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

Johannes C. Buggle[1,*] and Stephanos Vlachos[2,**]

Abstract: This paper examines the effect of campaign appearances in the context of the one-sided nationwide tour by William J. Bryan, the Democratic US Presidential candidate in 1896. During this electoral campaign, Bryan undertook an unprecedented whistle-stop train tour, while the Republican candidate followed a front-porch campaign. To identify the causal effect of campaign speeches, we exploit several estimation strategies, including a within-county difference-in-differences design and a neighbour-pair fixed effect estimator. We find that campaign visits by Bryan increased his vote share by about one percentage point on average. This increase likely stems from the persuasion of previously non-aligned industrial workers.

Keywords: Elections, campaign strategies, persuasive communication

Classification: D72, N41, N71, P48

1 Introduction

Despite the opportunities that modern communication technologies offer to reach an ever-growing audience, traditional campaign appearances remain an important strategy for candidates running for office (Shaw and Gimpel, 2012). This could be because personal contact is a key determinant of the effectiveness

*Correspondence address: Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. johannes.buggle@univie.ac.at

**Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. stephanos.vlachos@univie.ac.at

We are grateful to the editor and four anonymous referees for their helpful suggestions. We also thank Omar Bamieh, Mathieu Couttenier, Elena Esposito, Sophie Hatte, John Vernon Henderson, Seyhun Orcan Sakalli, Mathias Thoenig, and David Yanagizawa-Drott, as well as seminar participants for their valuable comments. Katja Bergonzoli provided excellent research assistance.

of persuasion. Reaching out personally to people during rallies and mass events seems a particularly important campaign strategy for politicians who position themselves as the candidate of the people against established elites (e.g. Kaltwasser *et al.*, 2017; De La Torre, 2018). Personal visits allow to signal the candidate's genuine interest in popular concerns and to display popular appeal to the wider electorate.

In this paper, we evaluate the persuasive effects of local campaign visits in a unique historical context: the first major one-sided nationwide speaking tour by a US Presidential candidate. In the 1896 election, due to limited funding, the Democratic candidate William Jennings Bryan broke with tradition and adopted an unprecedented campaign strategy by using the rail network to go on a national speaking tour. Bryan, also nicknamed 'the Great Commoner', ran on an anti-establishment platform that fought for the interests of the common men who had suffered from the economic depression of the early 1890s and against the elite capitalists, notably through the abolishment of the gold standard.[1] From July to early November 1896, Bryan, an exceptional orator, gave more than 700 speeches, addressing a staggering 4 million voters. In contrast, his opponent the Republican William McKinley followed a traditional front-porch campaign, inviting voters to his house. Bryan went on to lose the election, but his strategy was adopted by the Republican party by 1900.

While campaign appearances remain one of the oldest and most important electoral strategies, causal evidence of its effects on voting outcomes and the mechanisms driving such effects is still limited (Kalla and Broockman, 2018). Our setting has several attractive features that enable us to contribute to this discussion. First, the campaign, due to its novelty, drew immense crowds, increasing the scope for persuasive communication, while also informing on the potential first-mover advantage in adopting new technologies in electoral races. Crucially, as the candidate's goal was the persuasion of the urban working population, the setting allows disentangling the mobilizing and persuasive effects of local visits by politicians. Moreover, from an identification perspective, the focus on a one-sided campaign that is constrained by the lack of detailed information on voter preferences and the pre-existing railroad network helps us to overcome the methodological challenges that the estimation of campaign effects traditionally faces. Finally, while one-sided campaigns are unlikely in pluralist democracies nowadays, the setting provides evidence on the question of campaign effects in situations where one candidate has an exceptionally strong advantage in campaigning.[2]

[1] Contemporary observers have drawn parallels between the political platforms and campaigning styles of Donald J. Trump and William J. Bryan. See for example 'The Trump before Trump' in *The Wall Street Journal*, July 16, 2016. Indeed, both were relatively unknown in the political arena, campaigned in the aftermath of an economic crisis, bypassed traditional party structures, positioned themselves as the protector of the common people against the elite, and used modern technologies to communicate directly with voters (Twitter/railroad). An important difference is that Bryan is typically considered a left-wing populist (see e.g. Acemoglu *et al.*, 2013), while Trump a right-wing one.

[2] In all US Presidential campaigns since the 1960s, both the candidates of the Republican and Democratic parties made personal appearances (Althaus *et al.*, 2002; Wood, 2016). Personal appearances by the major party leaders have become an integral part of electoral campaigns in democracies nowadays, see e.g. Althaus (2009) for Germany, Middleton (2021)

To estimate the effect of speeches on electoral support, we exploit several estimation strategies. In our baseline approach, we estimate cross-sectional within-state OLS regressions. The results suggest that places where Bryan appeared showed increased support for the Democrats in the election, conditional on previous electoral outcomes and observable county characteristics.[3] Yet, the places visited by Bryan were not more likely to vote for the Democratic candidate in any of the previous three elections. Those counties are nonetheless different along several dimensions. Most notably, speech counties are more populated and industrial; as such, they could also differ in other aspects, which would not be captured by our control variables. To address this issue, we formally assess how important selection on unobservables has to be to explain away the effect of Bryan's speeches. We also consider a variety of additional estimation methods, such as neighbour-pair fixed effects estimations, matching counties on observable characteristics, balancing counties on pre-treatment covariates, and within-congressional district estimations. The results confirm the baseline cross-sectional findings.

We then make use of the panel structure of our data and estimate a within-county flexible difference-in-differences specification that accounts for time-invariant unobserved heterogeneity but also allows for the effect of observable characteristics to vary in time. Throughout all specifications, we estimate a positive effect of speeches that is significant and economically meaningful – it amounts to an increase in the Democratic vote share by about one percentage point. We also assess the sensitivity of the difference-in-difference to the specificity of the context. We first estimate the difference-in-differences model using the Republican vote share as the outcome. We then perform a placebo estimation that uses counties in the states that Bryan did not visit to evaluate whether speech counties would have voted for Bryan even in the absence of speeches.

Finally, we examine whether the gains resulted from persuading voters or mobilizing citizens that were already more supportive of the candidate's program. To disentangle these two components, we analyse the effect of speeches on two different sub-groups, industrial workers and farmers. We find that speeches had a larger effect in locations with many industrial workers, suggesting that Bryan succeeded to persuade new voters rather than mobilizing the already predisposed agricultural electorate. We also investigate the origin of the votes gained by focusing on the debate on silver. The findings suggest that the increase in the Democratic vote comes at the expense of 'pro-gold' parties, namely Republicans, the

for the UK, and Schmitt *et al.* (2015) for the European Parliamentary Elections. Some countries have relaxed campaign financing rules recently, such as the United States following the *Citizens United vs. FEC* ruling. Such decisions could, under some circumstances, give an exceptionally strong campaign advantage to one candidate.

[3] Democratic and Populist support in the 1892 election are very strong predictors of the 1896 Democratic vote share, despite the fact that the election of 1896 is often considered to be a critical juncture in American politics (Key Jr, 1955; Azari and Hetherington, 2016).

'narrow-gauge' Prohibition party, and Gold Democrats, while there is no impact on votes for other 'free silver' parties.

Related Literature. A large literature in political science and economics has studied different means to deliver political information and their implications for voter choices. Our paper relates in particular to previous work investigating the effect of campaign visits on voter preferences.[4] Early contributions developed the hypothesis that political campaigns have only minimal effects on votes ('minimal effect hypothesis'). However, this early literature is criticized for struggling in credibly identifying causal effects (DellaVigna and Gentzkow, 2010). In particular, many previous studies looked at the effects of two politicians campaigning simultaneously, which complicated the identification of aggregate effects, since campaigns might cancel each other out (Iyengar and Simon, 2000). The identification of campaign effects could furthermore be contaminated by previous electoral campaigns and by the presence of mass media. The latter might reduce the importance of information transmission of in-person appearances and lower the scope for belief-based persuasion, thus restricting the impact of campaign visits to non-informative dimensions (referred to as preference-based persuasion in DellaVigna and Gentzkow, 2010).

Papers that are more recent find mixed results. Some contributions document a positive effect for one of the candidates running (Jones, 1998; Herr, 2002; Holbrook, 2002). Heersink and Peterson (2017) that look at the campaigns of Truman and Dewey in 1948 attribute campaign effects to the higher quality of the candidate. Other studies find limited or no effects (Wood, 2016; Devine, 2018; Abramowitz and Panagopoulos, 2020). For example, Selb and Munzert (2018) document that speeches given by Hitler did not affect voting in the Reichstag elections, while Jäger (2020) finds that a one-sided re-vote campaign by a minor right-wing party (BIW) had long-lasting effects on votes in one of the precincts of Bremenhaven, Germany.[5]

Our paper contributes to these studies in several important ways: first, we are examining a setting where only one candidate, Bryan, went on the campaign trail - a situation that 'occurs rarely, if at all, in presidential campaigns, the races most often studied.' (Iyengar and Simon, 2000, p. 151). Thus, our setting eliminates the possibility of offsetting campaign messages through rival campaign visits and identifies a general equilibrium effect (see Kalla and Broockman, 2018).[6] Second, we try to improve on causal identification relative to previous studies by formally investigating selection on unobservables and

[4] See Kalla and Broockman (2018) and DellaVigna and Gentzkow (2010) for reviews of field and natural experiments, respectively.

[5] As in our paper, Selb and Munzert (2018) find positive effects in the only one-sided campaign in their sample taking place in 1932. Jäger (2020) looks at electoral outcomes in 5 Bremenhaven precincts over the 2003-2015 period. We improve upon these findings by using a much larger electoral campaign; while Hitler only gave 21 speeches and BIW campaigned in a single precinct, Bryan gave a full 746 speeches in 386 counties.

[6] We do not exclude that the campaign of the rival McKinley offset Bryan's effort through, for example, print propaganda. We analyse this possibility empirically.

by applying a flexible difference-in-differences estimator that effectively controls for local unobservable county characteristics. Moreover, we make progress on the mechanisms behind campaign effects: as Bryan's goal was the persuasion of the urban working population, our setting allows us to disentangle the mobilizing and persuasive effects of local visits by politicians, in the spirit of Pons (2018) and Spenkuch and Toniatti (2018). In our setting persuasion can take place via both informative and non-informative dimensions, as the election takes place in a period when the media is confined to local print media. Finally, since the campaign is the first of its sort, the setting also informs of the potential first-mover advantage in adopting new technologies in electoral campaigns.

Besides campaign visits, our study also relates to the literature on political communication and persuasion more broadly.[7] The expansion of the railroad network was a significant technological innovation in the late 19th century. We therefore also relate to the literature investigating the impact of new technologies on political outcomes, such as the radio (Strömberg, 2004; Adena *et al.*, 2015; Wang, 2021), television (Gentzkow, 2006), and the internet (Campante *et al.*, 2017; Guriev *et al.*, 2021). Lastly, our findings speak to studies documenting the influence of visits of leaders on outcomes such as fertility behaviour, the diffusion of religion, and nation-building (Bassi and Rasul, 2017; Becker *et al.*, 2020; Assouad, 2021).

2 Historical background & data

2.1 The 1896 electoral campaign[8]

The 1896 US Presidential Campaign followed the 1893 financial crisis that led to one of the worst recessions in US history, with decreasing prices, a threefold increase in unemployment, and widespread political discontent (Romer, 1986). The crisis put great pressure on the Government's gold reserves. Debates about whether to abolish the gold standard and to switch to the coinage of silver ('Free Silver') had already appeared in the previous decades. Free silver traditionally had a strong support base in the farming population that hoped for increases in crop prices and an easier payback of their credits. This agrarian movement went on to form the Farmer's Alliance, which became the People's Party in the early 1890s (also known as the Populist Party). The debates on free silver intensified after the financial crisis of 1893 and became a central topic of the 1896 election. 'Silverites' believed that an inflationary expansion

[7] See Gerber *et al.* (2011) and DellaVigna and Kaplan (2007) for two seminal contributions.
[8] This section draws on Jones (1964) unless otherwise stated.

of silver would help the economy to grow. On the other hand, bankers and other creditors, as well as entrepreneurs and industrial capitalists feared increasing costs of production and favoured deflation.

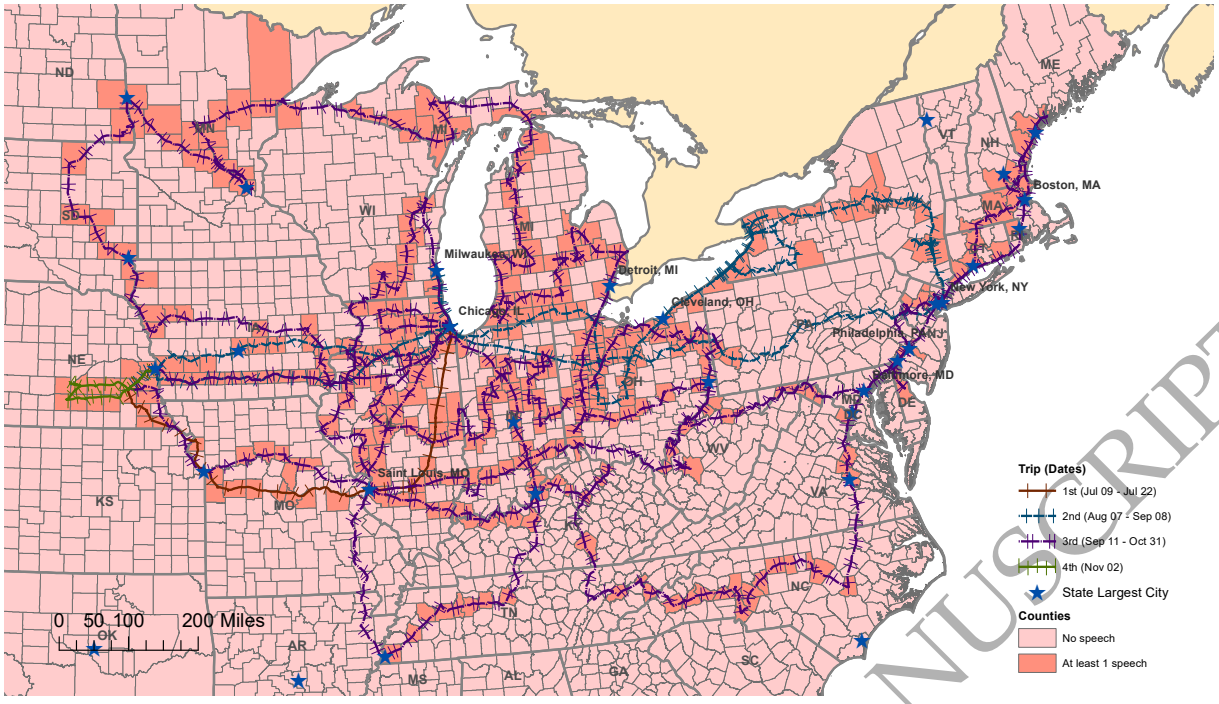
Running on a free silver platform and promoting the fight of farmers and labourers against elite capitalists, William Jennings Bryan became the presidential candidate of the Democrats, the Populist Party, and a branch of the Republicans ('Silver Republicans').[9] His opponent was William McKinley, a Republican that rejected inflationary policies. The electoral landscape was partly pre-determined. While Bryan had a strong voter base in the South and Mountain states, McKinley could be sure to win the Eastern states. The decisive battleground was therefore in the states of the Midwest and those of the border regions to Canada.

To win the presidential race, Bryan and McKinley followed very different campaigning strategies. McKinley could rely on the well-oiled fund-raising machinery of the Republican party that targeted donations from business owners and raised a total of about $3.6 million (Pixton, 1955). McKinley's campaign used the funds to finance rallies and print propaganda that intended to throw mud at the Democratic candidate, calling him a 'Popocrat'. In addition, McKinley invited about 500,000 voters to his home in Canton, Ohio, where he would address his visitors. Bryan's budget, on the other hand, was meagre and amounted to only $675,000 (Pixton, 1955). In light of this financial shortage, Bryan decided it would be cheaper for him to travel personally to the electorate than to bring people to him. Bryan – who appears to have been a mesmerizing orator – hoped that this way he could speak to a much greater number of voters, and convince them to vote for him. While McKinley's staff also considered going on tour, McKinley did not believe that his speaking abilities were on par with Bryan's and responded to his advisers: 'I might just as well put up a trapeze on my front lawn and compete with some professional athlete as go out speaking against Bryan.' (Jones, 1964, p. 277).

In an unprecedented campaigning strategy, Bryan became the first candidate to criss-cross the nation and to address large parts of the public face-to-face using the state-of-the-art mode of transportation, the railroad. Bryan travelled on four separate trips to 546 cities, in 386 counties, and 26 states, covering 18,009 miles according to his diary. He gave a total of 746 speeches in 113 days, addressing an estimated 4 million voters. Bryan's speech was almost always identical and focused heavily on the issue of free silver. Possibly, even more than the speech itself, the in-person contact with the public might have been the most significant feature of Bryan's visits (Young, 2017). The charismatic Democratic candidate attracted large and enthusiastic crowds that would celebrate his appearance with all-day-long festivities and demonstrations. The numerous attendances at his speeches also left an impression on the Republicans

[9] Other issues than silver that Bryan mentioned in his speeches included the income tax, states' rights, and the coercion of farmers and the labouring class by bankers and employers (Jones, 1964, p. 314).

Figure 1: *Bryan's Speeches in the 1896 Campaign*



Note: Map of William Jennings Bryan's presidential campaign in 1896 by trip. County limits as of 1900. Shaded counties are locations with at least one speech. Stars indicate the largest city in each state in the 1890 population census.

who responded with more fiercely attempts to discredit the Democratic candidate. In the early stages of his campaign, it looked like Bryan could secure a win in the important states of the Midwest.

Figure 1 illustrates the spatial extent of Bryan's electoral campaign and the counties in which he spoke at least once. All trips departed from his home base in Lincoln, Nebraska. As Bryan's platform already appealed to the farming population that favoured free silver, he focused a large part of his campaign on the East and Midwest where he intended to gain the votes of the urban labour population. Secret polls conducted months before the election revealed that the overall majority of workers in Illinois and Indiana intended to vote for McKinley (Jensen, 1971).[10] To win the election, Bryan therefore crucially needed to change the political preferences of the urban population. As evident from Figure 1, Bryan seemed to have targeted indeed the large cities where the manufacturing labourers were located.

Although the location of Bryan's appearances in the East and Middle West is consistent with his objective of persuading railroad and factory workers, there was nothing sophisticated about the Bryan's

[10] In the late 19th century only *state-level* straw polls could give some clues about voter preferences. By mid-October 1896 a large number of straw polls had been conducted, indicating favourable results for the Republicans among workers. The *Chicago Tribune* poll surveyed 13,182 working-men, 84% of which were in favour of McKinley; the *Minneapolis Journal* surveyed 1,513 workers, 79% of which declared their intention to vote for McKinley; and the *New York Evening World* surveyed 4,338 workers (59% for McKinley). These results were considered too favourable for the Republicans to be credible (Thies, 2018).

campaign tour. This was partly due to the absence of credible information about voter preferences and to his campaign manager having 'no experience in managing a national campaign' (Jones, 1964, p. 298). The tour of Bryan did not seem to have been carefully planned; it was mostly organized state-by-state and 'often had a thrown together look, with no central authority supervising the arrangements' (Harpine, 2006, p. 16). Bryan would travel on ordinary public trains; his schedule was not strict, even though he had a general, albeit very rough, itinerary plotted for each trip. Especially in the early trips, the planned timetable was often overturned from one day to another if a local leader asked for an appearance (Jones, 1964, p. 311). Bryan also detoured for private reasons. For example, after accepting his party's nomination in New York City, he travelled to up-state New York to visit a former teacher of his, giving several speeches along the way (Jones, 1964, p. 308).[11]

Moreover, due to the lack of credible data, Bryan might have misinterpreted public opinion (Geer, 1991). Possibly, as a result, Bryan campaigned even in states where his chances of winning were small. Harpine (2006, p. 135), for example, wonders about the '[...] second campaign trip to the East. This may have been a curious campaign decision, since he ended up campaigning in States that he must have known he could not hope to win.' Overall, Bryan lost the 1896 Presidential election with 176 electoral votes to 271 won by McKinley – or 46.7% of the popular vote versus a majority of 51.0%. Turnout was almost as high as 80%. Bryan could not win a single state in the Midwest, although he lost in some by close margins: Kentucky by 0.06% (277 ballots) and Indiana by 2.85%. He also lost in a couple of West-coast states by close margins: California (0.64%) and Oregon (2.09%). On the converse, he won in the state of South Dakota by 183 ballots or 0.22% of the vote.

2.2 Data

We employ several sources of data. Summary statistics of the variables described are provided in Table 1. The unit of analysis in this study is a harmonized county. In 1880, 1890, and 1900, there were 2,614, 2,799, and 2,848 counties, respectively (including territories). We harmonize the counties by keeping the largest common area over the three periods. This procedure leads to a set of 2,507 counties; 1,536 counties are in speech states; 1,422 of those had access to railroads in 1896.

[11] We found several historical newspaper articles in which announcements about the upcoming itinerary were made. They vary in detail and there is a fair deal of discrepancy between the speeches announced in newspapers and actual speeches. While some announcements detail most cities in which he plans to give a speech with the date and time, other announcements are coarser and only inform about the state ('through Ohio October 19th and 20th; through Indiana October 21st and 22nd; Illinois from October 23rd to 30th'). Interestingly, the content of the newspaper articles does not appear to be location specific, but the identical route description is given in the many newspapers all over the country.

Campaign speeches. 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' (Thomas III *et al.*, 2017).[12] The data contains for each of the four trips the city in which the speech was given, the date (day/year) and – for a limited number of observations – the time of the day. Information on the crowd size is only available for a small number of speech places. We aggregate the speech data on the county level to merge it with electoral outcomes and other covariates described in the following. As Table 1 reports, Bryan spoke in about 27% of the counties in our sample, which is composed of counties located in states where Bryan gave at least one speech and that have railroad access.[13]

Electoral outcomes. We use data on party vote shares from the *United States Historical Election Returns, 1824-1968* (ICPSR Study no.1) (ICPSR, 1999). The ICSPR dataset contains information on the vote tally of each party for the period from 1824 to 1968.[14] 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 when used as an outcome for the 1892 Presidential election and the 1890 to 1896 Congress. All data is harmonized to our county definitions. At first sight, the average Democratic vote share in the previous election and in 1896 is lower in speech places than in no-speech places, see Table 1. This can be explained by the fact that Bryan targeted his campaign to counties and states in which a Democratic majority was not guaranteed.

County characteristics. We control for several potential correlates of speech locations and voting behaviour measured at the county level. The first set of controls accounts for the geographic location of a county and captures proximity to the political and urban centres, as well as a county's accessibility. 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. We additionally control for the density of the railroad network (length over the area in 1896) using data from Atack (2013). Second, we calculate demographic characteristics of counties based on the decennial censuses available from the 'National Historical Geographic Information System' (NHGIS) (Manson *et al.*, 2021) and the 'Integrated Public Use Microdata Series' (IPUMS) full count datsets (Ruggles *et al.*, 2021). In particular, we compute the total (log) population per county, the population shares of whites, males above 21 years, individuals below 21, and the

[12] The data can be accessed via `http://railroads.unl.edu/topics/bryan.php`. Last accessed on Jan 21, 2022.

[13] Complete data on electoral outcomes and covariates is available for 1,410 of the 1,422 counties that had access to railroads in 1896. Bryan gave a speech in one county for which data is missing (District of Columbia). By restricting the sample to counties with railroads, we are comparing localities where Bryan could have gone and exclude places that he could not have possibly reached from the control group.

[14] Turnout is unavailable in the *United States Historical Election Returns, 1824-1968*. We retrieve this information from Clubb *et al.* (2006). Turnout in Clubb *et al.* (2006) is given as a fraction of the legally eligible electorate.

Table 1: *Summary statistics*

	(1) Mean	(2) Corr 1892 Dem and Pop	(3) Correlation *p-value*	(4) Speech mean	(5) No speech mean	(6) Difference in means	(7) Difference *p-value*
Speech (binary)	0.27	–0.04	0.10	1.00	0.00	1.00	.
Speeches per 10k pop	0.16	–0.02	0.38	0.59	0.00	0.59	0.00
Distance to speech (log)	3.71	0.03	0.29	2.72	4.07	–1.35	0.00
Democratic 1896 (%)	46.20	0.84	0.00	45.69	46.39	–0.70	0.33
Democratic 1892 (%)	43.10	0.54	0.00	44.28	42.66	1.62	0.07
Populist 1892 (%)	9.12	0.24	0.00	7.13	9.85	–2.73	0.00
Dem and Pop 1892 (%)	52.22	1.00		51.41	52.52	–1.11	0.10
Urban county (binary)	0.47	–0.16	0.00	0.77	0.35	0.42	0.00
Population (% of district)	15.83	–0.09	0.00	25.29	12.33	12.96	0.00
Population (log)	9.99	–0.07	0.01	10.50	9.80	0.71	0.00
White population (%)	92.43	–0.32	0.00	95.45	91.32	4.14	0.00
Male above 21 (%)	26.99	–0.38	0.00	28.18	26.56	1.62	0.00
Aged below 21 (%)	34.11	0.09	0.00	33.28	34.41	–1.14	0.00
Native population (%)	88.97	0.28	0.00	86.52	89.88	–3.36	0.00
Literacy (%)	88.53	–0.34	0.00	92.14	87.19	4.95	0.00
News circ (per elig voter)	0.14	–0.03	0.32	0.34	0.06	0.27	0.00
Railroad density	8.39	–0.13	0.00	11.96	7.07	4.89	0.00
Dist to state capital (log)	4.83	–0.01	0.76	4.59	4.92	–0.33	0.00
Dist to state largest city (log)	5.03	–0.01	0.76	4.66	5.17	–0.51	0.00
Latitude	40.44	–0.41	0.00	40.91	40.26	0.65	0.00
Longitude	–86.05	–0.05	0.08	–86.11	–86.03	–0.08	0.86
Farmer households (%)	47.92	0.20	0.00	37.96	51.61	–13.64	0.00
Manufacture workers (%)	3.92	–0.16	0.00	6.34	3.03	3.31	0.00
Crop value change (log)	–0.15	–0.37	0.00	–0.15	–0.15	0.00	0.28
Dist to silver mine (log)	5.72	–0.17	0.00	5.84	5.68	0.16	0.00
Dist to gold mine (log)	5.63	–0.14	0.00	5.86	5.55	0.31	0.00
Observations	1,410	1,410	1,410	381	1,029	1,410	1,410

Notes: All statistics at the county level. Campaign variables come from the University of Nebraska *Railroads and the Making of Modern America - A Digital History Project*. Election data from ICPSR Study no. 1 (ICPSR, 1999). Railroad data from Atack (2013). Population data from NHGIS and IPUMS (Manson *et al.*, 2021; Ruggles *et al.*, 2021). Descriptive statistics for all counties in Column (1). Correlation with the joint Democratic and Populist vote in 1892 in Columns (2) and (3). Characteristics of speech and no-speech counties in Columns (4) and (5), and test of means equality of these two groups in Columns (6) and (7).

share of the district population that lives in the county, to account for differences in the size of the voting population for Presidential and Congress Elections. Besides the share of whites, we capture identity politics with the share of natives. Moreover, we compute the literacy rate for individuals aged 21 and above (to capture educational differences and to approximate income), and whether a county is urban (defined as more than 25,000 inhabitants in urban centres). These censuses are available for 1880, 1890, and 1900. Following Gentzkow (2006), we linearly interpolate the data for inter-census years. Finally, to take into account differential access to the media as a source of information, we use newspaper circulation per eligible voter from Gentzkow *et al.* (2011).

Salience of the silver debate. We include several controls to account for the salience of the silver debate that was central to the campaign. Silver was supported heavily by farmers who worried about falling crop prices that they sought to counter with the coinage of silver to spur inflation. We empirically capture support for the agrarian movement by first controlling for the share of farmers and manufacturing workers in a county (the omitted category being the share of people employed in all other jobs) using information from the NHGIS. Moreover, we control for the local variation in agricultural revenues by combining county-level crop production data with yearly information on national crop prices.[15] Finally, miners also supported free silver. We take variation in mining into account in our empirical specification by controlling for the distance to gold and silver mines; data comes from Couttenier *et al.* (2017).

Observable differences. Table 1 shows that speech counties are different among several dimensions: they are more urban, significantly more populated, and more industrial than no-speech counties. This is in line with the observation that Bryan targeted the urban population which, according to the historical accounts, was not favourable to his central theme of free silver, at least less so than the farming population. The results of the balancedness test imply for our econometric strategy that selection on unobservables (if correlated positively with selection on observables) should work against finding an effect of Bryan's campaign on the Democratic vote share, and possibly introduce a downward bias. We will formally investigate selection on unobservables in the empirical section that follows.

In the Online Appendix, we also examine the determinants of speech locations. We first estimate state-level regressions that reveal that the most important predictor of speeches at the state level is the 1892 joint Democratic-Populist and Republican margin. The choice of states to visit was also driven by the state-level 1892 vote shares of the Democratic and Populist parties.[16] 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 this administrative level. We then turn to lower levels of administration by estimating the same specification within state at the district and county level. These regressions indicate that the main driver of Bryan's speeches was the presence of manufacturing workers. They also reveal that population and urban status were important, as in Kaslovsky (2022). Moreover, accessibility (as captured by railroad density) was an important determinant of the visits. To sum up, the 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) he then chose the counties within these states where there were potentially more voters to persuade, subject to the constraint of railroad accessibility.

[15] See Online Appendix Section **??** for additional information on the construction of the variable. The section documents that crop price fluctuations are a strong predictor of Populist support in the previous election of 1892.
[16] Additional details on the estimation can be found in Online Appendix Section **??**

3 Empirical strategy and results

3.1 Cross-sectional evidence

We begin our investigation into the effect of Bryan's speeches by estimating the following cross-sectional model by Ordinary Least Squares (OLS):

$$dem_c = \beta_0 + \beta_1 speech_c + \beta_2 dem_{c,t-1} + \beta_3 pop_{c,t-1} + \mathbf{z_c'}\gamma + \zeta_s + \varepsilon_c \quad (1)$$

where dem_c indicates the Democratic vote share in county c of state s in 1896. $speech_c$ is a measure of the speeches Bryan gave in county c (more details below), $dem_{c,t-1}$ and $pop_{c,t-1}$ are the vote shares of the Democratic and Populist party in c in the 1892 election, respectively; $\mathbf{z_c}$ is a vector of controls containing demographic and geographic controls, the railroad density, as well as the variables intended to capture the importance of the silver debate (share of farmers, manufacture workers, crop value change, distance to mines) as described in the previous section; ζ_s are state fixed effects. Standard errors ε_c are clustered at the level of the congressional district, since House elections take place on the same day. As the treatment is spatially correlated, we also compute standard errors adjusted for spatial correlation following Conley (1999), assuming a cut-off distance of 85 km which corresponds to the average surface of a district.[17] The coefficient of interest is β_1, the effect of Bryan's speeches on the Democratic party vote.

Table 2 reports the estimation results of Equation (1). In Column (1) we use a binary speech variable that takes the value 1 if Bryan gave at least one speech in the county. The estimated coefficient is positive and highly statistically significant, implying appearances increased the Democratic vote share, conditional on the vote shares in the previous election, demographic controls, and state fixed effects. In Column (2) we add geographic and economic controls, which increases the magnitude of the speech effect to 1.2 percentage points. The magnitude of the effect is comparable to an increase in the 1892 Democratic vote share by 1.5 percentage points (or 0.1 std.dev. of past vote). Note moreover that, the 1892 Democratic and Populist vote shares are very strong predictors of the 1896 Democratic vote share and statistically indistinguishable from each other.

The binary measure of speeches has the advantage of being easy to interpret but does not take into account the frequency of Bryan's visits. In Column (3) we present the results when using the number of speeches normalized by county population. The results are similar, albeit less precisely estimated. This

[17] We use the Stata package *acreg* by Colella *et al.* (2019) to compute Conley standard errors.

Table 2: *Cross-sectional evidence*

Dep. Variable *Democratic vote (%) in 1896*	(1) Speech (binary)	(2) Speech (binary)	(3) Speeches/ 10k pop	(4) Speeches/ 10k pop	(5) Distance (log km)
Speech variable	1.065	1.186	0.778	2.035	–0.277
	(0.316)***	(0.292)***	(0.403)*	(0.457)***	(0.120)**
	[0.331]***	[0.374]***	[0.537]	[0.615]***	[0.152]*
Speech variable squared				–0.583	
				(0.135)***	
				[0.131]***	
Democratic vote (%) in 1892	0.826	0.825	0.825	0.826	0.826
	(0.029)***	(0.028)***	(0.028)***	(0.028)***	(0.031)***
	[0.033]***	[0.030]***	[0.030]***	[0.030]***	[0.030]***
Populist vote (%) in 1892	0.898	0.844	0.843	0.843	0.844
	(0.035)***	(0.034)***	(0.034)***	(0.034)***	(0.038)***
	[0.035]***	[0.038]***	[0.039]***	[0.038]***	[0.044]***
Demographic controls	Yes	Yes	Yes	Yes	Yes
Geographic controls		Yes	Yes	Yes	Yes
Economic controls		Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	46.16	46.20	46.20	46.20	46.49
Mean exp. variable	0.27	0.27	0.16	0.16	3.88
Observations	1,420	1,410	1,410	1,410	1,260
Clusters	228	228	228	228	227
Adjusted R^2	0.85	0.86	0.85	0.86	0.85
Oster's δ	289.6	–20.7	5.0	4.0	2.6
Dem-Pop 1892 coefficient test					
F-statistic ($\hat{\beta}_{dem} - \hat{\beta}_{pop} = 0$)	3.60	0.22	0.20	0.19	0.17
p-value ($\hat{\beta}_{dem} - \hat{\beta}_{pop} = 0$)	0.06	0.64	0.66	0.66	0.68

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 clustered at the congressional district level in parentheses. Standard errors adjusted for spatial correlation in square brackets. * significant at 10%; ** at 5%; *** at 1%.

might be the case because the increase in candidates' popularity is an increasing and concave function of visits, as put forward by Stromberg (2008), an argument that we test in Column (4). Indeed, the point estimates imply that the effect of speeches decreases with the number of speeches (the maximum is reached at 1.75 per 10,000 inhabitants).

Bryan's speeches attracted large crowds, potentially from neighbouring counties. We thus allow the effect to spatially dissipate by using the (log) rail distance of a county to the nearest speech as our treatment variable. Column (6) documents that the Democratic vote share significantly decreases in the distance to a speech. This result suggests that besides the immediate impact in the county where a speech took place, Bryan also influenced vote shares in nearby populations.

Selection on unobservables. The cross-sectional evidence presented in Table 2 could suffer from omitted variable bias if Bryan targeted specific counties with characteristics unobservable to us. To formally test how strong selection on unobservables would have to be to explain away the effect of

Bryan's speeches, we adopt the method developed by Oster (2019) and compute the δ that puts a value on the strength of selection on unobservables. We evaluate the influence of adding additional controls on the stability of the coefficient compared to a regression that only controls for state fixed effect, assuming a maximal R-squared value of $R_{max} = 1$.[18] The δ values reported at the bottom of Table 2 are either negative (indicating that additional controls increase the treatment effect), or substantially larger than 1 (indicating that selection on unobservables has to be much larger than selection on observables). Overall, these results are consistent with the historical observation that Bryan targeted counties that were not inclined to vote for him. They increase our confidence that selection on unobservables does not bias the effect of Bryan's speeches away from zero. In the following, we will use the binary measure as our main treatment as it is easily interpretable, but the results are not driven by this choice.

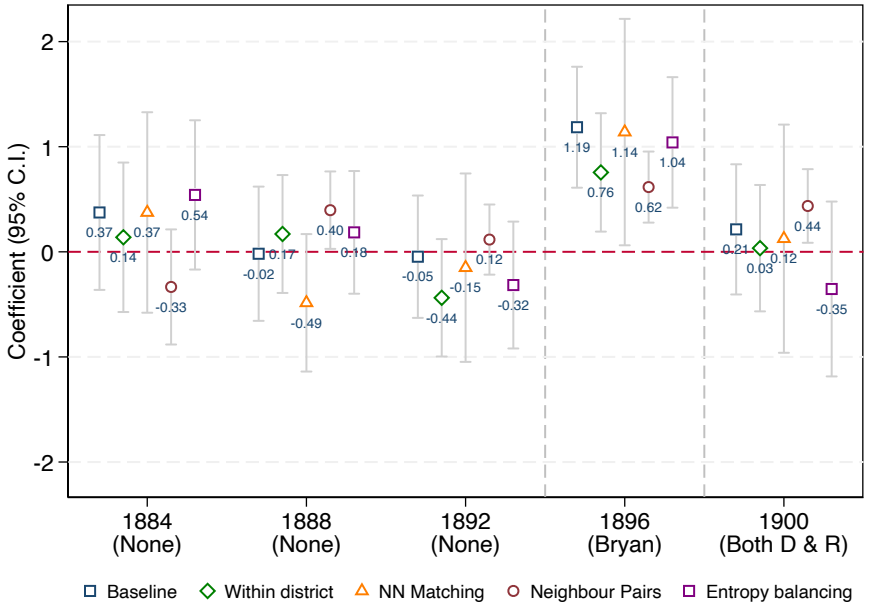
Cross-sectional falsification. To further test the concern that Bryan might have targeted counties with a pre-existing strong Democratic voter base, we estimate Equation (1) using Democratic votes in 1884, 1888, and the joint Democratic and Populist vote in 1892 as our dependent variable. The estimates are presented in Figure 2 (see Table **??** for the regression results). Reassuringly, when estimating the effect of speeches in the elections prior to 1896 (with the full set of controls and state fixed effects), speeches do not have any effect. Compared to the 1896 coefficient, the point estimates are much smaller in magnitude, partly even negative, and highly insignificant throughout. These findings support once more the observation that Bryan did not campaign in locations with pre-existing Democratic support. We also present the coefficient of the estimation for the 1900 election when controlling for the vote shares in the 1892 election. The results imply that the Democratic vote share reverts to its pre-1896 level.

Alternative estimators. Next, we consider a variety of additional estimation methods in the cross-sectional framework as alternatives to the within-state OLS model presented in Equation (1). To ensure that our coefficients do not capture potential differences in local campaigning (for the House election), we first estimate Equation (1) within congressional district (instead of within State). We also estimate the average treatment effect of speeches by nearest neighbour matching. To match speech counties to no-speech counties, we use the full set of controls of Equation (1) and only keep the closest match. The key advantage of this estimator is therefore that it increases the comparability of the treatment and control groups.[19]

[18] Oster (2019) proposes to calculate the $R_{max} = 1.3 \times \tilde{R}$, where $\tilde{R}$ is the R-squared of the fully controlled regression. In our case, $R_{max} = 1.3 \times 0.85 > 1$, which implies $R_{max} = 1$.

[19] Additional details on the different estimation methods and assumptions are provided in Online Appendix Section **??**. We also present the results of the matching estimator under different assumptions, such as allowing for more than one match, and propensity-score matching.

Figure 2: *Cross-sectional falsifications and alternative estimators*



Note: This Figure shows the results of estimating the cross-sectional specification of Equation (1). It displays the coefficients and 95% confidence intervals of the effect of a speech given by Bryan on the Democratic vote share obtained from five different estimation methods in five separate cross-sectional regressions, conditional on the full set of controls. The treatment year is 1896. Information in parentheses below the year indicate which of the candidates campaigned in-person. The 1892 coefficient presents the joint Democratic and Populist share. The corresponding regression results are provided in Table **??**.

In addition, we compare the effect of speeches across adjacent counties. This estimation strategy relies on comparing pairs of neighbouring counties, where only one was exposed to a speech. We identify all county-pairs and regress the Democratic vote share on the speech dummy, controlling for neighbour-pair dummies. The key advantage of the estimator is that it compares close-by counties that share similar (geographic) characteristics. It permits to control for these common characteristics by including fixed effects for pairs of neighbouring counties. However, a caveat of the estimator is that, in the presence of spillovers, the estimate of the speech effect should be downward biased.

Finally, we employ the entropy balancing procedure by Hainmueller (2012), as an additional method to achieve pre-treatment balancedness between treated (speech) counties and control (no-speech) counties. Entropy balancing is a pre-processing procedure that equalizes the moments of the covariates in the two groups. The sample is therefore perfectly balanced (we only balance the first moment). The treatment effect is then estimated on the balanced sample in a weighted regression using the balancing weights.

The results are displayed in Figure 2. The treatment year, 1896, is highlighted in the figure by the two vertical dashed lines. We also present the results for the 1884 to 1892 elections as a falsification.

Diamond symbols represent results from the within-district estimator, triangles represent the nearest neighbour average treatment effect, circles represent the within neighbour-pair estimators and squares the entropy balancing ones. To summarize, throughout all the different estimation methods we detect a significant effect of an 1896 speech that is positive and economically meaningful, as it is located in the range of 0.62 to 1.19 percentage points. The smaller magnitude of the effect for the within-district and the neighbour pair estimators is expected given the likely possibility of spillover effects. The falsification exercise, also displayed in Figure 2, shows that in all estimation methods speeches do not have any effect on the Democratic vote shares in the years prior to 1896.

Sensitivity analysis. In the Online Appendix, we evaluate the sensitivity of the results to alternative clustering rules and sub-samples. We first investigate alternative cut-off distances in computing standard errors that correct for spatial correlation. Figure **??** and Table **??** document that – compared to our baseline distance of 85km – standard errors become smaller when we increase the distance up to 160km or decrease the cut-off distance to 20km. Therefore, the treatment effect remains statistically significant at the 1% significance level.

In Table **??**, we estimate the effect of speeches in various samples. Excluding speeches that were planned (i.e. the more endogenous speeches), we find a somewhat smaller treatment effect ($\hat{\beta} = 1.10$); if instead we exclude non-planned speeches (i.e. the less endogenous speeches), we find a larger treatment effect ($\hat{\beta} = 1.53$). Excluding the direct neighbours of counties with a speech, i.e. eliminating direct spillovers, increases the treatment effect ($\hat{\beta} = 1.97$). The table also documents that the results are robust to excluding localities with multiple speeches, including localities without railroad access, and looking at the full sample of states, including those that Bryan never visited.

In Table **??**, we exploit announcements of the campaign in newspapers. We search the database *Chronicling America*, a collection of digitized newspapers provided by The Library of Congress (2007), for mentions of 'William J. Bryan', as well as 'Itinerary'. We then manually verify the articles to determine whether they indeed announce the upcoming planned campaign itineraries of Bryan. Overall, we are able to identify 78 stops and to draw segments of the itinerary plotted by the Bryan campaign. We then compare counties that lie on the plotted *but not* the actual itinerary to (i) counties in which Bryan gave a speech; (ii) counties in which he gave a speech but were not initially plotted; and (iii) counties in which he did not give any speeches. The results confirm the baseline findings.

Additional findings. Historical accounts suggest that the debate on silver was closely linked to crop price fluctuations (see e.g. Williams, 2010). We formally evaluate the importance of the debate on silver

in Table **??**. The results indicate that: (i) in counties where the value of the agricultural basket decreased between the 1888 and 1892 elections, 1892 Populist support increased, as one would expect, and that (ii) the control variables that are intended to capture the salience of the silver debate are not significant predictors of Democratic vote in 1896 once we control for 1892 Populist vote.

We investigate the effect on turnout in Table **??**. We do not find a difference in the number of people voting in counties that were visited and those that were not, a finding that is consistent with the impact of volunteer campaigns in Pons (2018) and television ads in Spenkuch and Toniatti (2018). Conversely, we document that campaign visits increased the gross Democratic vote share by 0.88 p.p. and decreased the gross Republican vote share by 1 p.p. There are two non-mutually exclusive ways to interpret the fact that, while the turnout rate appears to be unaffected by the campaign, the Democratic gross vote share increases and the Republican one decreases: (i) as evidence of persuasion by Bryan's speeches under the assumption that there are no defiers, as in Pons (2018), or (ii) as speeches mobilizing Democratic voters and demobilizing Republican ones to the same extent, as in Spenkuch and Toniatti (2018).

Finally, in Table **??**, we estimate the cross-sectional model of Equation (1) using the vote shares in the 1896 congressional election, as it took place on the same date as the Presidential election. As shown in the table, the positive effect of Bryan's speeches also materialized in more votes for the Democratic party in the Congress election.

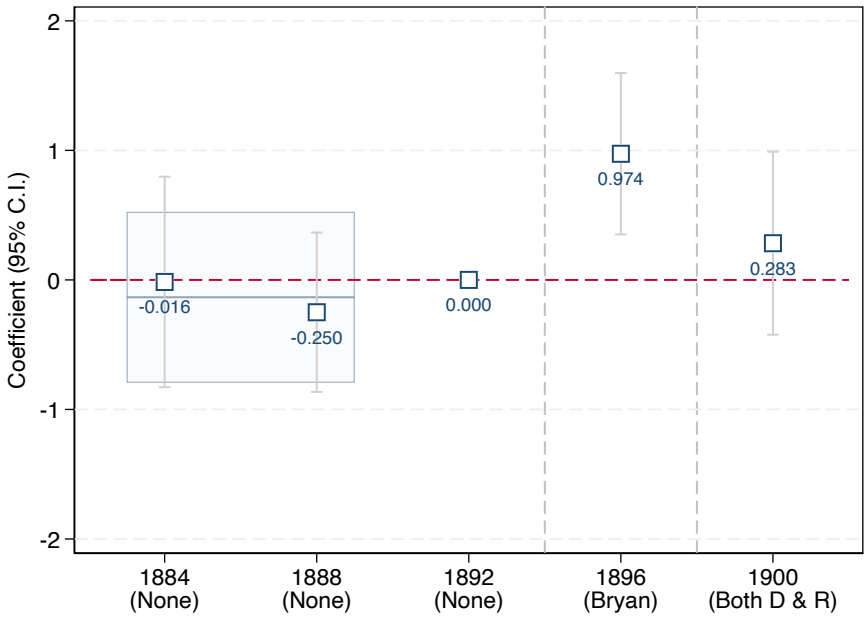
3.2 Difference-in-differences specification

To further address the threat that the estimated effect is driven by unobserved, time-invariant county characteristics, we exploit the panel structure of the data and estimate a difference-in-differences specification that includes county fixed effects. More precisely, we estimate the following model:

$$dem_{ct} = \lambda_{st} + \omega_c + \sum_{\tau=1884}^{1900} \beta_\tau \left(speech_c \cdot election_\tau \right) + \sum_{\tau=1884}^{1900} \left(\mathbf{x}'_{\mathbf{c}\tau} \cdot election_\tau \right) \gamma_\tau + \mathbf{z}'_{\mathbf{ct}} \delta + \varepsilon_{ct} \qquad (2)$$

where dem_{ct} is the vote share of the Democratic candidate in county c in election t (joint Democratic and Populist in 1892) and λ_{st} and ω_c are state $\times$ election and county fixed effects, respectively. The explanatory variable of interest is the interaction between the Bryan speech measure (which is time-invariant) and a full set of dummies for each election year from 1884 to 1900 (the omitted category is 1892). The coefficients β_{1884} and β_{1888} therefore are able to detect pre-existing trends in the Democratic vote share in 1884 and 1888 with respect to 1892. We also control for time-varying county controls

Figure 3: *Difference-in-differences estimation*

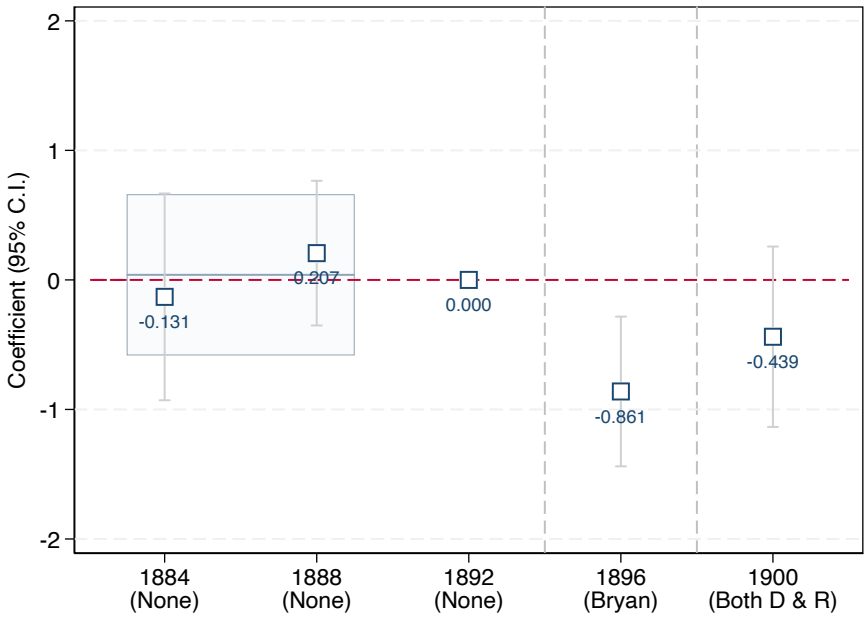


Note: This Figure shows the results of estimating the difference-in-differences specification of Equation (2). It displays the coefficients and 95% confidence intervals of the interaction between a speech given by Bryan and year dummies, conditional on county and state × election fixed effects, as well 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 0.59 and a corresponding p-value of 0.55. The corresponding regression results are provided in Column (1) of Table **??**. 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.

$\mathbf{z_{ct}}$, and allow for economic county controls ($\mathbf{x_{ct}}$) to have a differential effect in each election year (the coefficients γ_t).

Figure 3 summarizes the results of the difference-in-differences model by showing the β_t coefficients, i.e. the interaction between the speech binary variable and the election dummies. The corresponding regression results and point estimates are shown in Table **??**. The treatment year, 1896, is highlighted in the figure by the two vertical dashed lines. The figure illustrates two important results: first, counties with a speech experience an increase in the Democratic vote share of on average 0.97 percentage points in the 1896 election. Second, there are no pre-trends: speech counties were not voting differently in the elections before 1896. The coefficients of the interaction between speeches and the indicators for the years 1884 and 1888 (a test for potential pre-trends) are very close to zero and highly insignificant (p-values of 0.97 and 0.43, respectively, and of 0.55 jointly).

Figure 4: *Difference-in-differences estimation: Republican vote*



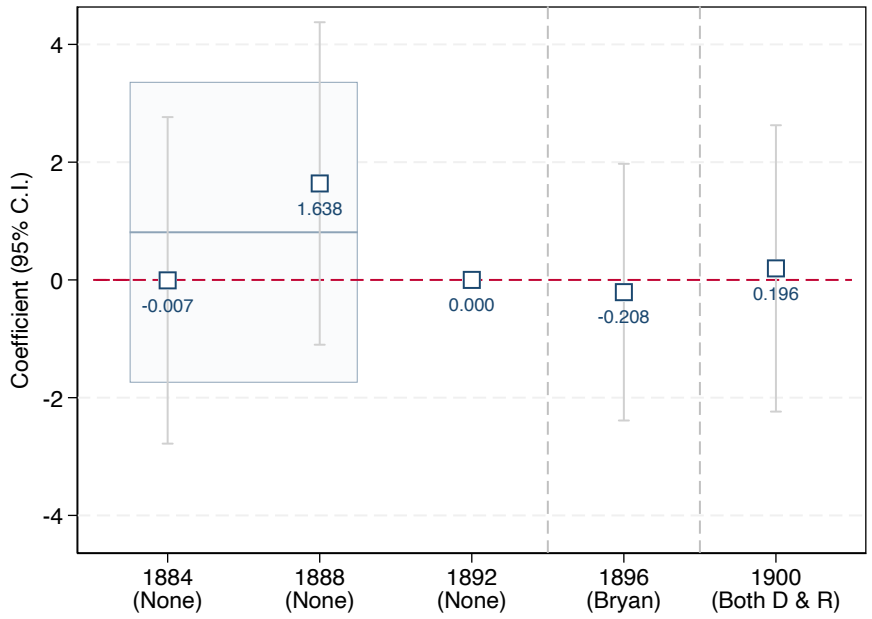
Note: This Figure shows the results of estimating the difference-in-differences specification of Equation (2) using the Republican vote share as the outcome. It displays the coefficients and 95% confidence intervals of the interaction between a speech given by Bryan and year dummies, conditional on county 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 0.86 and a corresponding p-value of 0.42. 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.

Republican vote. The specificity of the context, i.e. the fact that Bryan was the presidential candidate of the Democrats and Populists, challenges the validity of the identification strategy. The first question that arises is whether voting for the Democratic party before 1896 has the same meaning as voting for it in 1896.[20] If this assumption does not hold, then the coefficients on the 1884 and 1888 elections, which serve as pre-treatment falsifications, are not a valid test of the common trends assumption.

To overcome this challenge, we exploit the fact that the Republican party has been more stable in its positions and did not face the same 'structural break' as the Democratic party has, and estimate Equation (2) using the Republican vote as the outcome. Note that in all our estimations we are *not* using the two-party vote share, meaning that the vote shares of the two parties do not sum up to 100% of the vote. The results are presented in Figure 4 for simplicity. Reassuringly, the results on the Republican vote share do not display any pre-trends either. Moreover, the 1896 coefficient is negative, as expected, statistically significant, and closely mirrors the Democratic vote share coefficient.

[20] While in 1892 we use the total vote share of both Democrats and Populists, this issue might be still present for elections before 1892.

Figure 5: *Difference-in-differences estimation: falsification in no-speech States*



Note: This Figure shows the results of estimating the difference-in-differences specification of Equation (2) using counties in states where no speech took place. 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 × 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 0.74 and a corresponding p-value of 0.56. The corresponding regression results are provided in Column (5) of Table **??**. 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.

Pre-trends in Populist vote and falsification. A second threat stems from the possibility that Bryan targeted counties that would have swung for the Democrats in 1896 anyway, for example, because they favoured the Populist platform, or because of their social and economic characteristics.

We address this possibility in two ways. First, we investigate whether the counties, where Bryan gave a speech, revealed a preference for Populist platforms already before 1896. Performing the difference-in-differences estimation for the Populist party in presidential elections in unfortunately infeasible, since the party only presented a presidential candidate in the 1892 election. However, the Populist party was electable in three consecutive congressional elections during the 1890-1894 period. We therefore estimate the difference-in-differences specification for these elections. We find that counties where Bryan gave a speech in 1896 did not vote differently for the Populist party in the preceding Congress elections (see Figure **??** in Online Appendix).

Second, we take advantage of the fact that Bryan only campaigned in the North-East and perform a falsification exercise using the counties in states where he did not campaign (referred to as Eastern and Western states here for simplicity). The falsification exercise consists of identifying the Eastern county that has the most similar observables to a Western county, and attributing to the Western county the treatment status of its most comparable Eastern counterpart.[21] Equation (2) is then estimated using the falsified treatment and control counties in the Western states. The results are presented in Figure 5. In contrast to the actual treatment effect, displayed in Figure 3, there is no detectable effect in the falsified treatment in 1896. The coefficient for 1896 is negative, close to zero, and highly insignificant. The falsification exercise lends support to the argument that counties with similar characteristics to the ones visited by Bryan would not have voted in greater numbers for him in the absence of a speech.

Sensitivity analysis. In the Online Appendix, we report many sensitivity tests that assess the robustness of the results in different samples and the validity of the common trends assumption. We first test whether speeches of Bryan affected the number of people that turned out to vote. As shown in Figure **??** this does not seem to be the case.

We further assess the plausibility of the common trends assumption in Table **??**, in which we test for different slopes between the treatment and control groups in the years preceding the 1896 election, following the 'one step up' approach of Bilinski and Hatfield (2018) (see Section **??** for additional details). We find that the extended model is not superior to the baseline and that the linear trend is not different between the counties in which Bryan gave a speech and those where he did not. We also apply entropy weights to balance the treatment and control groups on pre-treatment (1892) characteristics, as in the cross-section. Applying entropy weights results in treatment effects that are positive and statistically significant, albeit slightly smaller in magnitude, with point estimates that range from 0.71 to 0.91 p.p.

We also investigate the effect of speeches in the differences-in-differences specification in various samples in Table **??**. The main result of a positive treatment effect is robust to (i) excluding speeches that were planned (i.e. the more endogenous speeches), (ii) excluding non-planned speeches (i.e. the less endogenous speeches), (iii) excluding the direct neighbours of counties with a speech, i.e. eliminating direct spillovers, (iv) including localities without a railroad, and (v) looking at the full sample of states, even those that Bryan never visited. The treatment effect varies slightly in the different samples and ranges from 0.80 to 1.28 p.p.

[21] To assign a 'false treatment' status we first estimate a nearest-neighbour matching regression. This allows us to identify, for every county in the Western states, the county in the Eastern states that has the most similar observable characteristics. Under this procedure, 89 of the 735 counties in the Western States are matched to speech counties. See Online Appendix Section **??** for additional information on the procedure.

Table 3: *Where do the votes come from? Silver and gold parties*

	(1)	(2)	(3)	(4)	(5)	(6)
	Free silver in 1896			Pro-gold in 1896		
Dep. Variable	Democratic	National Prohibition	Socialist Labor	Republican	National Democratic	Prohibition
Speech (binary)	1.186***	0.016	–0.004	–1.048***	–0.044	–0.066*
	(0.292)	(0.017)	(0.014)	(0.291)	(0.028)	(0.037)
Democratic vote (%) in 1892	0.825***	–0.002***	0.001**	–0.826***	0.011***	–0.009***
	(0.028)	(0.000)	(0.001)	(0.028)	(0.002)	(0.002)
Populist vote (%) in 1892	0.844***	0.000	0.002**	–0.843***	–0.001	–0.001
	(0.034)	(0.001)	(0.001)	(0.034)	(0.002)	(0.003)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	46.20	0.14	0.08	51.90	0.73	0.94
Mean exp. variable	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
Clusters	228	228	228	228	228	228
Adjusted R^2	0.86	0.71	0.42	0.85	0.67	0.48

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 clustered at the congressional district level in parentheses. * significant at 10%; ** at 5%; *** at 1%.

3.3 Where do the votes come from?

In this sub-section, we investigate who voted more for Bryan. We start by analysing from which parties Bryan gained votes, and focus in particular on the distinction between parties regarding their position in the silver (and gold) debate. We then analyse whether the effect of campaign speeches is strengthened or weakened depending on county characteristics, as well as the timing of a speech relative to the election date. Finally, we incorporate information on the campaign by McKinley into the estimation by looking at the coverage of the two candidates in local newspapers.

Silver and gold parties. In Table 3, we focus on the vote shares of other parties in the 1896 election to investigate the origin of the Democratic vote in speech counties. We classify parties based on their support for the 'free silver' or 'pro-gold' movement. The free silver parties in 1896 were the Democratic party, the National Prohibition party ('broad gauge' Prohibition, which also supported free silver and women's suffrage), and the Socialist Labor party. The pro-gold parties were the Republican party, the National Democratic party ('Gold Democrats'), and the Prohibition party ('narrow gauge' Prohibition, which only supported prohibition).

The findings suggest that the Democratic gains in votes in 1896 stem from 'pro-gold' parties: the vote shares of the Republican party and the 'narrow-gauge' Prohibition party declined significantly in places where Bryan gave a speech, while the effect on the vote share of Gold Democrats marginally fails to reach statistical significance (p-value=0.12). On the converse, the vote shares of neither minor free silver party were affected.

Heterogeneous impacts of Bryan's speeches. According to the historical literature, Bryan's main campaign goal was to convince the urban labour population (Jones, 1964). Column (1) of Table **??** documents that the effectiveness of his speeches indeed increased with the share of the industrial labour population, but not with the share of farmers. Interestingly, the main effect of *Workers (%)* is negative and significant, while the main effect of *Farmers (%)* is positive and significant. This implies that counties with a larger industrial labour population voted less for Bryan if he did not visit them. This interaction effect cannot be explained by a differential impact of his speeches in urban places, see Column (2), or by a preference for the Populist party in the previous election, see Column (3). As Column (4) shows, the effect of Bryan is not different in places with other sources of information measured by the number of newspapers (data from Gentzkow *et al.*, 2011). It is still possible however, that Bryan's speeches had a larger effect in industrial places because they informed workers about his platform. To test for such an informative dimension of speeches, 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.

Finally, speeches given by Bryan closer to the election date have a greater impact than speeches during the earlier trips. This is evident when splitting the speech indicator into the four trips (Column 6), and when interacting the speech with the number of days that passed since the start of the electoral campaign (Column 7). These results thus speak to the literature on short-lived effects of electoral campaigns, as in Gerber *et al.* (2011).

Rival campaign information. A further important question is whether the campaign of McKinley partly offset the campaign efforts of Bryan. Unfortunately, information on rallies or events in support of the Republican candidate is not available. We therefore turn to coverage of the candidates in local newspapers in the months preceding the election (June 1 to November 3, 1896), under the assumption that a greater coverage is indicative of stronger local support for the candidate.[22] Table **??** documents

[22] We draw on a collection of digitized newspapers (*Chronicling America* provided by The Library of Congress (2007)), in which we searched for the mentioning of 'William J. Bryan', as well as 'William McKinley'. As a first step, we document that Bryan was covered more frequently in the week after he gave a speech in a county, while the mentions of McKinley do not react to Bryan's campaign visits (see Figure **??**). This is in line with Shaw and Gimpel (2012) who show that Governor Perry's visits also increase his television and newspaper coverage.

that while the coefficients of the total newspaper coverage of the two candidates in the months before the election have the expected signs, they are very imprecisely estimated. However, we detect a negative interaction effect between speeches and a measure of newspaper coverage of McKinley relative to the coverage of Bryan. This provides some (weak) evidence that speeches had a smaller effect in places where McKinley was covered more widely. It also speaks to the idea that campaign efforts can offset each other.

3.4 Taking stock: Quantifications

This section tries to assess the overall impact of the Bryan campaign on the National popular vote, the Electoral College, and the House of Representatives composition. Our counter-factual 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 to Canton, Ohio. Bryan had a budget of $675,000. With this budget he was able 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).

To construct our counter-factual vote, we 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.[23] The state-by-state and National results indicate that the campaign increased the votes at the national level by roughly 60,000 (or roughly 10% of 600,000 in the National difference). Nationwide, this translates into an increase of approximately 0.5 percentage points in the popular vote. Bryan's campaign probably resulted in his very close win in South Dakota and its 4 Electors. No other state electoral outcome would have been overturned. Moreover, the Bryan campaign resulted in 7 to 10 additional seats for the Democratic party in the House of Representatives.

This result can also give a sense of the persuasion rate of the campaign speeches (DellaVigna and Kaplan, 2007). On average, 20% of the population of exposed counties attended a speech. Actual turnout in exposed counties was not different from turnout in the control ones (see Table **??**). The estimated counter-factual Democratic vote from the difference-in-differences estimation (Figure 3), would be 46.2%. This implies that the persuasion rate of the railroad campaign was 4% (sd = 5%). The within-state cross-sectional estimation gives a persuasion rate that is of a similar magnitude (5%). These persuasion rates are in line with the rates proposed in the literature (between 2% and 20%).

[23] In other words, in our counter-factual we assume that Bryan would have invited 100,000 voters from the locations in which he gave speeches. State-by-state results of this counter-factual exercise are presented in Table **??**.

4 Conclusion

In this paper, we evaluate the effect of campaign appearances relative to a 'front-porch' campaign in a unique historical context: the first whistle-stop tour by a Presidential candidate in the 1896 election. We show, through different identification strategies, that campaign visits increased the Democratic party vote share by about one percentage point on average. Our analysis suggests that it is unlikely that this effect is upward biased due to selection on unobservables, and we regard the effect size as a lower bound. We also find that the increase in votes stems from persuading the previously non-aligned industrial labour population, and is not just a mobilization effect that materializes in higher turnout.

Our results imply that campaign effects do not need to be minimal if campaigning efforts and skills are unbalanced. While one-sided electoral campaigns are unlikely in pluralist democracies, some countries have relaxed campaign financing rules recently, for example, the United States following the *Citizens United vs. FEC* ruling. Such decisions could, under some circumstances, give an exceptionally strong campaign advantage to one candidate. Our findings indicate that, in a context of severe political polarization, while lopsided campaigns matter, it is probably unrealistic to assume that campaign efforts can completely overturn an electoral result.

More broadly, our findings shed light on the survival of traditional campaign speeches in the age of mass and social media. Even today, in-person visits remain a very important electoral campaign strategy; during the 2016 Presidential election, Donald Trump addressed the crowds directly 73 times and Hillary Clinton made 79 campaign visits.[24] In the presence of mass media that reduce the informational content of campaign visits (belief-based persuasion), non-informative dimensions may matter more (preference-based persuasion), such as candidate attributes or the mere presence of a Presidential candidate. While this dimension is not thoroughly explored here, we do believe that this unique context could be exploited to distinguish between these two types of persuasive communication.

5 Supplementary data

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.6998089.

[24] Source: *The American Presidency Project*, University of California at Santa Barbara.

Affiliations

[1]Department of Economics, University of Vienna, AT.

[2]Department of Economics, University of Vienna, AT.

References

Abramowitz, A. and Panagopoulos, C. (2020). 'Trump on the trail: Assessing the impact of presidential campaign visits on voting behavior in the 2018 midterm elections', *Presidential Studies Quarterly*, vol. 50(3), pp. 496–506.

Acemoglu, D., Egorov, G. and Sonin, K. (2013). 'A political theory of populism', *The Quarterly Journal of Economics*, vol. 128(2), pp. 771–805.

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V. and Zhuravskaya, E. (2015). 'Radio and the rise of the Nazis in prewar Germany', *The Quarterly Journal of Economics*, vol. 130(4), pp. 1885–1939.

Althaus, M. (2009). 'German elections and modern campaign techniques', in (D. W. Johnson, ed.), *Routledge Handbook of Political Management*, pp. 295–316, Routledge, New York et al.

Althaus, S.L., Nardulli, P.F. and Shaw, D.R. (2002). 'Candidate appearances in presidential elections, 1972-2000', *Political Communication*, vol. 19(1), pp. 49–72.

Assouad, L. (2021). 'Charismatic leaders and nation building', Working Paper, Paris School of Economics.

Atack, J. (2013). 'On the use of geographic information systems in economic history: The american transportation revolution revisited', *The Journal of Economic History*, vol. 73(02), pp. 313–338.

Azari, J. and Hetherington, M.J. (2016). 'Back to the future? what the politics of the late nineteenth century can tell us about the 2016 election', *The ANNALS of the American Academy of Political and Social Science*, vol. 667(1), pp. 92–109.

Bassi, V. and Rasul, I. (2017). 'Persuasion: A case study of papal influences on fertility-related beliefs and behavior', *American Economic Journal: Applied Economics*, vol. 9(4), pp. 250–302.

Becker, S.O., Hsiao, Y., Pfaff, S. and Rubin, J. (2020). 'Multiplex network ties and the spatial diffusion of radical innovations: Martin Luther's leadership in the early reformation', *American Sociological Review*, vol. 85(5), pp. 857–894.

Bilinski, A. and Hatfield, L.A. (2018). 'Nothing to see here? non-inferiority approaches to parallel trends and other model assumptions', *arXiv preprint arXiv:1805.03273*.

Campante, F., Durante, R. and Sobbrio, F. (2017). 'Politics 2.0: The multifaceted effect of broadband internet on political participation', *Journal of the European Economic Association*, vol. 16(4), pp. 1094–1136.

Clubb, J.M., Flanigan, W.H. and Zingale, N.H. (2006). 'Electoral data for counties in the united states: Presidential and congressional races, 1840-1972', Inter-university Consortium for Political and Social Research [distributor], `https://doi.org/10.3886/ICPSR08611.v1`.

Colella, F., Lalive, R., Sakalli, S.O. and Thoenig, M. (2019). 'Inference with arbitrary clustering', IZA Discussion Paper No. 12584.

Conley, T.G. (1999). 'Gmm estimation with cross sectional dependence', *Journal of Econometrics*, vol. 92(1), pp. 1–45.

Couttenier, M., Grosjean, P. and Sangnier, M. (2017). 'The wild west is wild: The homicide resource curse', *Journal of the European Economic Association*, vol. 15(3), pp. 558–585.

De La Torre, C. (2018). 'Populism revived: Donald Trump and the latin american leftist populists', *The Americas*, vol. 75(4), pp. 733–753.

DellaVigna, S. and Gentzkow, M. (2010). 'Persuasion: Empirical evidence', *Annual Review of Economics*, vol. 2(1), pp. 643–669.

DellaVigna, S. and Kaplan, E. (2007). 'The Fox News effect: Media bias and voting', *The Quarterly Journal of Economics*, vol. 122(3), pp. 1187–1234.

Devine, C.J. (2018). 'What if Hillary Clinton had gone to Wisconsin? presidential campaign visits and vote choice in the 2016 election', *The Forum: A Journal of Applied Research in Contemporary Politics*, vol. 16(2), pp. 211–234.

Geer, J.G. (1991). 'Critical realignments and the public opinion poll', *The Journal of Politics*, vol. 53(2), pp. 434–453.

Gentzkow, M. (2006). 'Television and voter turnout', *The Quarterly Journal of Economics*, vol. 121(3), pp. 931–972.

Gentzkow, M., Shapiro, J.M. and Sinkinson, M. (2011). 'The effect of newspaper entry and exit on electoral politics', *The American Economic Review*, vol. 101(7), pp. 2980–3018.

Gerber, A.S., Gimpel, J.G., Green, D.P. and Shaw, D.R. (2011). 'How large and long-lasting are the persuasive effects of televised campaign ads? results from a randomized field experiment', *American Political Science Review*, vol. 105(01), pp. 135–150.

Guriev, S., Melnikov, N. and Zhuravskaya, E. (2021). '3g internet and confidence in government', *The Quarterly Journal of Economics*, vol. 136(4), pp. 2533–2613.

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

Harpine, W.D. (2006). *From the Front Porch to the Front Page: McKinley and Bryan in the 1896 Presidential Campaign*, vol. 13, Texas A&M University Press.

Heersink, B. and Peterson, B.D. (2017). 'Truman defeats dewey: The effect of campaign visits on election outcomes', *Electoral Studies*, vol. 49, pp. 49–64.

Herr, J.P. (2002). 'The impact of campaign appearances in the 1996 election', *Journal of Politics*, vol. 64(3), pp. 904–913.

Holbrook, T.M. (2002). 'Did the whistle-stop campaign matter?', *PS: Political Science and Politics*, vol. 35(1), pp. 59–66.

ICPSR (1999). 'United states historical election returns, 1824-1968', Inter-university Consortium for Political and Social Research [distributor], 1999-04-26. `https://doi.org/10.3886/ICPSR00001.v3`.

Iyengar, S. and Simon, A.F. (2000). 'New perspectives and evidence on political communication and campaign effects', *Annual Review of Psychology*, vol. 51(1), pp. 149–169.

Jäger, K. (2020). 'When do campaign effects persist for years? evidence from a natural experiment', *American Journal of Political Science*, vol. 64(4), pp. 836–851.

Jensen, R.J. (1971). *The Winning of the Midwest: Social and Political Conflict, 1888-1896*, vol. 2, University of Chicago Press.

Jones, J.M. (1998). 'Does bringing out the candidate bring out the votes? the effects of nominee campaigning in presidential elections', *American Politics Quarterly*, vol. 26(4), pp. 395–419.

Jones, S.L. (1964). *The Presidential Election of 1896*, University of Wisconsin Press.

Kalla, J.L. and Broockman, D.E. (2018). 'The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments', *American Political Science Review*, vol. 112(1), pp. 148–166.

Kaltwasser, C.R., Taggart, P.A., Espejo, P.O. and Ostiguy, P. (2017). *The Oxford Handbook of Populism*, Oxford University Press.

Kaslovsky, J. (2022). 'Senators at home: Local attentiveness and policy representation in congress', *American Political Science Review*, vol. 116(2), pp. 645–661.

Key Jr, V.O. (1955). 'A theory of critical elections', *The Journal of Politics*, vol. 17(1), pp. 3–18.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T. and Ruggles, S. (2021). 'Ipums national historical geographic information system: Version 16.0 [dataset]', Minneapolis, MN: IPUMS. `http://doi.org/10.18128/D050.V16.0`.

Middleton, A. (2021). *Communicating and Strategising Leadership in British Elections: Follow the Leader?*, Palgrave Studies in Political Leadership, Cham: Springer International Publishing.

Oster, E. (2019). 'Unobservable selection and coefficient stability: Theory and evidence', *Journal of Business & Economic Statistics*, vol. 37(2), pp. 187–204.

Pixton, J.E. (1955). 'Charles G. Dawes and the McKinley campaign', *Journal of the Illinois State Historical Society (1908-1984)*, vol. 48(3), pp. 283–306.

Pons, V. (2018). 'Will a five-minute discussion change your mind? a countrywide experiment on voter choice in france', *The American Economic Review*, vol. 108(6), pp. 1322–63.

Romer, C. (1986). 'Spurious volatility in historical unemployment data', *Journal of Political Economy*, vol. 94(1), pp. 1–37.

Ruggles, S., Fitch, C.A., Goeken, R., Hacker, J.D., Nelson, M.A., Roberts, E., Schouweiler, M. and Sobek, M. (2021). 'IPUMS Ancestry Full Count Data: Version 3.0 [dataset]', Minneapolis, MN: IPUMS, 2021.

Schmitt, H., Hobolt, S. and Popa, S.A. (2015). 'Does personalization increase turnout? Spitzenkandidaten in the 2014 european parliament elections', *European Union Politics*, vol. 16(3), pp. 347–368.

Selb, P. and Munzert, S. (2018). 'Examining a most likely case for strong campaign effects: Hitler's speeches and the rise of the Nazi party, 1927–1933', *American Political Science Review*, vol. 112(4), pp. 1050–1066.

Shaw, D.R. and Gimpel, J.G. (2012). 'What if we randomize the governor's schedule? evidence on campaign appearance effects from a texas field experiment', *Political Communication*, vol. 29(2), pp. 137–159.

Spenkuch, J.L. and Toniatti, D. (2018). 'Political advertising and election results', *The Quarterly Journal of Economics*, vol. 133(4), pp. 1981–2036.

Strömberg, D. (2004). 'Radio's impact on public spending', *The Quarterly Journal of Economics*, vol. 119(1), pp. 189–221.

Stromberg, D. (2008). 'How the electoral college influences campaigns and policy: The probability of being Florida', *The American Economic Review*, vol. 98(3), pp. 769–807.

The Library of Congress (2007). 'Chronicling America: Historic American newspapers', [Washington, D.C.: Library of Congress]. `https://chroniclingamerica.loc.gov/`.

Thies, C.F. (2018). 'Polls and elections: The Chicago record poll and the election of 1896', *Presidential Studies Quarterly*, vol. 48(1), pp. 127–138.

Thomas III, W.G., Healey, R. and Cottingham, I. (2017). 'Railroads and the making of modern america', Center for Digital Research in the Humanities, University of Nebraska–Lincoln, [distributor], `https://railroads.unl.edu`.

Wang, T. (2021). 'Media, pulpit, and populist persuasion: Evidence from Father Coughlin', *The American Economic Review*, vol. 111(9), pp. 3064–92.

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

Wood, T. (2016). 'What the heck are we doing in Ottumwa, anyway? presidential candidate visits and their political consequence', *The Annals of the American Academy of Political and Social Science*, vol. 667(1), pp. 110–125.

Young, J.C. (2017). *The Age of Charisma: Leaders, Followers, and Emotions in American Society, 1870–1940.*, Cambridge University Press.