yes	Yes	Yes
Observations	1,172	1,172	1,279	1,384	1,330
Clusters	201	202	206	202	202
Adjusted R^2	0.85	0.91	0.88	0.88	0.88
Panel D. Entropy balancing estimation					
Speech (binary)	0.541	0.185	–0.316	1.041***	–0.354
	(0.360)	(0.296)	(0.306)	(0.315)	(0.422)
State FE	Yes	Yes	Yes	Yes	Yes
Observations	1,194	1,196	1,305	1,410	1,357
Clusters	223	226	232	228	229
Adjusted R^2	0.83	0.86	0.87	0.89	0.87
Panel E. Nearest neighbor matching estimation					
Speech (binary)	0.367	–0.490	–0.153	1.149**	0.125
	(0.487)	(0.335)	(0.450)	(0.548)	(0.554)
Observations	1,194	1,196	1,305	1,410	1,357

Notes: OLS regressions. The unit of observation is a county. All regressions include the full set of controls presented in Table 1 with the exception of the Matching estimator. Counties matched using the set of controls presented in Table 1 in Panel E. Panels A, C, and D: standard errors clustered at the congressional district level in parentheses; Panel B: at the neighbor-pair level. * significant at 10%; ** at 5%; *** at 1%.

B.3 Sensitivity to sample selection

In Table B.5 we estimate the effect of speeches in various samples. We first exploit information on Bryan's campaign that describes whether a speech took place in a train depot or a different location (convention hall, city hall, city park, hotel etc.). We interpreted speeches in locations such as city halls as targeted speeches (that we refer to as "planned" speeches), and the ones in train depots as stops on the road (speeches "not planned"). In the data there are 133 counties with planned speeches (35% of speech counties), and 248 counties where speeches took place in train depots (65%).

Table B.5: Sensitivity to sample selection

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable	Democratic vote (%) in 1896							
Sample	Excluding planned	Excluding not planned	Excluding >1 speeches	Excluding neighbors	Including no-railroad	Full sample	No-speech states	Excluding Pop strong
Speech (binary)	1.096***	1.529***	1.202***	1.969***	1.200***	2.028***		1.228***
	(0.332)	(0.474)	(0.326)	(0.574)	(0.295)	(0.506)		(0.282)
Dem (%) in 1892	0.832***	0.823***	0.831***	0.799***	0.837***	0.709***	0.614***	0.845***
	(0.028)	(0.031)	(0.028)	(0.040)	(0.029)	(0.039)	(0.079)	(0.028)
Pop (%) in 1892	0.832***	0.840***	0.834***	0.810***	0.842***	0.669***	0.333***	0.908***
	(0.033)	(0.036)	(0.034)	(0.047)	(0.037)	(0.071)	(0.108)	(0.043)
Population (log)	–0.340	–0.384	–0.496	–0.070	–0.356	–0.394	0.045	–0.363
	(0.461)	(0.512)	(0.460)	(0.553)	(0.439)	(0.585)	(1.361)	(0.458)
Urban county	–0.103	–0.062	–0.004	–0.203	–0.099	–0.480	–1.224	–0.056
	(0.394)	(0.432)	(0.406)	(0.526)	(0.380)	(0.662)	(1.154)	(0.375)
Farmers (%)	0.078***	0.071***	0.066***	0.068***	0.080***	0.066*	0.124*	0.064***
	(0.025)	(0.026)	(0.025)	(0.026)	(0.024)	(0.036)	(0.069)	(0.023)
Workers (%)	–0.068	–0.088	–0.090	–0.060	–0.038	–0.235*	0.018	–0.022
	(0.071)	(0.076)	(0.070)	(0.088)	(0.068)	(0.126)	(0.271)	(0.062)
Railroad (binary)					0.005	–2.332**		
					(0.770)	(1.095)		
Demo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes		Yes	Yes
Mean dep. variable	46.52	46.02	46.38	46.20	46.36	54.04	66.81	45.52
Mean exp. variable	0.19	0.11	0.17	0.45	0.25	0.16	0.00	0.29
Observations	1,276	1,162	1,241	856	1,510	2,348	735	1,194
Clusters	210	220	211	218	229	305	76	219
Adjusted R^2	0.86	0.85	0.86	0.84	0.86	0.72	0.70	0.87

Notes: OLS regressions. The unit of observation is a county. All regressions include the full set of controls presented in Table 1. All regressions include state fixed-effects. Standard errors clustered at the congressional district level in parentheses. * significant at 10%; ** at 5%; *** at 1%.

In Columns (1) and (2) of Table B.5 we present the results when excluding planned speeches and speeches not planned, respectively. The results indicate that there was an increase in Democrat vote in both types of locations when comparing these counties to other counties in the same state. The effect when excluding planned speeches is very similar to the baseline effect. This is reassuring, since for this group of counties selection should be less of an issue. The effect when excluding speeches not planned is larger, as one would expect, since for these counties selection might be more of an issue.

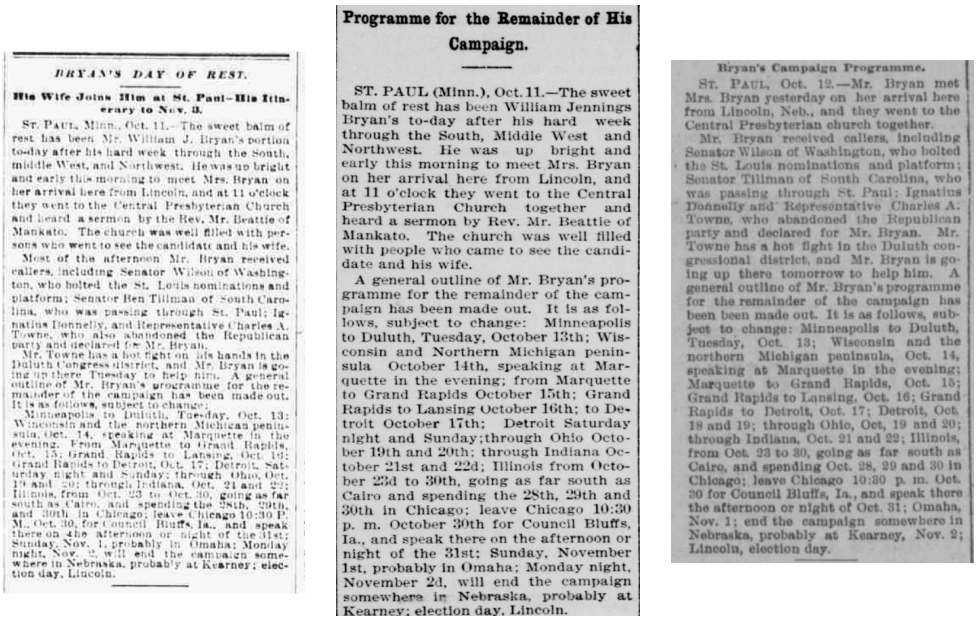
In Column (3) we only focus on counties visited at most once, since counties visited several times are more likely to be targeted. The results remain unaltered. In Column (4) we exclude the direct neighbors of counties with a speech, i.e., eliminating direct spillovers. This increases the the treatment effect. In Column (5) we expand the sample to counties that do not have railroad access in 1896 (and introduce consequently a railroad binary variable). In Column (6) we exploit information on the counties in which Bryan did not give a speech. The inclusion of these counties only serves in the estimation of the coefficients on the control variables, since there is no identifying variation of speeches in these states. Column (7) only focuses on no-speech states; the results of this estimation bring credence to the anecdotal evidence that indicated that Bryan was very popular among farmer populations. Finally, in Column (8) we exclude Populist strongholds from the estimation. We define a county as a populist stronghold, if the Populist/Democrat log ratio in 1892 exceeded the average Populist/Democrat log ratio by more than 1 standard deviation (roughly 12% of the sample).

B.4 General itinerary plotted

As an additional robustness, we attempted to find reports of the general itinerary that the campaign team of Bryan had scheduled. We were unable to find a source that explicitly states the initial plan. We thus searched for newspaper mentions of the campaign itinerary. For this purpose, we made use of the database Chronicling America, a collection of digitized newspapers provided by the *The Library of Congress*. We searched the database for mentions of "William J. Bryan", as well as ""Itinerary". We then manually verified the articles to determine whether they referred to the forthcoming Bryan speeches.

Data collection. We were able to identify several newspaper articles that refer to the itinerary at different points in time. The announcements vary in details and there is a fair deal of discrepancy between the speeches announced in newspapers and actual speeches. While some announcements detail almost all cities in which he plans to give a speech with date and time, other announcements are coarser and only inform about the city or state. Interestingly, the content of the newspaper articles does not appear to be location specific. For example, exactly the same route description is given in the the Sun (based in New York City, NY), the Record Union (based in Sacramento, CA), and the Evening Journal (based in Wilmington, DE), as shown in Figure B.4.

Figure B.4: General itinerary plotted: Data collection examples



BRYAN'S DAY OF REST.

His Wife Joins Him at St. Paul—His Itinerary to Nov. 3.

ST. PAUL, Minn., Oct. 11.—The sweet balm of rest has been Mr. William J. Bryan's portion to-day after his hard week through the South, middle West, and Northwest. He was up bright and early this morning to meet Mrs. Bryan on her arrival here from Lincoln, and at 11 o'clock they went to the Central Presbyterian Church and heard a sermon by the Rev. Mr. Beattie of Mankato. The church was well filled with persons who went to see the candidate and his wife.

Most of the afternoon Mr. Bryan received callers, including Senator Wilson of Washington, who bolted the St. Louis nominations and platform; Senator Ben Tillman of South Carolina, who was passing through St. Paul; Ignatius Donnelly, and Representative Charles A. Towne, who also abandoned the Republican party and declared for Mr. Bryan.

Mr. Towne has a hot fight on his hands in the Duluth Congress district, and Mr. Bryan is going up there Tuesday to help him. A general outline of Mr. Bryan's programme for the remainder of the campaign has been made out. It is as follows, subject to change:

Minneapolis to Duluth, Tuesday, Oct. 13; Wisconsin and the northern Michigan peninsula, Oct. 14, speaking at Marquette in the evening. From Marquette to Grand Rapids, Oct. 15; Grand Rapids to Lansing, Oct. 16; Grand Rapids to Detroit, Oct. 17; Detroit, Saturday night and Sunday; through Ohio, Oct. 19 and 20; through Indiana, Oct. 21 and 22; Illinois, from Oct. 23 to Oct. 30, going as far south as Cairo, and spending the 28th, 29th, and 30th in Chicago; leave Chicago 10:30 P. M., Oct. 30, for Council Bluffs, Ia., and speak there on the afternoon or night of the 31st; Sunday, Nov. 1, probably in Omaha; Monday night, Nov. 2, will end the campaign somewhere in Nebraska, probably at Kearney; election day, Lincoln.

Programme for the Remainder of His Campaign.

ST. PAUL (Minn.), Oct. 11.—The sweet balm of rest has been William Jennings Bryan's to-day after his hard week through the South, Middle West and Northwest. He was up bright and early this morning to meet Mrs. Bryan on her arrival here from Lincoln, and at 11 o'clock they went to the Central Presbyterian Church together and heard a sermon by Rev. Mr. Beattie of Mankato. The church was well filled with people who came to see the candidate and his wife.

A general outline of Mr. Bryan's programme for the remainder of the campaign has been made out. It is as follows, subject to change: Minneapolis to Duluth, Tuesday, October 13th; Wisconsin and Northern Michigan peninsula October 14th, speaking at Marquette in the evening; from Marquette to Grand Rapids October 15th; Grand Rapids to Lansing October 16th; to Detroit October 17th; Detroit Saturday night and Sunday; through Ohio October 19th and 20th; through Indiana October 21st and 22d; Illinois from October 23d to 30th, going as far south as Cairo and spending the 28th, 29th and 30th in Chicago; leave Chicago 10:30 p. m. October 30th for Council Bluffs, Ia., and speak there on the afternoon or night of the 31st; Sunday, November 1st, probably in Omaha; Monday night, November 2d, will end the campaign somewhere in Nebraska, probably at Kearney; election day, Lincoln.

Bryan's Campaign Programme.

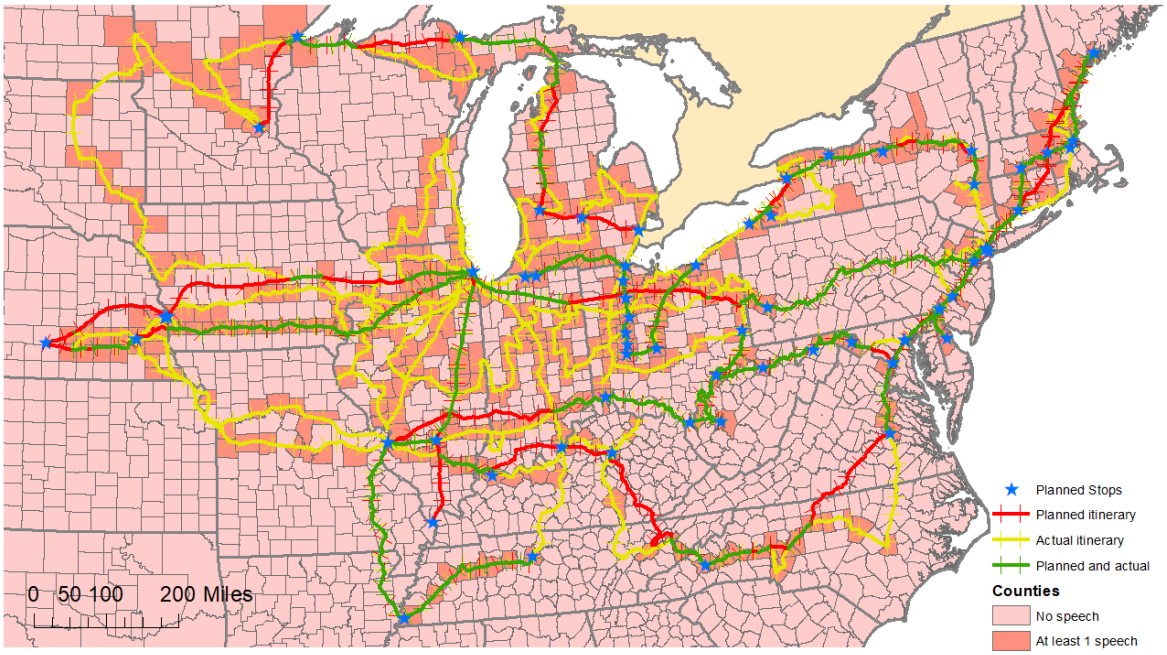
ST. PAUL, Oct. 12.—Mr. Bryan met Mrs. Bryan yesterday on her arrival here from Lincoln, Neb., and they went to the Central Presbyterian church together.

Mr. Bryan received callers, including Senator Wilson of Washington, who bolted the St. Louis nominations and platform; Senator Tillman of South Carolina, who was passing through St. Paul; Ignatius Donnelly and Representative Charles A. Towne, who abandoned the Republican party and declared for Mr. Bryan. Mr. Towne has a hot fight in the Duluth congressional district, and Mr. Bryan is going up there tomorrow to help him. A general outline of Mr. Bryan's programme for the remainder of the campaign has been been made out. It is as follows, subject to change: Minneapolis to Duluth, Tuesday, Oct. 13; Wisconsin and the northern Michigan peninsula, Oct. 14, speaking at Marquette in the evening; Marquette to Grand Rapids, Oct. 15; Grand Rapids to Lansing, Oct. 16; Grand Rapids to Detroit, Oct. 17; Detroit, Oct. 18 and 19; through Ohio, Oct, 19 and 20; through Indiana, Oct. 21 and 22; Illinois, from Oct. 23 to 30, going as far south as Cairo, and spending Oct. 28, 29 and 30 in Chicago; leave Chicago 10:30 p. m. Oct. 30 for Council Bluffs, Ia., and speak there the afternoon or night of Oct. 31; Omaha, Nov. 1; end the campaign somewhere in Nebraska, probably at Kearney, Nov. 2; Lincoln, election day.

Data description. The newspaper articles mention 78 stops that Bryan would perform, in several distinct parts, spanning from the Aug 7, 1896 to Aug 12, 1896 period, up to the Oct 13, 1896 to Nov 3, 1896. Large parts of the campaign remain undocumented. This is in some cases due to our incapacity to find articles that mention the campaign, as for example from Aug 12, 1896 to Aug 25, 1896, during the 2nd Bryan trip, and Sep 4, 1896, to Sep 12, 1896 (3rd trip in Wisconsin). In other cases, the description is just too vague, as for the Oct 18, 1896, to Oct 28, 1896, part of the campaign where newspapers mention that Bryan will visit Ohio and Indiana without any further details.

We use these 78 stops to calculate the shortest itinerary from one stop to another using the railroad network. We also perform the same exercise for the actual Bryan campaign visits. Both itineraries are presented in Figure B.5. Large parts of the planned itinerary overlap with the itinerary Bryan ended up doing. There are however also some differences between the itinerary that was mentioned in newspapers and the one that Bryan followed. Moreover, there are several parts of the campaign that are undocumented, either because we missed out on the articles, or because they were not reported in newspapers.

Figure B.5: General itinerary plotted: Data description



Note: Map of the itinerary as mentioned in newspapers and the actual itinerary followed by William J. Bryan. Information on the itinerary of Bryan as mentioned in newspapers comes from the database Chronicling America. Green lines refer to the itinerary that was both mentioned and followed; yellow lines refer to routes for which we did not find any newspaper mentions; the red itinerary refers to routes that were mentioned but not followed in the end.

Estimation results. We thus exploit the discrepancy between the plotted itinerary and the actual one in Table B.6. In Column (1) we focus on counties that were on the plotted itinerary, but that Bryan did not visit, and compare the voting outcome in these counties to the voting outcome in counties where Bryan actually gave a speech. The results from restricting our baseline cross-sectional specification to this sample confirm the findings of Table 2 in that, in counties

where Bryan actually gave a speech, the Democratic vote share increased by roughly 1.3pp. In Column (2) we exclude the counties whose localities are explicitly mentioned. These localities are usually large cities (see Figure B.4 for an example), and are likely endogenously chosen. Excluding counties with cities exactly mentioned does not alter the results. Finally, in Column (3) we restrict our sample ever further by excluding altogether places that were both mentioned in the plotted itinerary and where a speech took place. By doing so, we only compare counties that Bryan was initially planning to visit but did not, to counties that are not on mentioned routed and in which Bryan gave a speech. Interestingly, the point estimate in Column (3) is indistinguishable from the baseline coefficient of Table 2 Column (2).

Table B.6: General itinerary plotted: Estimation results

	(1)	(2)	(3)	(4)	(5)
Dep. Variable	Plotted	Excl. exact	Excl. plotted	Excl. speech	Excl. route
Democratic vote (%) in 1896	or speech	mentions	with speech	counties	performed
Speech (binary)	1.348***	1.340***	1.133**		
	(0.407)	(0.424)	(0.479)		
Plotted itinerary (binary)				–0.448	–0.538
				(0.418)	(0.750)
Lag Democratic (%)	0.814***	0.823***	0.793***	0.809***	0.820***
	(0.039)	(0.037)	(0.043)	(0.032)	(0.033)
Lag Populist (%)	0.892***	0.860***	0.855***	0.817***	0.825***
	(0.059)	(0.066)	(0.078)	(0.039)	(0.039)
Demographic controls	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	45.75	46.23	46.85	48.09	48.49
Mean exp. variable	0.68	0.65	0.50	0.17	0.08
Observations	545	484	348	1,023	826
Clusters	181	157	140	199	178
Adjusted R^2	0.88	0.89	0.88	0.87	0.88

Notes: OLS regressions. The unit of observation is a county. Demographic controls: Urban (binary), District population (%), Log population, White (%), Male above 21 (%), Native (%). Geographic controls: Railroad density (km per sq. km), Distance to state capital (log), Distance to state largest city (log), Latitude, Longitude. Economic controls: Farmers (%), Manufacture workers (%). All regressions include State fixed-effects. Standard errors clustered at the Congressional district level in parentheses. Standard errors adjusted for spatial correlation in square brackets. * significant at 10%; ** at 5%; *** at 1%.

Finally, in the last two columns of Table B.6 we perform placebo estimations. In Column (4) we compare counties on the plotted itinerary to counties where no speech was given. Counties on the plotted itinerary did not vote differently than counties where Bryan gave no speech. In Column (5) we further restrict the sample to counties that were not on the route that Bryan

actually followed. The election outcome in counties on the plotted (but not followed) itinerary was not different from neither the outcome in counties that were on the plotted, nor the actual itinerary.

B.5 Effects on turnout and Republican vote

Table B.7 presents the results from estimating Seemingly Unrelated Regressions (SUR) of the effect of speeches on Turnout and Democratic and Republican gross vote shares (as a fraction of total eligible voters). Panel A presents the results for the Democratic vote share; Panel B for the Republican one; Panel C for turnout. The results from Column (1) of Panel A indicate that the vote share of the Democratic party increased in counties where Bryan gave a speech; conversely, in these same counties, the vote share of the Republican party decreased, see Column (1) of Panel B; finally, the results when using turnout as the outcome indicate that, the speeches by William Jennings Bryan had no statistically significant effect on turnout, see Column (1) of Panel C.

There are two non-mutually exclusive ways to interpret these results, either as evidence of persuasion, or as evidence of mobilization Democratic-leaning voters and demobilization of Republican-leaning ones. The first interpretation assumes that there are no "defiers", in other words, the speeches of Bryan did not demobilize neither Republican-leaning, nor Democratic-leaning voters. Under this assumption, the results of Table B.7 can be interpreted as evidence of persuasion by Bryan's speeches; the pool of voters did not change significantly, but the voting choice of the voters switched from the Republican ballot to the Democratic one. Such an intepretation of the absence of effects on turnout is put forward in Pons (2018).

The second interpretation assumes that the speeches by Bryan had a demobilizing effect on Republican-leaning voters, and a mobilizing one of equivalent extent for Democratic-leaning ones. As such, these two effects cancel out, resulting in no observable effect on turnout but effects on the respective vote shares. Such an intepretation of the absence of effects on turnout is put forward in Spenkuch and Toniatti (2018). While Spenkuch and Toniatti (2018) attribute their effect to negative advertising, and we have been unable to find any account of negative mentions of McKinley in our reading of Bryan speeches, we are reluctant to take a stance, given how virulent the Republican campaign was.

Table B.7: Effects on turnout and Republican vote

Dep. Variable	(1) Speech (binary)	(2) Speeches/ 10k pop	(3) Speeches/ 10k pop	(4) Distance (log km)
Democratic vote (% eligible) in 1896				
Speech variable	0.883***	0.686*	1.955***	–0.272**
	(0.256)	(0.379)	(0.413)	(0.116)
Speech variable squared			–0.589***	
			(0.103)	
Democratic vote (% eligible) in 1892	0.784***	0.784***	0.785***	0.785***
	(0.024)	(0.024)	(0.024)	(0.023)
Populist vote (% eligible) in 1892	0.741***	0.739***	0.742***	0.740***
	(0.046)	(0.046)	(0.046)	(0.049)
Republican vote (% eligible) in 1892	–0.162***	–0.163***	–0.162***	–0.157***
	(0.023)	(0.023)	(0.023)	(0.023)
Republican vote (% eligible) in 1896				
Speech variable	–1.010***	–0.614*	–1.322**	0.236*
	(0.341)	(0.348)	(0.516)	(0.130)
Speech variable squared			0.329**	
			(0.138)	
Democratic vote (% eligible) in 1892	0.066	0.067	0.066	0.074
	(0.077)	(0.077)	(0.077)	(0.075)
Populist vote (% eligible) in 1892	0.054	0.055	0.054	0.056
	(0.068)	(0.068)	(0.068)	(0.067)
Republican vote (% eligible) in 1892	1.090***	1.090***	1.090***	1.101***
	(0.077)	(0.077)	(0.077)	(0.075)
Turnout (% eligible) in 1896				
Speech variable	–0.261	–0.002	0.557	0.004
	(0.416)	(0.394)	(0.666)	(0.160)
Speech variable squared			–0.260*	
			(0.155)	
Democratic vote (% eligible) in 1892	0.877***	0.877***	0.878***	0.887***
	(0.102)	(0.102)	(0.102)	(0.097)
Populist vote (% eligible) in 1892	0.812***	0.811***	0.812***	0.814***
	(0.096)	(0.096)	(0.096)	(0.096)
Republican vote (% eligible) in 1892	0.949***	0.949***	0.949***	0.966***
	(0.100)	(0.101)	(0.101)	(0.097)
Demographic controls	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Mean Democratic vote	39.41	39.41	39.41	39.72
Mean Republican vote	43.62	43.62	43.62	43.57
Mean turnout	84.60	84.60	84.60	84.86
Observations	1,410	1,410	1,410	1,260

Notes: Seemingly unrelated regressions (SUR). The unit of observation is a county. All regressions include the full set of controls presented in Table 1. All regressions include state fixed-effects. Standard errors clustered at the congressional district level in parentheses. * significant at 10%; ** at 5%; *** at 1%.

B.6 Congress election

Table B.8 presents the results from estimating the cross sectional model of Equation 1 with vote in favor of the Democrat or Populist candidates in the 1896 Congress election as the outcome. The difference in the observations between the baseline sample and the Congress election sample comes from the following states: 1896 Kentucky (missing 52 counties), 1894 Delaware (missing 3 counties), and 1892 Virginia (missing 81 observations).

Table B.8: Congress election

	(1)	(2)	(3)	(4)
Dep. Variable	Speech	Speeches/	Speeches/	Distance
Dem or/and Pop vote (%) in 1896 (House)	(binary)	10k pop	10k pop	(log km)
Speech variable	1.434**	1.300**	2.111**	–0.471**
	(0.621)	(0.552)	(0.879)	(0.234)
Speech variable squared			–0.370*	
			(0.210)	
Democratic vote (%) in 1894 (House)	0.325***	0.325***	0.326***	0.326***
	(0.049)	(0.048)	(0.048)	(0.051)
Populist vote (%) in 1894 (House)	0.028	0.028	0.028	0.038
	(0.087)	(0.087)	(0.087)	(0.087)
Democratic vote (%) in 1892 (House)	0.323***	0.323***	0.323***	0.327***
	(0.065)	(0.065)	(0.065)	(0.067)
Populist vote (%) in 1892 (House)	0.206*	0.206*	0.206*	0.203*
	(0.111)	(0.111)	(0.111)	(0.116)
Demographic controls	Yes	Yes	Yes	Yes
Geography controls	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Mean dep. variable	43.27	43.27	43.27	43.54
Observations	1,274	1,274	1,274	1,135
Clusters	216	216	216	215
Adjusted R^2	0.77	0.77	0.77	0.77

Notes: OLS regressions. The unit of observation is a county. All regressions include the full set of controls presented in Table 1. All regressions include State fixed-effects. Standard errors clustered at the Congressional district level in parentheses. * significant at 10%; ** at 5%; *** at 1%.

The results indicate that Bryan appearances also had an effect on the outcome of the Congress election. A counterfactual exercise (explained in detail in Section D.4) indicates that the campaign might have resulted in 7 additional representatives in the House.

C Difference-in-differences specification: Additional results

C.1 Difference-in-differences evidence

Table C.1: Difference-in-differences evidence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Variable	Democratic vote (%)				Rep (%)	Dem (%)	Turnout (%)
Speech variable	Speech (binary)	Speeches/ 10k pop	Speeches/ 10k pop	Distance (log km)	Speech (binary)	Speech (false)	Speech (binary)
Speech × 1884	–0.016	0.593	0.299	–0.171	–0.131	–0.007	0.794
	(0.414)	(0.566)	(1.043)	(0.265)	(0.407)	(1.412)	(0.688)
Speech × 1888	–0.250	0.047	0.018	–0.199	0.207	1.638	0.415
	(0.314)	(0.449)	(0.767)	(0.218)	(0.285)	(1.395)	(0.656)
Speech × 1892 (omitted)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Speech × 1896	0.974***	0.735*	2.148***	–0.471**	–0.861***	–0.208	0.474
	(0.318)	(0.424)	(0.499)	(0.203)	(0.295)	(1.110)	(0.491)
Speech × 1900	0.283	0.566*	0.743	–0.539**	–0.439	0.196	0.565
	(0.360)	(0.339)	(0.552)	(0.238)	(0.355)	(1.239)	(0.521)
Speech squared × 1884			0.290				
			(0.652)				
Speech squared × 1888			0.117				
			(0.473)				
Speech sq. × 1892 (omitted)			0.000				
Speech squared × 1896			–0.663***				
			(0.131)				
Speech squared × 1900			–0.090				
			(0.150)				
Full controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × Farmers (%)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × Workers (%)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	47.71	47.71	47.71	47.71	49.80	65.09	81.52
Observations	6,829	6,829	6,829	6,829	6,829	3,633	6,829
Clusters	1,410	1,410	1,410	1,410	1,410	735	1,410
Adjusted R^2	0.89	0.89	0.89	0.89	0.89	0.82	0.86
Pre-trends joint significance							
F-statistic ($\hat{\beta}_{1884} = \hat{\beta}_{1888} = 0$)	0.59	0.79	0.50	0.42	0.88	1.44	0.67
p-value ($\hat{\beta}_{1884} = \hat{\beta}_{1888} = 0$)	0.55	0.45	0.74	0.66	0.42	0.24	0.51

Notes: OLS regressions. The unit of observation is a county × election. Time-varying controls: Year × Farmers (%), Year × Manufacture workers (%). All regressions include the full set of controls presented in Table 1, except time invariant geographical characteristics. All regressions include Year × State fixed-effects and county fixed-effects. Standard errors clustered at the county level in parentheses. * significant at 10%; ** at 5%; *** at 1%. *Pre-trends joint significance* is a F-test that the 1884 and 1888 are jointly different from the 1892 coefficient.

Table C.1 presents the difference-in-difference estimation results. Column (1) presents our baseline results presented in Figure 3. Column (2) to (4) present the estimation of Equation (2)

using alternative measures of Bryan speeches. The results remain qualitatively unaltered.

Column (5) presents the point estimates when using the Republican vote share as the outcome (presented in Figure 4). Column (6) present the point estimates from the difference-in-differences falsification, presented in Figure 5, and described in detail in Section C.2. Finally, in Column (7) we present the results when using turnout as the outcome (discussed in Section C.3 and shown in Figure C.3).

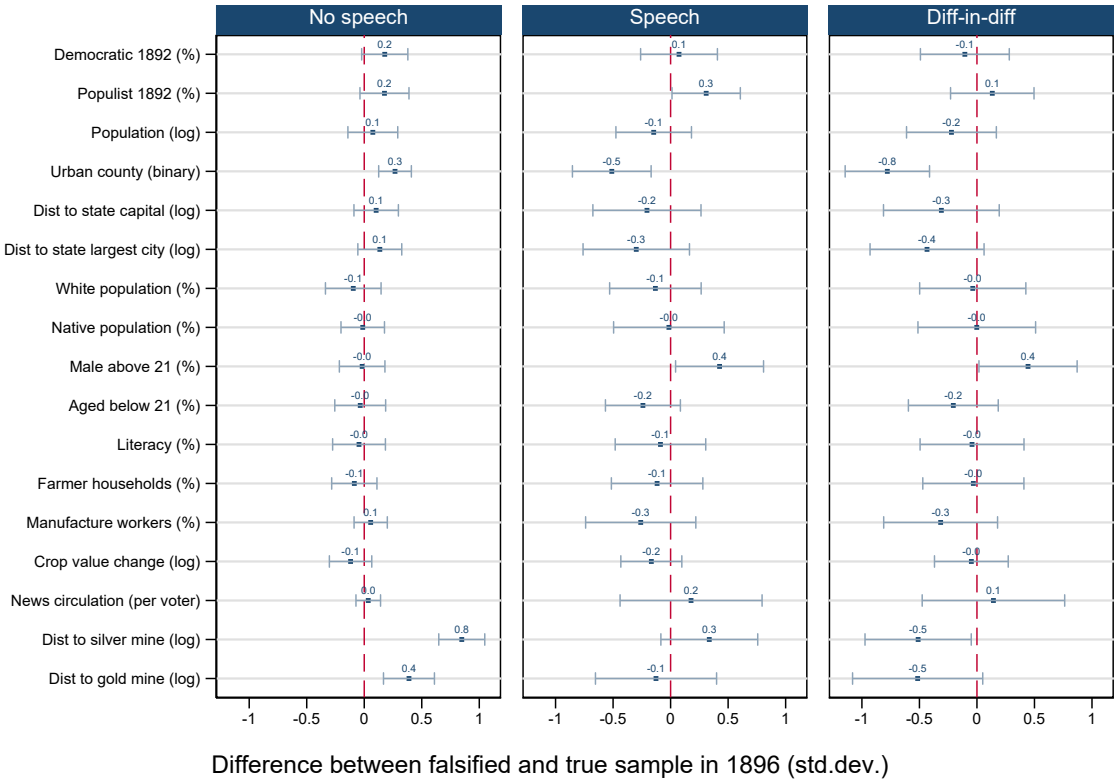
C.2 Falsification in no-speech states

To evaluate whether our difference-in-differences estimates are capturing unobservable trends in the vote of the Democratic candidate, we perform a falsification exercise in the 18 States in which he did not give any speeches. To perform this falsification and assign a potential "speech" and "no-speech" status to the counties in these States we proceed in two steps.

Matching procedure. We first estimate a nearest-neighbor matching regression to assign a "false treatment" status to each county. In this estimation we match every country located in no-speech states in 1896 to a county located in one of the 27 speech-states in the same year; we restrict the sample once again only to counties that have a railroad in 1896. The variables we use to perform the match are the control variables of our baseline cross-sectional estimation (with the exception of latitude and longitude); we once again use Euclidean distance to match on the continuous variables. This allows us to identify, for every county in the no-speech states, the county in the speech states that has the most similar observable characteristics. We then attribute to this county the speech/no-speech status of the most similar county in our true sample. There are 812 counties in the 18 States in which Bryan did not give any speeches; we have information on all covariates for 735 of those. Under this procedure, 89 of the 735 counties in no-speech states (12%) are associated to speech counties, and therefore the false treatment.

We then estimate the difference-in-differences estimation of Equation (2) using the false treatment. The results, presented in Figure 5 of the main paper and in Column (6) of Table C.1, are reassuring. None of the coefficients on both pre- and post-1892 elections is significant, meaning that counties that are more similar to the speech counties did not vote differently than counties that are more similar to non-speech counties, neither before, nor after 1896. A joint significant test for all four coefficients not being different than zero ($\hat{\beta}_{1884} = \hat{\beta}_{1888} = \hat{\beta}_{1896} =$

Figure C.1: Falsification in no-speech states: Procedure validation



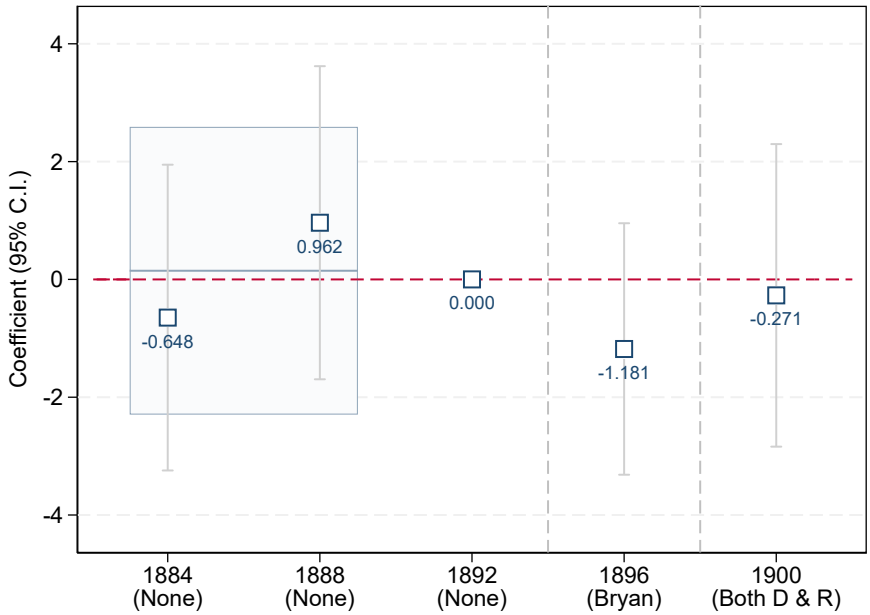
Note: Panel A ("No speech"): standardized differences between false control and actual control counties; Panel B ("Speech"): standardized differences between false treatment and actual treatment counties. Each coefficient is obtained from a separate regression that regresses a dummy taking on the value of 1 if the county belongs to a non-speech state. Panel C ("Diff-in-diff"): difference in the differences of Panels A and B. Standard errors clustered at the county level.

$\hat{\beta}_{1900} = 0$) yields a F-statistic of 0.74 and a corresponding p-value of 0.56.

Procedure validation. To evaluate the quality of the matching we compare observable county characteristics for false treatment/control counties to those of actual treatment/control counties. The results are presented in Figure C.1. The matching procedure performs well for both the "control" and "treatment" counties, with the exception of distances to silver and gold mines. We also look at the "difference-in-differences" between treatment-control and false-true treatment counties. The difference-in-differences reveal that the main characteristic on which we fail to properly match the treatment-control false counties is urbanization. More precisely, our false treatment counties are less urban (compared to our false control counties), that the actual treatment ones (compared to the actual control ones.

Sensitivity analysis. While results in Table B.5 indicate that, if anything, Bryan received a lower share of the vote in urban localities, meaning that, the difference in urbanization would increase the likelihood of the falsification yielding a positive result, we try to formally account for this in Figure C.2. This figure replicates the baseline result from Figure 5 of the paper, but weights observations by their similarity to their counterpart, where similarity to the counterpart is measured as the inverse distance between the no-speech county and its speech counterpart. Accounting for remaining differences does not change the results; a joint significance test for the 1884, 1888, 1896, and 1900 coefficients in the baseline regression yields an F-statistic of 0.74 and a corresponding p-value of 0.56, while it yields a an F-statistic of 0.89 and a corresponding p-value of 0.47 in the weighted regressions.

Figure C.2: Falsification in no-speech states: Sensitivity analysis

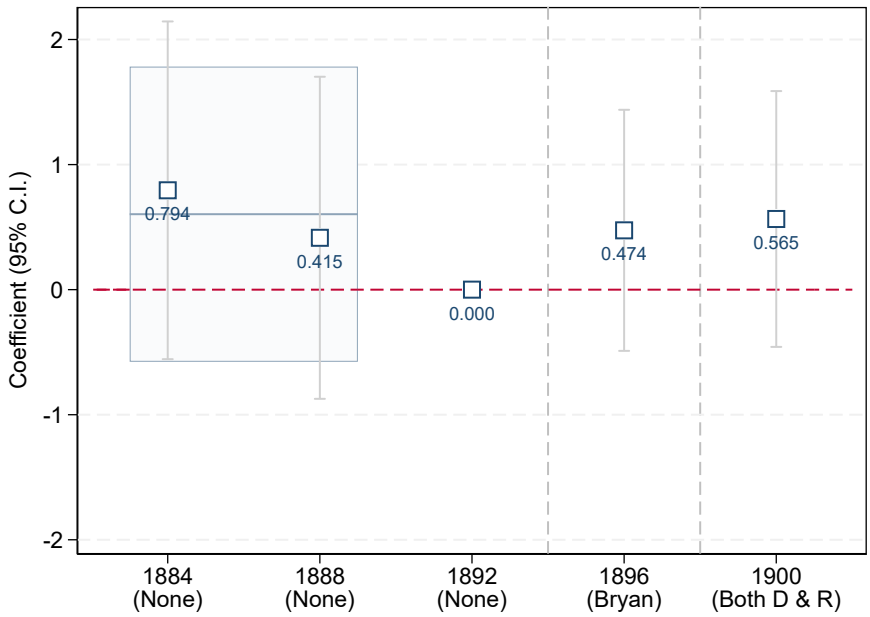


Note: This Figure replicates the results of Figure 5 when using the inverse distance between the matched false/true treatment/control counties as weights. Speech treatment is attributed by matching counties in no-speech states to their closest counterpart in speech states, and attributing them the treatment status of their counterpart. The Figure displays the coefficients and 95% confidence intervals of the interaction between a falsified speech and year dummies, conditional on county and state $\times$ election fixed effects, as well as controls. The treatment year is 1896. Information in parentheses below the year indicate which of the candidates campaigned in-person. A joint significance test for the 1884, 1888, 1896, and 1900 coefficients ($\hat{\beta}_{1884} = \hat{\beta}_{1888} = \hat{\beta}_{1896} = \hat{\beta}_{1900} = 0$) yields an F-statistic of 0.89 and a corresponding p-value of 0.47. The surrounding bounding box presents the 1884-1888 coefficient and 95% confidence interval of the difference-in-differences estimation using the interaction between a speech given by Bryan and 1884-1888, 1892, and 1896-1900 group binary variables.

C.3 Turnout

Figure C.3 presents the year-specific coefficients from estimating Equation (2) with turnout as the outcome. The results, also presented in in Column (7) of Table C.1, confirm the cross-sectional evidence in that there appears to be no significant effect of speeches on turnout.

Figure C.3: Turnout



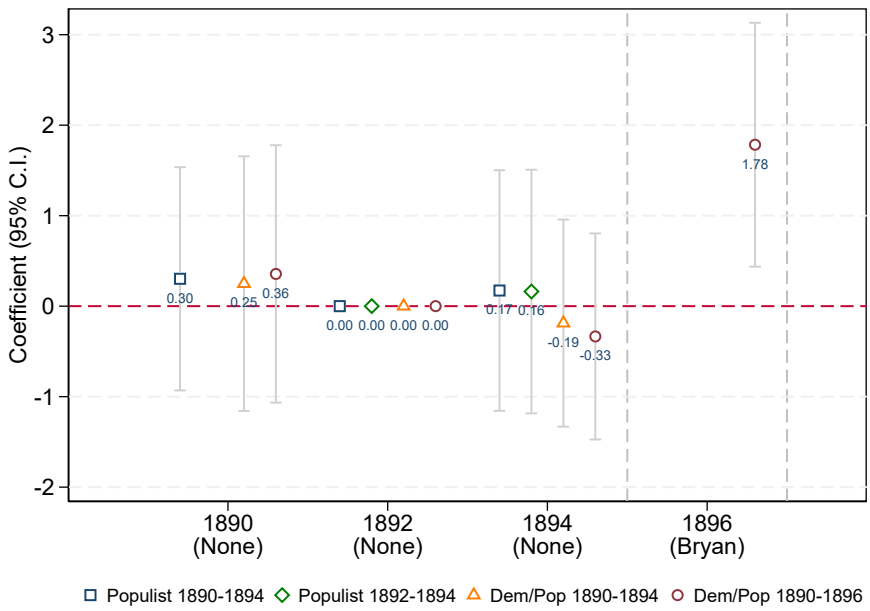
Note: This Figure shows the results of estimating the difference-in-differences specification of Equation (2) using voter turnout as the outcome. It displays the coefficients and 95% confidence intervals of the interaction between a speech given by Bryan and year dummies, conditional on county and state × election fixed effects, as well as controls. The treatment year is 1896. Information in parentheses below the year indicate which of the candidates campaigned in-person. A joint significance test for the 1884 and 1888 coefficients ($\hat{\beta}_{1884} = \hat{\beta}_{1888} = 0$) yields an F-statistic of 0.67 and a corresponding p-value of 0.51. The corresponding regression results are provided in Column (7) of Table C.1. The surrounding bounding box presents the 1884-1888 coefficient and 95% confidence interval of the difference-in-differences estimation using the interaction between a speech given by Bryan and 1884-1888, 1892, and 1896-1900 group binary variables.

C.4 Populist vote in Congress

Performing the difference-in-differences estimation for the Populist party in presidential elections in unfortunately unfeasible, since the party only presented a presidential candidate in the 1892 election. The Populist party was nonetheless present in three consecutive Congressional elections during the 1890-1894 period. We therefore exploit these electoral outcomes and estimate our difference-in-differences specification.

The results when using Populist Congressional elections vote as the outcome are presented in Figure C.4. In this figure, we present estimation results when using two different outcomes and three different periods. The first set of coefficients presents the results when estimating Equation (2) with populist vote as the outcome and the three elections during the 1890-1894 period. The results indicate that counties visited by Bryan in 1896 were not voting differently counties where Bryan did not go, neither in 1890, nor in 1894 (1892 is the omitted election).

Figure C.4: Populist vote in Congress



Note: This Figure shows the results of estimating the difference-in-differences specification of Equation (2) using (i) the Populist vote share in the 1890-1894 Congress election as the outcome; (ii) the Populist vote share in the 1892-1894 Congress election as the outcome; (iii) the joint Democratic and Populist vote share in the 1890-1894 Congress election as the outcome, and (iv) the joint Democratic and Populist vote share in the 1890-1896 Congress election as the outcome. It displays the coefficients and 95% confidence intervals of the interaction between a speech given by Bryan and year dummies, conditional on county and state × election fixed effects, as well as controls. The treatment year is 1896. Information in parentheses below the year indicate which of the candidates campaigned in-person. A joint significance test for the 1890 and 1894 coefficients ($\hat{\beta}_{1890} = \hat{\beta}_{1894} = 0$) yields F-statistics (p-values) of (i) 0.12 (0.88), (iii) 0.24 (0.79), and (iv) 0.45 (0.64). The corresponding regression results are provided in Table C.2.

Since in the 1890 Congress election the Populist was still marginal, and only present in 157 counties in our sample, we also estimate Equation (2) using only 1892 and 1894. This estimation has the advantage of overcoming the issue of how to treat counties where there was not Populist candidate (1,024 and 1,064 counties have such a candidate in 1892 and 1894, respectively). This comes however at the cost of reducing the number of elections for which one can test the common trends assumption. Still, the results from this estimation, presented in

diamonds in Figure C.4, confirm the previous findings, both quantitatively, and qualitatively. There appears to be no differential trend between speech and no-speech counties between 1892 and 1894.

Table C.2: Populist vote in Congress

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable	Populist			Democrat and/or Populist		
	Vote (%)		Candidate	Vote (%)		Candidate
Sample years	1890-1894	1892-1894	1890-1894	1890-1894	1890-1896	1890-1894
Speech (binary) × 1890	0.303		0.021	0.249	0.356	–0.013
	(0.629)		(0.032)	(0.717)	(0.725)	(0.009)
Speech (binary) × 1892 (omitted)	0.000	0.000	0.000	0.000	0.000	0.000
Speech (binary) × 1894	0.173	0.162	0.006	–0.188	–0.335	0.000
	(0.678)	(0.686)	(0.032)	(0.583)	(0.580)	(0.001)
Speech (binary) × 1896					1.784***	
					(0.687)	
Full controls	Yes	Yes	Yes	Yes	Yes	Yes
Year × State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year × Farmers (%)	Yes	Yes	Yes	Yes	Yes	Yes
Year × Workers (%)	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	6.44	7.62	0.54	48.60	46.97	0.96
Observations	4,159	2,820	4,159	4,158	5,568	4,159
Clusters	1,410	1,410	1,410	1,410	1,410	1,410
Adjusted R^2	0.66	0.68	0.64	0.84	0.80	0.91
Joint significance test						
F-statistic ($\hat{\beta}_{1890} = \hat{\beta}_{1894} = 0$)	0.12		0.17	0.43	0.27	1.08
p-value ($\hat{\beta}_{1890} = \hat{\beta}_{1894} = 0$)	0.89		0.84	0.65	0.76	0.34

Notes: OLS regressions. The unit of observation is a county × election. Time-varying controls: Year × Farmers (%), Year × Manufacture workers (%). All regressions include the full set of controls presented in Table 1, except time invariant geographical characteristics. All regressions include Year × State fixed-effects and county fixed-effects. Standard errors clustered at the county level in parentheses. * significant at 10%; ** at 5%; *** at 1%. *Joint significance test* is a F-test that the 1890 and 1894 are jointly different from the 1892 coefficient.

We then turn to the joint Democrat and Populist vote share in Congress elections. Democrats or Populists are present in the quasi-totality of the counties in our sample. The coefficients from estimating our difference-in-difference specification for the 1890-1894 (in triangles in Figure C.4) indicate once again that there were not differential trends between the two groups of counties. We then turn to the full 1890-1896 period, thus testing if there were spillovers from the Bryan campaign on the vote tally in Congress elections. The results (in circles in Figure C.4) confirm the evidence presented in Table B.8: Democratic and Populist candidates significantly increased their vote shares in counties where Bryan gave a speech.

Finally, to formally assess whether Bryan gave speeches in counties where Populist candi-

dates were more likely to run, we also estimate the difference-in-differences specification using a binary variable for the presence of a Populist candidate. The results are presented in Columns (3) and (6) of Table C.2.

C.5 Sensitivity analysis

The difference-in-differences specification crucially relies on the common trends assumption. The common trends assumption stipulates that, in the absence of treatment, the treated and control counties would have voted in the same way, and that the control counties are a good counterfactual for the treatment ones. The procedure we follow to test the parallel trends assumption closely follows Bilinski and Hatfield (2018) and consists of two distinct tests: (i) a test for different time trends between the treated and control group, and (ii) a placebo test that uses pre-treatment voting outcomes (which closely follows what we already do in the paper). Moreover, a recent debate has emerged on whether one should just look at pre-trends or whether one should also focus on differences in level before the treatment. The main argument of proponents of also focusing on levels is that there is no apparent reason for the mechanism behind the differences in levels not to simultaneously influence the trends.

Parallel trends test (slope). The first test of the parallel trends assumption we perform is a test of different slopes between the treatment and control groups. Note that the specification presented here is arguably inferior to the more flexible specification that we present as our baseline. This test is nonetheless useful to understand whether the pre-treatment coefficients we present in our baseline exhibit a trend that could be driving the post-treatment difference.

The test we perform to assess whether there are different trends between the control and treated counties closely follows what Bilinski and Hatfield (2018) refer to as a "one step up" approach. We first estimate a base model that imposes that all pre-treatment periods have the same coefficient. In our case, this boils down to estimating our basic difference-in-differences Equation (2) with the interaction between the Bryan speeches and dummies for the 1896 and 1900 elections. In other words, the basic specification in our context (Equation (1) in Bilinski and Hatfield, 2018) becomes

$$dem_{ct} = \lambda_{st} + \omega_c + \sum_{\tau=1896}^{1900} \beta_\tau \left(speech_c \cdot election_\tau \right) + Z'_{ct}\delta + \varepsilon_{ct} \tag{2}$$

This specification is then tested against a specification that includes a linear trend difference.[1] In our context, the specification become (Equation (2) in Bilinski and Hatfield, 2018)

$$dem_{ct} = \lambda_{st} + \omega_c + \sum_{\tau=1896}^{1900} \beta_\tau \left(speech_c \cdot election_\tau \right) + \theta \cdot speech_c \cdot t + Z'_{ct}\delta + \varepsilon_{ct} \tag{3}$$

The parallel trends test then requires that the coefficient on the linear trend is not statistically significant ($\hat{\theta} = 0$). In other words, there should be no different slope between the counties in which Bryan gave a speech and those where he did not.

The results from performing the aforementioned test are presented in Columns (1) and (2) of Table C.3. In Column (1) we estimate the model of Equation (2). The coefficient for 1896 is statistically significant and close in magnitude to the baseline coefficient (which is presented once again in Column [3]). In Column (2) we introduce a linear trend difference. The 1896 coefficient remains unchanged; this comes as no surprise, since as shown by the $Speech \times Year$ coefficient (p-value=0.97), both groups of counties have the same time trends.

Parallel trends test (placebo). The second parallel trends test proposed in Bilinski and Hatfield (2018) closely follows the approach we adopt in the paper. The test stipulates that one should test whether there is a significant treatment effect in periods prior the interventions, which in our context translates into counties where Bryan gave a speech already voting differently in elections before 1896.

As already mentioned, we already perform this test in our baseline difference-in-differences specification. Our baseline results are presented once again in Column (3) of Table C.3; the test of both the 1884 and 1888 coefficients being simultaneously not different from the 1892 coefficient yields an F-statistic of 0.59 (p-value=0.55). This means that the two coefficients as statistically indistinguishable from the 1892 coefficient, as one would expect if the parallel trends assumption holds. Note moreover that this failure reject the null is not the result of imprecise estimates, but rather because the estimates are small in magnitude.

In Column (4) we replicate the approach of Bilinski and Hatfield (2018) that suggest estimating the difference-in-difference specification only on pre-treatment periods (i.e. for elections before 1896). The 1884 and 1888 coefficients remain jointly insignificant (p-value=0.54). This

[1] We omit the $\sum_{\tau=1884}^{1900} \left(X'_{c\tau} \cdot election_\tau \right) \gamma_\tau$ part of our baseline difference-in-differences Equation (2) for ease of notation. Nonetheless, both the basic and the expanded model presented here include the term.

Table C.3: Sensitivity analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable	Democratic vote (%)							
Parallel trends test	Slope		Placebo			Balancing (entropy weights)		
	No slope	Linear slope	Baseline	Placebo	Extended pre-treat	1892 no elections	1892 elections	1892 full
Sample years	1884-1900	1884-1900	1884-1900	1884-1892	1880-1900	1884-1900	1884-1900	1884-1900
Speech × 1880					–0.580			
					(0.497)			
Speech × 1884			–0.016	–0.155	0.104	–0.494	–0.177	–0.278
			(0.414)	(0.423)	(0.417)	(0.455)	(0.407)	(0.453)
Speech × 1888			–0.250	–0.312	–0.217	–0.222	–0.433	0.032
			(0.314)	(0.309)	(0.319)	(0.397)	(0.296)	(0.429)
Speech × 1892 (omit.)			0.000	0.000	0.000	0.000	0.000	0.000
Speech × 1896	1.060***	1.043**	0.974***		0.973***	0.782*	0.710**	0.905**
	(0.344)	(0.406)	(0.318)		(0.319)	(0.432)	(0.311)	(0.438)
Speech × 1900	0.369	0.343	0.283		0.332	–1.260*	0.087	–0.964
	(0.371)	(0.594)	(0.360)		(0.363)	(0.728)	(0.355)	(0.644)
Speech × Year		0.002						
		(0.052)						
Full controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × Farmers (%)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × Workers (%)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	47.71	47.71	47.71	49.00	47.51	47.51	47.51	47.51
Observations	6,829	6,829	6,829	4,012	8,161	6,651	6,651	6,651
Clusters	1,410	1,410	1,410	1,338	1,410	1,374	1,374	1,374
Adjusted R^2	0.89	0.89	0.89	0.91	0.87	0.88	0.88	0.87
Pre-trends joint test								
Pre-trends F-stat		0.00	0.59	0.61	1.20	0.64	1.40	0.46
Pre-trends p-value		0.97	0.55	0.54	0.31	0.53	0.25	0.63

Notes: OLS regressions. The unit of observation is a county × election. Time-varying controls: Year × Farmers (%), Year × Manufacture workers (%). All regressions include the full set of controls presented in Table 1, except time invariant geographical characteristics. All regressions include Year × State fixed-effects and county fixed-effects. Standard errors clustered at the county level in parentheses. * significant at 10%; ** at 5%; *** at 1%. *Pre-trends joint significance* is a F-test that the 1884 and 1888 are jointly different from the 1892 coefficient.

might nonetheless be because we only have three pre-treatment periods. For example, Jaeger et al. (2020) show that, once expanding the reference period from 2005-2010 to 2001-2010 in the Kearney and Levine (2015) analysis, they fail to reject the null of no pre-treatment differences. To see whether this is the case in our estimation, we expand our baseline sample with the electoral outcome of 1880. Note that, by including electoral results that are that far in the past (16 years before the campaign), we increase the likelihood that there might be differences in voting outcomes. The results are presented in Column (5). The coefficients for the 1884 and

1888 elections are still small in magnitude and statistically insignificant; the coefficient on the 1880 election, while somewhat larger in magnitude, is very imprecisely estimated. A test of joint significance of the 1880, 1884, and 1888 coefficients still fails to reject the null that they are all three jointly not statistically significant (p-value=0.31).

Pre-treatment balancedness. As an additional method to achieve balancedness in pre-trends between treated speech counties and un-treated control counties we employ the entropy balancing procedure by Hainmueller (2012), as detailed in Section B.2.

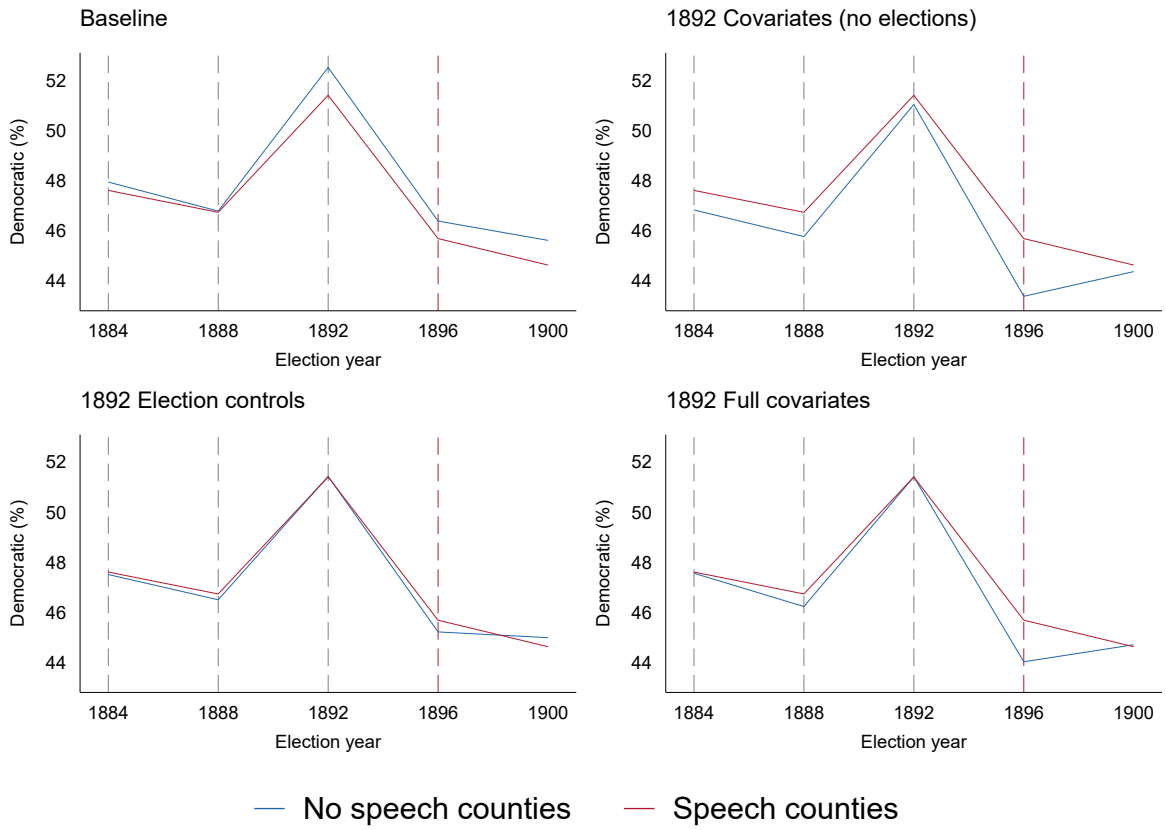
In Figure C.5 we plot the Democratic vote share for the years 1884 - 1900 separately for speech counties and non-speech counties in the baseline sample and in the samples that we pre-processed using entropy weights. The figure allows to visually inspect differences in levels in the outcome variable and associated differences in pre-trends. We observe that compared to the baseline, unweighted sample, the pre-processed samples achieve a much better balance in the pre-treatment Democratic vote share in 1892 (the level). In particular, when we balance on a set of co-variates that includes 1892 election controls, the Democratic vote share in the previous election appears to be perfectly balanced across treatment and control. The figure also shows that the averages of the Democratic vote share in the treatment and control group are co-moving more closely when the sample is pre-processed than in the case of the baseline, suggesting pre-trends that are more similar.

Columns (6) to (8) of Table C.3 display results using entropy weights for different sets of pre-treatment covariates. We balance on all 1892 pre-treatment characteristics without electoral variables (Column [6]), only on pre-treatment electoral variables (Column [7]), and on all pre-treatment covariates (Column [8]). We estimate treatment effects that are positive and highly significant, albeit the magnitude of the coefficients is slightly smaller than in the unweighted estimation.

C.6 Sensitivity to sample selection

In Table C.4 we investigate the effect of speeches in the differences-in-differences specification in various samples. We first exclude speeches that were planned (i.e., the more endogenous speeches), and find a smaller treatment effect of $\hat{\beta} = 0.80$ (Column [2]). If instead we exclude non-planned speeches (i.e., the less endogenous speeches), we find a larger treatment effect of

Figure C.5: Pre-treatment balancedness



Note: The figure displays the Democratic vote share for the years 1884 - 1900 separately for speech counties and non-speech counties in the baseline sample and in the samples that we pre-processed using entropy weights. In 1892 the figure displays the joint Democratic and Populist vote share.

$\hat{\beta} = 1.38$ (Column [3]). When excluding the direct neighbors of counties with a speech, i.e., eliminating direct spillovers, the treatment effect increases to $\hat{\beta} = 1.28$ (Column [4]). We also find an effect that is comparable to the baseline when including localities without a railroad, and looking at the full sample of states, even those that Bryan never visited (Columns [5] and [6]).

D Where do the votes come from? Additional results

D.1 Salience of debate on silver

Historical accounts suggest that the debate on silver was closely linked to crop price fluctuations (see e.g. Williams, 2010). The agrarian movement (Farmer's Alliance) on which the

Table C.4: Sensitivity to sample selection

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable	Democratic vote (%) in 1896					
Sample	Baseline sample	Excluding planned	Excluding not planned	Excluding neighbors	Including no-railroad	Full sample
Speech (binary) × 1884	–0.016	0.345	–0.934	–0.020	–0.029	0.662
	(0.414)	(0.415)	(0.823)	(0.652)	(0.418)	(0.436)
Speech (binary) × 1888	–0.250	–0.077	–0.702	–0.313	–0.318	0.550
	(0.314)	(0.341)	(0.524)	(0.503)	(0.321)	(0.356)
Speech (binary) × 1892 (omitted)	0.000	0.000	0.000	0.000	0.000	0.000
Speech (binary) × 1896	0.974***	0.801**	1.376**	1.280***	0.941***	1.249***
	(0.318)	(0.335)	(0.571)	(0.495)	(0.331)	(0.352)
Speech (binary) × 1900	0.283	0.370	–0.267	0.645	0.259	0.325
	(0.360)	(0.384)	(0.610)	(0.563)	(0.365)	(0.373)
Full controls	Yes	Yes	Yes	Yes	Yes	Yes
Year × State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year × Farmers (%)	Yes	Yes	Yes	Yes	Yes	Yes
Year × Workers (%)	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	47.71	47.89	47.68	47.68	47.83	54.39
Observations	6,829	6,168	5,616	4,133	7,319	11,452
Clusters	1,410	1,276	1,162	856	1,510	2,348
Adjusted R^2	0.89	0.90	0.89	0.88	0.90	0.88
Pre-trends joint significance						
F-statistic ($\hat{\beta}_{1884} = \hat{\beta}_{1888} = 0$)	0.59	1.14	0.96	0.44	0.88	1.37
p-value ($\hat{\beta}_{1884} = \hat{\beta}_{1888} = 0$)	0.55	0.32	0.38	0.65	0.41	0.25

Notes: OLS regressions. The unit of observation is a county × election. Time-varying controls: Year × Farmers (%), Year × Manufacture workers (%). All regressions include the full set of controls presented in Table 1, except time invariant geographical characteristics. All regressions include Year × State fixed-effects and county fixed-effects. Standard errors clustered at the county level in parentheses. * significant at 10%; ** at 5%; *** at 1%. *Pre-trends joint significance* is a F-test that the 1884 and 1888 are jointly different from the 1892 coefficient.

populist movement was built favored the coinage of silver to spur inflation, counter falling crop prices and to raise prices for farm products. The debate about free silver, which was heavily supported by farmers and miners, intensified after the crop failure of 1890 that preled the panic of 1893. The variation in crop price movements across time and across space seems therefore a crucial determinant of a county's support for free silver and the populist movement.

To account for the salience of silver we have constructed two types of controls. First, we construct a county specific value of crop baskets, following the work of Eichengreen et al. (2019). For this, we used information about the crop production by county in 1890 from the US Agricultural Census. The basket of crops contains corn, barley, oats, wheat, Irish potatoes, sweet potatoes, hay, rye, buckwheat, cotton, and tobacco. To this basket, we apply yearly crop prices measured at the national level. This gives a yearly measure of the value of all the crops pro-

duced in a given county (see Section A.5 for more details on the data). In all regressions, we now control for the change in the basket values: in the cross-section we include the change in the (log) average value of the crop basket between the current election period (defined as the election year and the three years preceding it) and the previous election period, while we add to the difference-in-differences specification the (log) of the basket value in the election period. Second, to take into account the support of miners for the free silver movement we control in all regressions for the (log) distance to the closest gold and silver mine.

We formally evaluate whether our control variables indeed capture the debate on silver and a latent support for Populist ideas in Table D.1. In Column (1) we present the results of the cross sectional estimation when using the 1892 Populist vote share as the outcome (all control variables are also from 1892). The results indicate that, as expected, in counties when the value of the agricultural basket decreased between the 1888 and 1892 elections, Populist support increased.[2] Column (2) serves as a falsification, by investigating whether Bryan visited counties with a stronger Populist support; this does not appear to be the case.

In Column (3) we investigate whether our control variables that are intended to capture the salience of the debate on silver are significant predictors of Democratic vote in 1896, when omitting to control for 1892 Populist vote. Indeed, these control variables are jointly statistically significant (see the p-value of the joint significance test). Once we control for past Populist vote however, these control variables are no longer statistically significant, as shown in Column (4). Note that Column (4) is our baseline cross-sectional regression.

In Columns (5) to (9) we turn to the vote share of other parties running in the 1896 election. Since Columns (4) to (9) present the vote shares of all parties running in 1896, this results also inform us on the origin of the votes in favor of the Democrats in counties where Bryan gave a speech. In the 1896 convention, the Prohibition party split. The "narrow gaugers", that only supported an alcohol prohibition won the presidential ticket, leading the "broader gaugers" that also wanted to include free silver and women suffrage in the party program to form a separate party, the National Prohibition party. Similarly, Democrats that were disillusioned by Bryan's candidacy, also known as Gold Democrats, went on to form the National Democratic party. The results indicate that the Democratic gains come from "pro-gold" parties. The vote share of Republicans and the "narrow-gauge" Prohibition party dropped significantly, while

[2]The different number of observations in the 1892 election stems from the fact that North Dakota and South Dakota were admitted in the states on November 2, 1889.

Table D.1: Salience of debate on silver

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Populist (1892)		Democratic (1896)		Pro-silver (1896)		Pro-gold (1896)		
Dep. Variable	Pop	Pop	Dem	Dem	Prohib	Labor	Rep	Dem	Prohib
Speech (binary)		0.342	1.580***	1.186***	0.016	–0.004	–1.048***	–0.044	–0.066*
		(0.358)	(0.434)	(0.292)	(0.017)	(0.014)	(0.291)	(0.028)	(0.037)
Lag Democratic (%)	0.004	0.004	0.694***	0.825***	–0.002***	0.001**	–0.826***	0.011***	–0.009***
	(0.021)	(0.021)	(0.038)	(0.028)	(0.000)	(0.001)	(0.028)	(0.002)	(0.002)
Lag Populist (%)				0.844***	0.000	0.002**	–0.843***	–0.001	–0.001
				(0.034)	(0.001)	(0.001)	(0.034)	(0.002)	(0.003)
Railroad (binary)	0.839	0.823							
	(0.987)	(0.983)							
Crop value change	–35.582***	–35.726***	–12.050	–1.576	–0.388**	0.137	2.179	–0.027	–0.198
	(10.332)	(10.351)	(8.859)	(4.398)	(0.168)	(0.190)	(4.727)	(0.292)	(0.834)
Dist to silver mine	–0.734	–0.733	–1.426**	–0.350	0.034	–0.019	0.440	–0.017	–0.106**
	(0.569)	(0.568)	(0.696)	(0.425)	(0.023)	(0.016)	(0.421)	(0.035)	(0.051)
Dist to gold mine	1.683***	1.689***	1.972***	0.904*	–0.016	0.003	–1.006**	–0.067**	0.179***
	(0.495)	(0.492)	(0.665)	(0.492)	(0.013)	(0.011)	(0.494)	(0.029)	(0.057)
Demo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	7.45	7.45	46.20	46.20	0.14	0.08	51.90	0.73	0.94
Mean exp. variable	0.28	0.28	0.27	0.27	0.27	0.27	0.27	0.27	0.27
Observations	1,341	1,341	1,410	1,410	1,410	1,410	1,410	1,410	1,410
Clusters	232	232	228	228	228	228	228	228	228
Adjusted R^2	0.77	0.77	0.72	0.86	0.71	0.42	0.85	0.67	0.48
Salience of silver									
F-statistic	6.88	6.96	5.12	1.28	1.64	3.14	2.01	3.52	0.74
p-value	0.00	0.00	0.00	0.28	0.18	0.03	0.11	0.02	0.53

Notes: OLS regressions. The unit of observation is a county. All regressions include the full set of controls presented in Table 1. All regressions include State fixed-effects. Standard errors clustered at the Congressional district level in parentheses. Standard errors adjusted for spatial correlation in square brackets. * significant at 10%; ** at 5%; *** at 1%.

the effect on the vote share of Gold Democrats marginally fails to reach statistical significance (p-value=0.12).

Accounting for unemployment. While the variables that are intended to capture the salience of the debate on silver are important predictors of vote shares, it can still be the case that the impact of the economic crisis is not completely captured by these variables. We therefore also attempt to capture the intensity of the crisis using unemployment rates. Unemployment information comes from the "Integrated Public Use Microdata Series" (usa.ipums.org) 1880 10% and 1900 5% samples. While the data is not perfect in that it does not cover all counties, and that there probably is a break in unemployment in 1893, we nonetheless exploit it to assess the

sensitivity of our results to directly accounting for unemployment.

Table D.2: Accounting for unemployment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Populist (1892)		Democratic (1896)		Pro-silver (1896)		Pro-gold (1896)		
Dep. Variable	Pop	Pop	Dem	Dem	Prohib	Labor	Rep	Dem	Prohib
Speech (binary)		0.340	1.569***	1.337***	0.018	–0.002	–1.195***	–0.050*	–0.068*
		(0.358)	(0.449)	(0.302)	(0.018)	(0.015)	(0.299)	(0.029)	(0.039)
Lag Democratic (%)	0.004	0.004	0.709***	0.830***	–0.002***	0.001**	–0.830***	0.010***	–0.009***
	(0.021)	(0.021)	(0.037)	(0.028)	(0.000)	(0.001)	(0.028)	(0.002)	(0.002)
Lag Populist (%)				0.847***	0.000	0.002***	–0.846***	–0.002	–0.001
				(0.037)	(0.001)	(0.001)	(0.038)	(0.002)	(0.004)
Railroad (binary)	0.824	0.809							
	(0.984)	(0.980)							
Unemployment rate	–0.005	–0.004	0.057	0.074***	0.002	–0.000	–0.065**	–0.008**	–0.003
	(0.038)	(0.038)	(0.035)	(0.026)	(0.002)	(0.001)	(0.027)	(0.003)	(0.005)
Crop value change	–35.312***	–35.460***	–4.750	–0.782	–0.438**	0.121	1.280	–0.065	0.057
	(10.368)	(10.388)	(9.422)	(5.159)	(0.181)	(0.213)	(5.554)	(0.331)	(0.967)
Dist to silver mine	–0.711	–0.710	–1.074*	–0.186	0.031	–0.013	0.284	–0.018	–0.114**
	(0.565)	(0.564)	(0.646)	(0.399)	(0.022)	(0.016)	(0.399)	(0.036)	(0.051)
Dist to gold mine	1.692***	1.697***	2.380***	1.233**	–0.019	0.007	–1.328***	–0.074**	0.177***
	(0.491)	(0.489)	(0.683)	(0.476)	(0.015)	(0.012)	(0.482)	(0.032)	(0.064)
Demo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	7.42	7.42	46.16	46.16	0.14	0.08	51.89	0.77	0.95
Mean exp. variable	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28
Observations	1,339	1,339	1,339	1,339	1,339	1,339	1,339	1,339	1,339
Clusters	232	232	226	226	226	226	226	226	226
Adjusted R^2	0.77	0.77	0.74	0.86	0.71	0.43	0.85	0.66	0.49
Salience of silver									
F-statistic	6.89	6.97	4.94	2.55	2.77	2.73	2.22	3.28	0.32
p-value	0.00	0.00	0.00	0.06	0.04	0.04	0.09	0.02	0.81

Notes: OLS regressions. The unit of observation is a county. All regressions include the full set of controls presented in Table 1. All regressions include State fixed-effects. Standard errors clustered at the Congressional district level in parentheses. Standard errors adjusted for spatial correlation in square brackets. * significant at 10%; ** at 5%; *** at 1%.

The results are presented in Table D.2 that replicates Table D.1 while adding the unemployment rate. Indeed, unemployment is an important predictor of 1896 Democratic vote even after accounting for fluctuation in crop prices. It is also a significant, negative predictor of Republican and Golden Democratic vote. Once unemployment is included, the effect of speeches on Democratic vote becomes larger ($\hat{\beta} = 1.337$ instead of 1.186), see Columns 4 of Tables D.2 and D.1. Overall, the effects for all parties qualitatively (and quantitatively) identical as without including unemployment. The fact that we are missing unemployment information for 71

counties, has motivated our decision not to include it in the baseline controls.

D.2 Heterogeneous impacts of Bryan's speeches

In Table D.3 we examine the hetereogeneous impacts of Bryan's speeches in the cross-sectional model. We first interact speeches with population shares of workers and farmers, urbanization, populist support in the previous election in 1892, as well as newspaper circulation. We find that speeches had a larger effect in counties with a larger population of workers. This effect cannot be explained by a differential impact of speeches in urban counties, or in counties that voted more strongly for the Populist party in 1892.

We do not find that Bryan's speeches had a differential impact in counties with more or fewer newspapers. However, it is still possible that Bryan speeches had a larger effect in industrial places because of their informative dimension. To test this idea, we estimate in Column (5) the triple interaction effect between speeches, industrial workers, and newspapers. Indeed, the results suggest that his speeches had a larger effect in industrial places when alternative sources of information (newspapers) were absent. The table also documents results from estimating the effect of each of the four trips that Bryan did separately, and from interacting speeches with the days passed since the start of the campaign. We consistently find that speeches closer to the election day mattered more than speeches early in the campaign.

Table D.3: Heterogeneous impacts of Bryan's speeches

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Variable			Democratic vote (%) in 1896				
Speech (binary)	–1.522	–1.740	–1.728	1.233***	–1.309		0.176
	(1.703)	(2.068)	(2.061)	(0.281)	(1.801)		(0.657)
Speech × Workers (%)	0.249***	0.249***	0.249***		0.356***		
	(0.094)	(0.094)	(0.094)		(0.114)		
Speech × Farmers (%)	0.036	0.038	0.035		0.029		
	(0.031)	(0.034)	(0.035)		(0.032)		
Speech × Urban county		0.153	0.190				
		(0.685)	(0.676)				
Speech × Populist 1892 (%)			0.015				
			(0.027)				
Speech × Newspapers circulation				–0.404	–0.217		
				(0.803)	(1.311)		
Speech × Workers (%) × News circ					–0.311*		
					(0.168)		
Speech × 1st trip						0.373	
						(0.538)	
Speech × 2nd trip						1.113**	
						(0.539)	
Speech × 3rd trip						1.107***	
						(0.319)	
Speech × 4th trip						1.664	
						(1.149)	
Speech × Days since campaign start							0.013*
							(0.008)
Democratic vote (%) in 1892	0.826***	0.826***	0.826***	0.825***	0.828***	0.825***	0.825***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.027)	(0.029)	(0.028)
Populist vote (%) in 1892	0.843***	0.843***	0.841***	0.843***	0.843***	0.844***	0.842***
	(0.034)	(0.034)	(0.033)	(0.034)	(0.033)	(0.034)	(0.034)
Workers (%)	–0.151*	–0.150*	–0.151*	–0.060	–0.276***	–0.058	–0.059
	(0.082)	(0.082)	(0.083)	(0.065)	(0.103)	(0.067)	(0.066)
Farmers (%)	0.063**	0.062**	0.063**	0.069***	0.063**	0.070***	0.071***
	(0.026)	(0.026)	(0.026)	(0.024)	(0.026)	(0.024)	(0.024)
Urban county	0.080	0.042	0.045	–0.131	0.164	–0.085	–0.102
	(0.375)	(0.448)	(0.448)	(0.372)	(0.378)	(0.376)	(0.375)
Workers (%) × News circulation					0.341**		
					(0.133)		
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	46.20	46.20	46.20	46.20	46.20	46.20	46.20
Observations	1,410	1,410	1,410	1,410	1,410	1,410	1,410
Clusters	228	228	228	228	228	228	228
Adjusted R^2	0.86	0.86	0.86	0.86	0.86	0.86	0.86

Notes: OLS regressions. The unit of observation is a county. Demographic controls: Urban (binary), District population (%), Log population, White (%), Male above 21 (%), Native (%). Geographic controls: Railroad density (km per sq. km), Distance to state capital (log), Distance to state largest city (log), Latitude, Longitude. Economic controls: Farmers (%), Manufacture workers (%). All regressions include State fixed-effects. Standard errors clustered at the Congressional district level in parentheses. * significant at 10%; ** at 5%; *** at 1%.

D.3 Rival campaign information

In this sub-section, we study whether the campaign of McKinley partly offset the campaign efforts by Bryan. We incorporate the intensity of local newspaper coverage of the two candidates in the months preceding the election in the regression of vote shares on speeches. For this, we compute the number of newspaper mentions per county for each of the two candidates from June 1, 1896, to November 3, 1896. See Section A.6 for more details on the data. Most counties do not have any mention of the two candidates, i.e., the measures are zero. The absence of mentions could either stem from the fact that local newspapers did not report on any of the two candidates, or that Chronicling America does not cover the newspapers.

Table D.4: Rival campaign information

	(1)	(2)	(3)	(4)	(5)
Dep. Variable	Democratic vote (%) in 1896				
Sample	Full sample	Full sample	Full sample	Full sample	At least 1 mention
Speech (binary)	1.242***	1.188***	1.241***	1.384***	2.844**
	(0.306)	(0.294)	(0.306)	(0.346)	(1.274)
McKinley newspaper mentions (per elig. voter)		–13.814			
		(95.611)			
Bryan newspaper mentions (per elig. voter)		43.756			
		(89.844)			
McKinley to Bryan news mentions (ratio)			–0.115	–0.046	0.111
			(0.107)	(0.166)	(0.216)
Speech × McKinley to Bryan news mentions				–0.140	–0.598**
				(0.194)	(0.299)
Demographic controls	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	46.20	46.21	46.20	46.20	46.28
Observations	1,410	1,399	1,410	1,410	164
Clusters	228	227	228	228	97
Adjusted R^2	0.85	0.86	0.85	0.85	0.81

Notes: OLS regressions. The unit of observation is a county. Demographic controls: Urban (binary), District population (%), Log population, White (%), Male above 21 (%), Native (%). Geographic controls: Railroad density (km per sq. km), Distance to state capital (log), Distance to state largest city (log), Latitude, Longitude. Economic controls: Farmers (%), Manufacture workers (%). All regressions include State fixed-effects. Standard errors clustered at the Congressional district level in parentheses. * significant at 10%; ** at 5%; *** at 1%.

Table D.4 reports the results. We first consider the total number of mentions of McKinley and Bryan (divided by the number of eligible voters). We find that the former shows the ex-

pected negative and the later the expected positive sign, but neither coefficient is statistically different from zero (Column [2]). The main effect of speeches is robust to the inclusion of these variables. As an alternative proxy for the relative support for McKinley we consider the ratio of the number of newspaper mentions of McKinley (+1) divided by the number of newspaper mentions of Bryan (+1). As Column (3) shows, this measure is also negative but not significant. Lastly, we interact the relative measure with speeches, to test whether the effect of speeches is smaller in places where McKinley was covered more widely. The interaction effect is negative, but only significantly different from zero when we restrict the sample to counties in which one of the candidates was mentioned at least once (Columns [4] and [5]). Overall, the results provide (weak) evidence that in places were McKinley was covered more widely speeches had a smaller effect, and support the idea that campaign efforts can offset each other.

D.4 Quantification

The quantification exercise that we describe in the paper is presented in Table D.5. As explained in the paper, our counterfactual experiment consists of assuming Bryan performed a traditional "front-porch" campaign, as did McKinley and both Democratic and Republican candidates in campaigns before him. With a budget of $3.6 million, McKinley invited about 500,000 voters in Canton, Ohio. Bryan had a budget of $675,000. This allowed him to reach roughly 4 million voters using the railroad network instead of roughly 100,000 voters had he allocated his budget in a similar way as McKinley ($0.2 instead of $7.2 per voter).

We thus assume a constant persuasion rate (regardless of the campaign type) and recalculate the vote share by county with the effect of speeches being 1/40 of its actual effect. We focus our counterfactual on the findings of the difference-in-difference estimation and of the baseline (within-state) cross-sectional estimation. To calculate the predicted change in vote, we estimate Equations (2) and (1) (Tables C.1 and 2), predict the Democratic vote share, then replace the speech indicator by 1/40, and predict the Democratic vote share once again. The difference between the predicted values of the estimation and the predicted values when replacing a speech by 1/40 give the counterfactual difference.

The results are presented in Table D.5. Columns (1) to (4) present the actual, state-specific outcome of the election. Column (5) presents the change in vote from the difference-in-difference estimates by state; Column (6) presents the results from the cross-sectional estimates. Aggre-

Table D.5: Quantification

	(1)	(2)	(3)	(4)	(5)	(6)
	Actual outcomes				Counterfactual vote if no campaign	
	Speeches	Electoral votes	Democratic vote	Democratic margin	Difference-in-differences	Within-state cross-section
Connecticut	10	6	32.5	–30.7	–0.77	–0.93
Delaware	5	3	43.1	–10.1	–0.74	–0.90
Illinois	135	24	42.7	–13.0	–0.73	–0.89
Indiana	68	15	48.0	–2.9	–0.58	–0.71
Iowa	55	13	42.9	–12.6	–0.33	–0.40
Kentucky	17	13	48.9	–0.1	–0.29	–0.35
Maine	3	6	29.2	–38.7	–0.26	–0.31
Maryland	4	8	41.6	–13.1	–0.56	–0.68
Massachusetts	10	15	26.3	–43.1	–0.70	–0.85
Michigan	78	14	43.5	–10.3	–0.71	–0.86
Minnesota	31	9	40.9	–15.7	–0.41	–0.50
Missouri	29	17	54.0	8.7	–0.37	–0.45
Nebraska	43	8	51.5	5.3	–0.41	–0.49
New Hampshire	3	4	25.9	–42.8	–0.35	–0.42
New Jersey	12	10	36.0	–23.7	–0.52	–0.63
New York	49	36	38.7	–18.9	–0.63	–0.77
North Carolina	14	11	52.6	5.8	–0.25	–0.30
North Dakota	2	3	43.6	–11.9	–0.18	–0.22
Ohio	85	23	47.1	–4.8	–0.64	–0.78
Pennsylvania	16	32	36.3	–24.7	–0.32	–0.39
Rhode Island	1	4	26.4	–41.9	–0.70	–0.85
South Dakota	14	4	49.7	0.2	–0.28	–0.34
Tennessee	16	12	52.1	5.8	–0.22	–0.27
Virginia	6	12	52.5	6.6	–0.11	–0.13
West Virginia	15	6	46.8	–5.4	–0.36	–0.44
Wisconsin	23	12	37.0	–22.9	–0.47	–0.57
Popular vote (1,000)	744	447	7112.1	–601.3	–55.7	–67.8
Popular vote (%)	744	447	46.7	–4.3	–0.50	–0.61
Electoral college	744	320	65	–190	–4	–4
House representatives	591	283	116	–48	–10	–7

Notes: Counterfactual vote in the absence of campaign speeches, based on the results of Tables 2 and C.1 (Presidential) and Tables B.8 and C.2 (Congress). Columns (1) to (4): actual outcomes; Column (5) presents the counterfactual vote in the absence of the campaign using the difference-in-differences point estimates; Column (6) using the within-state cross-sectional point estimates.

gate results indicate that if Bryan had performed a "front-porch" campaign, where he proportionally invited people from the counties where he gave speeches, his vote tally would be lower by 55 to 68 thousand, or 0.5 to 0.6 percentage points. The only state that he would have lost would have been South Dakota (which he won by 183 ballots).

We also perform the exercise for the Congress election to evaluate whether the campaign affected the composition of the House. For our counterfactual exercise, we use the results of Tables B.8 and C.2. The results indicate that the campaign of Bryan resulted in an increase of 7 to 10 congressional representatives for the Democratic party.

References

Atack, Jeremy, "On the use of geographic information systems in economic history: The American transportation revolution revisited," *The Journal of Economic History*, 2013, *73* (02), 313–338.

Bilinski, Alyssa and Laura A Hatfield, "Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions," *arXiv preprint arXiv:1805.03273*, 2018.

Clubb, Jerome M, William H Flanigan, and Nancy H Zingale, *Electoral data for counties in the United States: Presidential and congressional races, 1840-1972*, Inter-university Consortium for Political and Social Research, 1987.

Couttenier, Mathieu, Pauline Grosjean, and Marc Sangnier, "The wild west is wild: The homicide resource curse," *Journal of the European Economic Association*, 2017, *15* (3), 558–585.

DellaVigna, Stefano and Ethan Kaplan, "The Fox News effect: Media bias and voting," *The Quarterly Journal of Economics*, 2007, *122* (3), 1187–1234.

Eichengreen, Barry, Michael Haines, Matthew Jaremski, and David Leblang, "Populists at the polls: Economic factors in the US presidential election of 1896," in "Research in Economic History," Emerald Publishing Limited, 2019.

Gentzkow, Matthew, "Television and voter turnout," *The Quarterly Journal of Economics*, 2006, *121* (3), 931–972.

_, Jesse M Shapiro, and Michael Sinkinson, "The effect of newspaper entry and exit on electoral politics," *The American Economic Review*, 2011, *101* (7), 2980–3018.

Haines, Michael, Price Fishback, and Paul Rhode, *United States Agriculture Data, 1840 - 2012*, Inter-university Consortium for Political and Social Research, 2018.

Hainmueller, Jens, "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies," *Political Analysis*, 2012, *20* (1), 25–46.

_ and Yiqing Xu, "Ebalance: A Stata package for entropy balancing," *Journal of Statistical Software*, 2013, *54* (7).

Jaeger, David A, Theodore J Joyce, and Robert Kaestner, "A cautionary tale of evaluating identifying assumptions: did reality TV really cause a decline in teenage childbearing?," *Journal of Business & Economic Statistics*, 2020, *38* (2), 317–326.

Kearney, Melissa S and Phillip B Levine, "Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing," *The American Economic Review*, 2015, *105* (12), 3597–3632.

Leip, Dave, *Dave Leip's Atlas of the US Presidential Elections*, Dave Leip, 1999.

Pons, Vincent, "Will a five-minute discussion change your mind? A countrywide experiment on voter choice in France," *The American Economic Review*, 2018, *108* (6), 1322–63.

Shaw, Daron R and James G Gimpel, "What if We Randomize the Governor's Schedule? Evidence on Campaign Appearance Effects From a Texas Field Experiment," *Political Communication*, 2012, *29* (2), 137–159.

Spenkuch, Jörg L and David Toniatti, "Political advertising and election results," *The Quarterly Journal of Economics*, 2018, *133* (4), 1981–2036.

Williams, R.H., *Realigning America: McKinley, Bryan, and the Remarkable Election of 1896*, University Press of Kansas, 2010.