



Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Does fake news affect voting behaviour? ☆

Michele Cantarella ^{a,b,*}, Nicolò Fraccaroli ^{c,d}, Roberto Volpe ^e^a University of Helsinki, Finland^b University of Modena and Reggio Emilia, Italy^c Brown University, United States of America^d University of Rome Tor Vergata, Italy^e LUISS Guido Carli, Italy

ARTICLE INFO

Keywords:

Fake news

Mass media and political communication

Voting behaviour

Populism

ABSTRACT

We study the impact of fake news on votes for populist parties in the Italian elections of 2018. Our empirical strategy exploits the historical variation in Italian-speaking and German-speaking voters in the Italian region of Trentino Alto-Adige/Südtirol as an exogenous source of assignment to fake news exposure. Using municipal data, we compare the effect of exposure to fake news on the vote for populist parties in the 2013 and 2018 elections. To do so, we introduce a novel indicator of populism using text mining on the Facebook posts of Italian parties before the elections. Our findings support the view that exposure to fake news favours populist parties regardless of prior support for populist parties, but also that fake news alone cannot explain most of the growth in populism.

1. Introduction

Over the last years, elections in Western democracies have been preceded by the dissemination of online political fake news — i.e., intentionally fabricated misinformation with politically-charged content. This proliferation, aided by the growth of personalised filter bubbles online (Sunstein, 2018), has been followed by the electoral success of populist candidates. The spread of fake news before the 2016 US Presidential election was largely in favour of Donald Trump (Allcott and Gentzkow, 2017), whose campaign has been consistently characterised by a populist rhetoric (Norris and Inglehart, 2019). In the 2017 French Presidential election hoaxes took an unambiguous stance in favour of Front National's candidate Marine Le Pen (Barrera et al., 2020), who

reached the runoff.¹ Misinformation is also believed to have supported the electoral growth of the far right party Alternative Für Deutschland in the 2017 German election (Zimmermann and Kohring, 2020), which became the third most voted party in the country. Fake news also spread before the 2018 Italian election, which led to the first populist majority in Western Europe (D'Alimonte, 2019). In this context, fact-checkers reported that within the three months preceding the election the second most shared news online was false and directed against the incumbent party.²

Does fake news have a causal effect on the electoral success of populist parties? We address this question focusing on the case of the 2018 Italian election in the region of Trentino Alto-Adige/South Tyrol. Located on the border with Austria, this region hosts two main

☆ The authors would like to thank Thorsten Beck, Markus Cappello, Paloma Díaz Topete, Stefano Gagliarducci, Stephen Hansen, Alessandro Pizzigolotto, Alberto Franco Pozzolo, Chris Redl, Marco Schito, Giancarlo Spagnolo, Arthur Turrell, Anzelika Zaiceva and two anonymous referees for their invaluable suggestions and advice. We are grateful to participants to the Bank of England-King's College conference 'Modelling with Big Data and Machine Learning: Interpretability and Model Uncertainty', the 13th CESifo Workshop on Political Economy, the 2021 American Political Science Association (APSA) Annual Meeting in Seattle (WA), and the SMYE 2021 as well as to participants to the seminars at the University of Helsinki, University of Rome Tor Vergata, the IRVAPP of the Bruno Kessler Foundation in Trento, the University of Modena and Reggio Emilia and the 'Economics of Populism' seminar series at the University of Rome 'La Sapienza' for their comments. All remaining errors are ours.

* Corresponding author at: University of Helsinki, Finland.

E-mail address: michele.cantarella@helsinki.fi (M. Cantarella).

¹ BBC 'Fake news: Five French election stories debunked', 15 March 2017. Link: <https://www.bbc.com/news/world-europe-39265777> (last accessed on 30 April 2020).

² Bruno, Nicola (2018); 'Satira e fake news: gli articoli più condivisi delle Elezioni 2018'. Available at: <https://tg24.sky.it/politica/2018/03/08/fake-news-elezioni-2018.html> (last accessed on 30 April 2020).

³ More precisely, our control group is made of Italian voters speaking a language other than Italian. While this group is dominated by German-speaking voters, other groups in the region speak Ladin, Mocho and Cimbrian. As we explain in further detail in the paper, our instrumental variable is the share of Italian-speaking voters in each municipality, which is therefore compared to the share of non-Italian speaking voters.

<https://doi.org/10.1016/j.respol.2022.104628>

Received 5 May 2021; Received in revised form 23 June 2022; Accepted 5 September 2022

Available online 20 September 2022

0048-7333/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

communities of Italian voters: one that is Italian-speaking and the other that is German-speaking. Since only a small portion of the local population is proficient in both languages (Baggio, 2020; Ebner, 2016; Abel et al., 2012), the membership to each linguistic group serves as an ideal exogenous factor to determine voters' exposure to fake news during the electoral campaign. This source of random allocation is explained by the fact that, in the months preceding the elections, fake news on the Italian elections were in Italian only, and not in German. As a result, Italian-speaking voters in the region were more likely to be exposed to misinformation than their German-speaking peers. In this unique setting the group of German-speaking Italian voters³ serves as an ideal control group to study the effect of fake news on Italian-speaking voters (the treated group).

We measure exposure to fake news by computing for each municipality of the region the number of likes to Facebook pages that disseminated fake news in the months before the elections. Using the historical proportion of Italian-speaking voters in each municipality as an instrumental variable (IV), we test the effect of an additional like to a fake news Facebook page – our treatment variable – on the electoral support for populist parties in the 2018 elections. We combine the IV with municipality fixed effects to compare this outcome with the one of the 2013 Italian election. This instrumented difference-in-differences approach allows us to compare the voting trends of the two linguistic communities and hence to filter out community-specific political preferences, including the time-invariant preference in the German-speaking community for German-speaking parties, such as the Südtiroler Volkspartei.

By exploiting the multi-party structure of the Italian political system, we look at the effect of fake news on populist voting on a continuous scale. To this aim, we construct a text-based indicator of populist rhetoric. We apply this measure to the universe of Facebook posts published by the running political parties and their leaders three months before the 2013 and 2018 election. In this way, we obtain a continuous indicator of populist rhetoric that characterises each party with different intensity. This gives us more precise estimates of populism than the ones derived from binary classifications of populist and non-populist candidates generally used in other works.

Our results have two implications. On the one hand, they show that fake news has had a positive effect on the votes for populist candidates in the 2018 Italian election, and that this effect cannot be explained by prior political beliefs. On the other hand, they show that fake news alone cannot explain most of the growth in populist preferences. In particular, we find that in 2018 an additional like to a Facebook page that disseminated fake news resulted in a 0.14 average increase in the electoral support for populists. Therefore, other economic and political factors may have played a more prominent role in leading the recent increase in support for populists (for a review of these drivers, see Guriev and Papaioannou, 2020 and Guriev, 2018).

The main contribution of this work is to identify the causal link between fake news and populist vote. Existing works on the topic highlight a strong correlation between exposure to misinformation and populist voting (Allcott and Gentzkow, 2017; Swire et al., 2017; Guess et al., 2018; Gunther et al., 2019). However, they do not account for the selection bias linked to fake news exposure: voters of populist parties can self-select into misinformation bubbles and consume fake news that confirms their political preferences (Liberini et al., 2020; Boutyline and Willer, 2016). Other papers overcome this issue by relying on data from surveys or experiments (Zimmermann and Kohring, 2020; Barrera et al., 2020; Munger et al., 2020). The limitation of this second set of works is that they capture the effect of hoaxes on voting intentions rather than actual voting behaviour. Voting intentions may in fact be short-lived and hence tend to overestimate actual voting (Gerber et al., 2009, 2011; Chiang and Knight, 2011). Our work overcomes this limitation by looking at the impact of fake news on the actual electoral outcomes in each municipality.

In addition, this paper is the first to provide empirical evidence on the role of fake news in the 2018 Italian election. This election is an interesting case study since it was the first in Western Europe where populist parties won a majority (D'Alimonte, 2019). Furthermore, while in the other cases there was a single populist party or candidate (e.g. Trump, Le Pen, etc.), in the 2018 Italian election there were at least two large populist parties. This enables to reduce party-specific confounding factors when studying the impact of misinformation on populism.

Finally, our results enrich the literature on the role of media and the internet on electoral behaviour. There is extensive evidence on how traditional media, such as newspaper (Gerber et al., 2009; Chiang and Knight, 2011; Drago et al., 2014), television (DellaVigna and Kaplan, 2007; Gerber et al., 2011; Peisakhin and Rozenas, 2018) and the internet (Jaber, 2013; Larcinese and Miner, 2012; Miner, 2015; Falck et al., 2014; Lelkes et al., 2017; Gavazza et al., 2018; Poy and Schüller, 2020; Munger et al., 2020) affect voting behaviour.

Recent works highlight how media favoured the growth of populism in Italy. Durante et al. (2019) show that individuals with early access to Berlusconi's private TV network in the 1980s were more likely to vote for Berlusconi's party, Forza Italia, in 1994. Moreover, they find that these voters were made more vulnerable to populist rhetoric, as the municipalities exposed to Berlusconi's TV shifted their support to the Five Star Movement (M5s, hereafter) in 2013. Campante et al. (2017) find that in 2013 broadband internet favoured the electoral growth of the M5s, as it mobilised through the internet those voters disenchanted with establishment parties. Our findings add to this literature, showing that fake news contributed to this link by mobilising populist voters.

A caveat that applies to our results is that, in exploiting the unique presence of two linguistic communities in a specific region, the effects we identify are naturally limited in their generalisability. However, these findings are key to understanding how individuals react to political misinformation beyond the context of a survey or an experiment.

The remainder of this paper is organised as follows. Section 2 provides relevant background on the diffusion of fake news in the 2018 Italian election and on the Trentino and South Tyrol region. Section 3 describes the data, while Section 4 outlines how we use social media and electoral data to construct the indicator of populism, which will be our dependent variable. Section 5 describes the empirical strategy. Section 6 presents and discuss the results. Section 7 concludes.

2. Background

2.1. Fake news in the 2018 Italian general election

According to several journalistic⁴ and institutional⁵ sources, the campaign period leading to the 2018 Italian general election saw a remarkable spread of 'fake news'. Such misinformation had the common feature of being intentionally fabricated and published on social media non-institutional outlets. While not all fake news stories had a politically-charged content, those with a clear political target were highly diffused and had the highest reach.

By tracking a set of politically-charged keywords via a content analysis tool, an Italian news channel⁶ reported that among the top 100 articles in Italian for social media engagement, five were hoaxes, while another ten were classified as reporting real events but out of context or omitting relevant details. While few in number, these hoaxes

⁴ See BuzzFeed News (2017); One Of The Biggest Alternative Media Networks In Italy Is Spreading Anti-Immigrant News And Misinformation On Facebook; available at: <https://www.buzzfeed.com/albertonardelli/one-of-the-biggest-alternative-media-networks-in-italy-is>; last accessed: 25 May 2019.

⁵ See Autorità per le Garanzie nelle Comunicazioni (2018).

⁶ Bruno, Nicola (2018); see footnote 2.

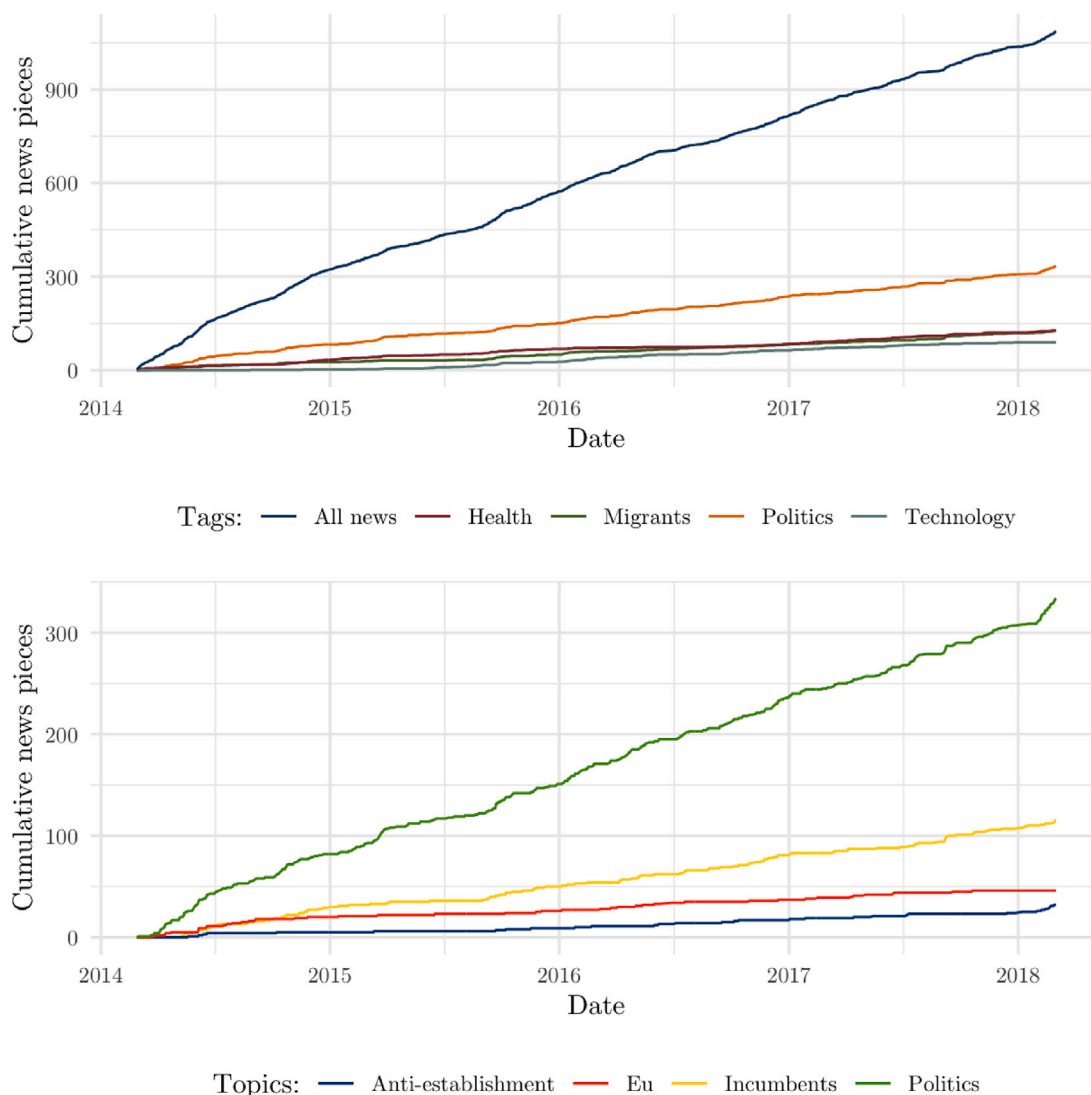


Fig. 1. Cumulative growth of fake news pieces reported by debunking websites between 2013 and 2018 by tags and topic.

had a significant outreach. The second-most shared online news in the database, published on the day before the election, received more than 140,000 interactions, mostly on Facebook. The news consisted of an entirely unsubstantiated report of voter fraud planned by the incumbent Democratic Party in Sicily.

‘Purely false’ news, as the latter example, and non-traditional information sources have a predominantly ‘anti-establishment’ (and, by extension, ‘anti-incumbent’) character. The report by Giglietto et al. (2018) provides a detailed classification of news sources in the lead-up to 2018 general election based on partisanship of their news content and evidences that the vast majority of ‘non-institutional’ websites are biased in favour of Lega and M5s. Crucially, comparable biased sources supporting other parties (including smaller ‘anti-establishment’ ones) captured much less social media attention than pro-M5s and Lega sources. The report stops short of establishing a link between the spread of false information in the electorate and the support for Italian populist parties and their policy stances. Nonetheless, an investigation by Avaaz⁷ provides evidence in support of this link, at least as far as Facebook is concerned.

⁷ La Repubblica (2019); Facebook chiude 23 pagine italiane con 2.4 milioni di follower; available at: https://www.repubblica.it/tecnologia/social-network/2019/05/12/news/facebook_chiude_23_pagine_italiane_con_2_4_

As a complement to these findings, we provide further evidence for the persistence of an anti-establishment bias in misinformation by studying their contents. We do so by analysing entry metadata scraped from an independent Italian debunking website (Butac.it) and retrieving all post tags, along with the text of all available entry headlines and fake news transcriptions. Our sample comprises all fake news propagated between February 2014, the date the first available debunking article was published on Butac.it, and March 2018, on the day of the 2018 elections.

We sort debunking articles by their tags, distinguishing generic fake news pieces from those related to politics, migration, health and technology. Using a simple text mining technique, we search for recurring mentions of parties and leaders of incumbent/establishment and challenging/anti-establishment platforms, along with mentions of EU institutions, across headlines and fake news transcriptions.

Fig. 1 plots the cumulative growth of fake news pieces over the years. Interestingly, the rate of growth of all fake news pieces is linear after 2015, albeit this might be related to the debunking website reaching its full capacity to process and fact check all entries. The

milioni_di_followers_diffondevano_fake_news_e_parole_d_odio-226098817/; last accessed: 13 May 2019.

top figure highlights how, while the bulk of false information is not political in its nature, political misinformation has played a large role in contributing to the growth of the fake news phenomenon, with over 300 pieces of fake news of a political nature being published since 2014. Increases in the count of politically-charged misinformation can, nonetheless, be detected in proximity with the 2018 general elections.

The bottom figure supplements this information by looking at the contents of each fake story in the political subset. Nearly than 1/3 of all pieces of misinformation with a political slant directly targeted the incumbents. More importantly, starting from late 2015, the rate of growth of fake headlines related to the incumbent platform increases significantly, signalling a change in the strategy of disseminators. The growth of false news pieces related to the anti-establishment platform is, instead, much more contained.

In addition, the few pieces of misinformation that mention explicitly an anti-establishment party or candidate take a broadly supportive stance in their favour, meaning that the actually damaging headlines are even sparser than the figure suggests. Indeed, between 2016 and 2018, only six⁸ among all headlines reported by Butac.it could be said to have an anti-M5s or anti-Lega slant, while only two were directed against the neo-fascist parties Casapound and Forza Nuova.

Moreover, the descriptive evidence presented so far shows that in terms of social media exposure misinformation in support of anti-establishment parties, and in particular of the Lega and the M5s, spread much more successfully than that of their pro-establishment counterparts in the run-up to the 2018 general election. While this does not necessarily mean that the spread of fake news was an integral part of populists' strategy and campaign, it further confirms the pro-populist character of fake news, in line with what has been found for other countries.

2.2. Trentino and South Tyrol: Political and sociolinguistic background

The region of Trentino Alto-Adige/Südtirol is subdivided into two autonomous provinces (*province autonome*) with very similar characteristics: Trento and Bolzano – Alto-Adige (in Italian) or Bozen – Südtirol (in German; 'South Tyrol' henceforth). Both provinces became part of the Italian unitary state after World War 1, when the Treaty of Saint-Germain-en-Laye assigned the southern part of Tyrol, composed by the present-day provinces of Trento and Bozen, to the Kingdom of Italy. Before that, the two provinces shared a common history as part of the Princely County of Tyrol until 1804, as a crown land of the Austrian Empire from 1804 to 1867 and as a Cisleithanian crown land of Austria-Hungary until 1919.

Other than history, the two provinces share socio-economic similarities. With 540,000 inhabitants (Istat, 2018) Trentino is only slightly more populated than South Tyrol, which has 530,000 inhabitants. Both provinces are highly rural, with a large share of the population living outside of the few mid-sized urban centres,⁹ scattered across hundreds of very small municipalities. Moreover, they are comparable in terms of size of the electorate and income per capita (for more details see Table 1). From an administrative perspective, both provinces enjoy a large degree of self-government compared to other Italian local authorities, with the *provincia autonoma* having significant legislative, fiscal and budgetary autonomy.

While similar in many regards, the two provinces are characterised by a long-standing linguistic diversity. The most spoken language is German in South Tyrol, whereas in Trentino is Italian. This linguistic divide has existed well before the 19th century. When they were part of the Austrian crown land of Tyrol, 96.4% of the 362,684 inhabitants province of Trento was Italian-speaking, while only 2.9% was

German-speaking. On the other hand, 83.4% of the inhabitants of South Tyrol was German-speaking, 13.2% Ladin-speaking and 2.6% Italian-speaking (Boelitz, 1930). The linguistic divide persists today: in 2011, 69.7% of South Tyrol's population declared German as its first language (Astat, 2011), while Italian speakers represent slightly less than one quarter of all inhabitants (118,000 people). In contrast, Trentino is mostly Italian-speaking with some small Ladin-speaking communities in the North-Eastern side. Italian and German are by far the most represented languages. The average municipality in the region has 60% of its inhabitants who speak Italian, 33.4% German, and 5% Ladin, while the remaining 2% speaks another language.

While both Italian and German are compulsory subjects for members of both language groups from age six, effective bilingualism is not widespread. A number of studies (for a review see Ebner, 2016), point to a relatively low second language proficiency (L2) of the South Tyrolean population. In particular, less than 10% of 17–18 year old high school pupils of either language group are proficient in the other language (above C1 CEFR level).

The linguistic diversity in the region is likely to affect media consumption. As a striking example, in 2016 *Dolomiten*, the main German-language newspaper of South Tyrol, boasted a circulation of 42,103,¹⁰ which is more than four times higher than that of its local Italian-language counterpart, *Alto Adige*, and over ten times that registered by the most common mainstream newspapers in Italy. German-speaking voters are hence likely to be exposed to a different media bubble than the Italian-speaking ones. This historical linguistic divide provides for an ideal exogenous allocation into two separate online media bubbles.

The language divide has also affected the political dynamics of the two provinces. While the focus on devolution certainly favoured the emergence of pro-autonomy local parties in the Trento province, its voting history, especially in national elections, does not dramatically diverge from neighbouring areas. Conversely, in South Tyrol the party system has been dominated by the South Tyrolean People's Party (Südtiroler Volkspartei, 'SVP') since the birth of the Italian Republic in 1946. The SVP is a catch-all party featuring both conservative and social-democratic wings and its voting patterns reflect the weight of the German (and Ladin) community. The party received a majority of the votes cast in the Province in every election between 1948 and 2008. In spite of its clear ethnolinguistic character, the SVP can hardly be considered an 'anti-establishment' platform, as it has regularly coalesced with traditional political forces such as the Christian Democracy party (*Democrazia Cristiana*) before 1992 and with the centre-left *Partito Democratico* (PD) and its predecessors in more recent times.

In parallel with the electoral decline of the centre-left in Trentino, the provincial vote of the SVP has progressively eroded. In general elections, the SVP has mostly faced provincial-wide competition from national parties only. In the last decades, many disenfranchised voters have dropped their support for the SVP, while its supremacy has been challenged by the localist radical-right party *Die Freiheitlichen* (DF), which openly advocated for secession or reunification with Austria. While the DF forfeited in 2018, the SVP party still suffered a remarkable decrease in absolute votes (from 147,000 to 127,000) as turnout in the areas with the highest representation of German-speaking voters fell from over 80% to slightly more than 50% of the electorate. A large number of spoilt and blank ballots was also recorded.¹¹

Since to a certain extent linguistic communities spill over political preferences, our research design will compare the trends rather than the levels in support for populism. For this reason, our data, outlined in the next section, will focus on both the 2013 and 2018 election. Moreover, differences in voting trends will affect our empirical approach, and will be discussed again in Section 5.

¹⁰ Accertamenti Diffusione Stampa (ADS) (2016), *Dati medi annuali territoriali per testata: diffusione cartacea Italia*, available at: <http://www.adsnotizie.it/grafici.asp?>; Last accessed 08/2018.

¹¹ Italian Ministry of the Interior, Archivio storico delle elezioni. Available at: <https://elezionistorico.interno.gov.it/>.

⁸ See Appendix D.

⁹ Istat (2015); Principali dimensioni geostatistiche e grado di urbanizzazione del Paese; istat.it; Available at: <https://www.istat.it/it/archivio/137001>; Last accessed: 25/05/2019.

Table 1
Summary statistics by province and year.

	(1) Bolzano/Bozen			(2) Trento			(3) Difference		
	2013	2018	Average	2013	2018	Average	2013	2018	2018–2013
Populist score (total)	1070.366 (2397.416)	1169.605 (2835.977)	1119.985 (2620.651)	739.040 (2164.160)	1057.371 (2771.920)	898.205 (2488.243)	331.326 (275.970)	112.234 (336.141)	–118.440 (262.998)
Exposed to fake news	0.000 (0.000)	376.876 (2366.432)	188.438 (1680.340)	0.000 (0.000)	521.874 (3459.335)	260.937 (2456.570)		–144.998 (340.984)	
Broadband connections	2469.032 (7542.113)	3007.955 (7661.528)	2738.494 (7590.388)	1987.041 (6763.404)	2217.626 (7030.532)	2102.334 (6889.395)	481.991 (866.188)	790.329 (887.056)	–0.012 (0.045)
Income per capita (natural log)	9.828 (0.142)	9.928 (0.141)	9.878 (0.150)	9.743 (0.140)	9.817 (0.106)	9.780 (0.129)	0.085*** (0.017)	0.112*** (0.015)	0.032* (0.013)
Electorate size	3245.284 (7548.803)	3332.853 (7683.594)	3289.069 (7600.120)	2276.602 (6986.644)	2330.716 (7112.244)	2303.659 (7039.726)	968.682 (876.694)	1002.138 (892.387)	538.524 (695.252)
Observations	232			352			292		

Notes: Mean coefficients, standard errors in parentheses.

*p < .05; **p < .01; ***p < .001

3. Data

Due to the lack of individual-level micro-data on fake news exposure and voting preferences, we use municipalities as the principal unit of analysis of this study.

We collect information on voting preferences, exposure to fake news, and important socio-economic features such as linguistic group shares, average income levels and broadband internet coverage from a number of different sources. Moreover, we generate information on the populist stance of each party running for election in 2013 and 2018, using data from electoral campaigns on social media. Summary statistics on the core variables used in our analysis are presented in Table 1. The table also uses independent sample T-tests (columns set 3), showing the differences between the two provinces for each election year, and the differences in the changes between the two elections. Alone, these statistics are not particularly informative without controlling for electorate size, but it is worth noting that the magnitude (and significance, in case of income) of these figures is greatly reduced when looking at the differences across the years (column 2018–2013), hinting at the presence of parallel trends between the two provinces which motivate our empirical strategy.

Appendix A provides an overview on how these sources were gathered and harmonised. With regards to our main variables of interest, populist voting preferences are produced using the methodologies presented in Section 4. Exposure to fake news is defined as the number of likes to Facebook pages that disseminate fake news for each municipality, and further information on the collection process is available in Appendix B.

4. Measuring populism

4.1. Two text-based indicators of populism

To estimate the impact of fake news on populism, we construct a dependent variable that captures the electoral performance of populist parties. Identifying a party as populist is not a trivial task, since the literature offers no univocal definition of populism (Norris, 2019). We therefore use a minimalist definition of populism as a form of discourse based on its tendency to adopt an emotional language (Taggart, 2000; Rooduijn, 2014; Caiani and Graziani, 2016; Bischof and Senninger, 2018). As a robustness check, we will adopt an alternative definition of populism based on parties' anti-elite/establishment rhetoric, which is another aspect of populist rhetoric highlighted in the literature (Mudde, 2004; Albertazzi and McDonnell, 2008; Pauwels, 2011; Rooduijn and Pauwels, 2011; Kaltwasser et al., 2017; Norris and Inglehart, 2019; Hopkin, 2020).

To capture these two aspects, we create two indexes that measure the degree of populist rhetoric of each party based on the text of their

Facebook posts. The first “assertive” indicator counts the number of exclamation marks as a proxy of an emotional tone. Kumar and Sebastian (2012) show that the number of exclamation marks used in a post is a good predictor of emotional statements on social media. The second index is based on a dictionary of 23 key terms which capture a “anti-establishment/aggressive” tone in the Italian political language, such as ‘establishment’ or ‘caste’. This index is similar to the ones used in previous works to measure populism on party manifestos (Rooduijn and Pauwels, 2011; Pauwels, 2011) and parliamentary speeches (Decadri and Boussalis, 2019).

We compute these indexes on the text of all the posts of Italian parties (12,159 posts) and their political leaders (8,164 posts) published three months before the Italian elections of 2013 and 2018, which we web-scrape from Facebook. After pre-processing the text,¹² we match the words in the dictionaries with those in the Facebook posts collected. Through the match, we obtain two alternative scores of populism (one based on emotional language, the other on anti-establishment rhetoric) for each party, leader and election.

To obtain a more precise measure of a party's degree of populism, we compute for each text bag the following indicator:

$$Popscore_{tk} = \sum_{d=1}^D (Count_{tkd}/T_{tkd}) * Engagement_{tkd} \quad (4.1)$$

where *Count* is the number of words matched for each text bag in each Facebook post $d = \{1, \dots, D\}$ published by each party/leader $k = \{1, \dots, K\}$ three months ahead of each election $t \in \{2013, 2018\}$. *T* is the total number of words contained in each Facebook post *d*. We divide the number of matches by *T* to avoid that longer (shorter) posts score higher (lower) due to their length. *Engagement* is the sum of all Facebook shares and likes a given post has received.

We multiply the scores by *Engagement* to assign higher values to those posts that were more visible on Facebook before the elections, due to their higher number of shares and likes. In this way, our index measures not only the degree of populism of each post, but also their online visibility, through the number of likes and shares given to each post. Hence, exogenous factors affecting the nation-wide demand for populism are taken into account: in line with the intuition of Guiso et al. (2017) that populism results from the interplay of factors of demand and supply, Facebook posts which received little engagement should not contribute in defining the perception of a party as “populist” as much as posts with high engagement and diffusion.

In a nutshell, *Popscore* is given by the sum of the number of matches of exclamation marks (or populist words for other index) in each post

¹² To allow for matching, we tokenise the text, remove stopwords, punctuation (except exclamation marks), numbers and white spaces and transform all terms in lower case. We also translate each term of the text-bag in German to compute scores also for German-speaking parties.

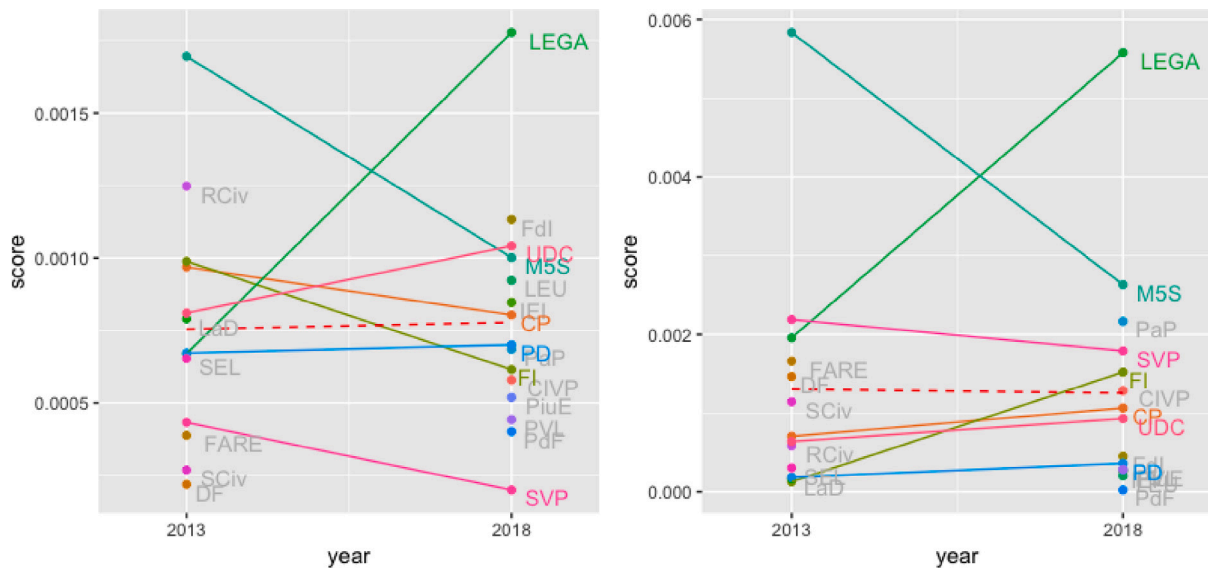


Fig. 2. Text analysis scores $\times 1000$ of social media posts from parties and their leaders during the 2013 and 2018 elections campaigns. Text bags: left figure, “anti-establishment/aggressive”; right figure, “aggressive”. Parties in grey have only took part in one of the two elections. The dashed lines refer to election-specific averages.

(Count), weighted by the total number of words (T) and by the number of interactions ($Engagement$) for each party.

This measure has the advantage of assigning to each party a continuous score of populism, rather than separating them into two distinct groups. This allows for comparing different intensities of populism both between and within parties. Capturing parties’ within-variation in their supply of populism is particularly important for our research question. It could be argued that the traditional dominance of the SVP in German-speaking municipalities might have limited the penetration of populist parties such as M5S and Lega in such municipalities. By adopting a continuous measure of populism, we overcome this limitation, as we do not exclude that the SVP might have adopted a populist rhetoric as well. Our measure captures both the degree of populist rhetoric supplied by SVP as well as the intensity of social media engagement such rhetoric received on Facebook. Our approach does not preclude to any party – even those not traditionally regarded as populist – the possibility of having engaged in populist rhetoric.

Fig. 2 shows values of $Popscore$ for each party and elections for both the dictionary (panel on the left) and the exclamation marks (right). The two indexes display similar trends, suggesting that anti-establishment discourses and emotional tones generally go hand in hand in political communication. The Lega features as the most populist party in 2018 for both indexes, as a result of a steep increase in populist language compared to 2013. On the other hand, M5S shows a lower degree of populist language in 2018 than in 2013 in both indexes. The Democratic Party (PD) displays relatively low levels of populism. Smaller fringe parties, such as the far right Casapound (CP) and the far left Potere al Popolo (PaP), report high values in at least one indicator, in line with previous findings in the literature (Ernst et al., 2017).

A few counter-intuitive results can be explained through a deeper analysis of political dynamics. The clearest example is the Christian-democratic UDC: this party presents a lower score in 2013, when it ran with a centrist pro-incumbent coalition, compared to 2018, when it coalesced with the opposition. An inspection of the words matched under the first methodology shows a clear change in communication strategy, with terms like ‘scandal’ or ‘shame’ featuring more frequently in 2018.

These results are in line with the widespread categorisation of parties such as Lega and M5S as populist. As a further test, we compare our results to the values assigned by political experts in the Chapel Hill Expert Survey (Polk et al., 2017) to a number of parties for the

variable ‘People vs the Elites’, which measures the ‘salience of anti-establishment and anti-elite rhetoric’. Text analysis scores display a positive correlation with the anti-establishment indicator based on experts’ perceptions (Fig. E.1 in Appendix E).

4.2. Combining populist scores with electoral results

Since our aim is to measure the electoral success of populist parties, we multiply the populist score for party k in election t ($Popscore_{tk}$) by the votes gained by the same party p in municipality i during election t . Formally, we aggregate the scores as follows:

$$Y_{it} = \sum_{k=1}^K Popscore_{tk} * Votes_received_{tki} \tag{4.2}$$

where $Votes_received$ is the number of votes received by each party k in each election t in each municipality i . Y_{it} will be our indicator of populism at municipal level and our main dependent variable. Given these computations, the absolute values of the score are low by construction. For this reason, we focus on the relative distance in the scores between parties.

As our final index combines a number of elements including the votes gained by each party and its engagement on Facebook, Table 2 decomposes the index for each party. Most parties enhanced their presence of social media in 2018, increasing the number of posts published and the number of likes and shares received by their followers. The table shows raw results for the populist score (scaled by a factor of 100). In order to make the results more easily interpretable, in the regressions we standardise the scores such that we can measure the relative distance in populist rhetoric of each party from the most populist party in the sample (the Lega in 2018). In particular, we divide the score of each party by the score of the most populist party to compute their distance from the highest supply of populism.¹³

The parties that published the most on Facebook in 2018 are by far the Lega and the M5S, in line with the findings of Giglietto et al. (2018). The Democratic Party ranks third in number of Facebook posts, but scores low in terms of populist language compared to the Lega and M5S, reflecting its moderate stances and position as incumbent. The

¹³ For instance, the score of Forza Italia in 2018 is $0.152/0.558 = 0.272$, where 0.152 is the populist score of Forza Italia in 2018, whereas 0.558 is the score of Lega in 2018.

Table 2
Elements composing the populist score, by political party and year.

Party	Votes received at national level (%)		Populist Score × 10k (assertive approach)		Number of posts (Party + Leader)		Engagement (Party + Leader)	
	2013	2018	2013	2018	2013	2018	2013	2018
Movimento 5 Stelle (M5S)	25.56	32.68	0.584	0.263	1798	2020	4,247,119	12,739,192
Partito Democratico (PD)	25.43	18.76	0.018	0.036	719	1062	559,794	3,889,217
Lega	4.09	17.35	0.196	0.558	615	1994	37,211	14,322,120
Forza Italia ^a (FI)	21.56	14.00	0.013	0.152	197	668	968,151	1,730,078
Fratelli d'Italia (FdI)	1.96	4.35	N/A	0.045	N/A	1580	N/A	4,570,784
Liberi e Uguali (LeU)	N/A	3.39	N/A	0.020	N/A	356	N/A	546,399
Più Europa (PiuE)	N/A	2.56	N/A	0.028	N/A	500	N/A	640,781
Unione di Centro (UDC)	1.79	1.30	0.064	0.093	304	396	103,722	194,736
Potere al Popolo (PaP)	N/A	1.13	N/A	0.217	N/A	992	N/A	650,732
Casapound (CP)	0.14	0.95	0.070	0.107	753	835	146,040	1,930,393
Italia Europa Insieme (IEI)	N/A	0.58	N/A	0.026	N/A	541	N/A	100,496
Civica Popolare (CIVP)	N/A	0.45	N/A	0.128	N/A	429	N/A	184,056
Südtiroler Volkspartei (SVP)	0.43	0.41	0.219	0.179	385	373	3371	10,368
Il Popolo della Famiglia (PdF)	N/A	0.67	N/A	0.002	N/A	420	N/A	160,875
Partito Valore Umano (PVL)	N/A	0.15	N/A	0.029	N/A	41	N/A	13,388
Die Freiheitlichen (DF)	0.14	N/A	0.146	N/A	270	N/A	2200	N/A
Scelta Civica (SC)	8.30	N/A	0.114	N/A	494	N/A	420,434	N/A
Sinistra, Ecologia e Libertà (SEL)	3.20	N/A	0.030	N/A	829	N/A	847,083	N/A
Rivoluzione Civile (RCiv)	2.25	N/A	0.058	N/A	607	N/A	319,163	N/A
Fare per Fermare il Declino (FARE)	1.12	N/A	0.166	N/A	526	N/A	570,311	N/A
La Destra (LaD)	0.65	N/A	0.016	N/A	519	N/A	68,383	N/A

Note: N/A denotes when a party did not exist, ceased to exist, or did not run in an election. Engagement is the sum of the likes, shares and comments received on each posts by the party and leader Facebook page.

^aForza Italia ran as Popolo delle Libertà in 2013.

low engagement of Forza Italia, the third most voted party in 2013, is not surprising. The party’s leader, Silvio Berlusconi, had a very low social media presence and opened his Twitter account just in October 2017 (Giglietto et al., 2018). Forza Italia is outperformed by small fringe parties, such as Casapound, which was very active on Facebook.

5. Empirical framework

Our aim is to understand if the diffusion of fake news in 2018 was responsible for the electoral growth of populist parties. The null hypothesis is that fake news exposure has had no effect on votes for populist parties, as populist voters tended to be more exposed to fake news due to their ex-ante political preferences.

In order to test these hypotheses, we first estimate OLS regressions of fake news exposure on electoral support for populism, controlling for a vector of covariates that can predict the growth of populism. Formally, we estimate the following equation:

$$Y_{it} = \alpha + (F_i \times P_t)\lambda + P_t\zeta + X'_{it}\gamma + \delta_i + e_{it} \tag{5.1}$$

This is a standard difference-in-differences equation with two-way fixed effects. The outcome variable is Y_{it} , which captures the electoral success of populist parties for each municipality i and electoral year $t \in [2013; 2018]$. In our baseline model Y is the indicator computed in Eq. (4.2), standardised so that the LEGA party has a score of 1, with all other parties following. We will replace this indicator with different measures of populist voting to test the robustness of our results. Municipality fixed effects are captured by δ_i , while the time effects are controlled by P , a dummy that equals 1 for electoral year 2018, i.e. the election preceded by the diffusion of fake news. X' is a vector of covariates including broadband connection, income per capita, electorate size, abstentions and invalid votes.

F_i is the treatment indicator and captures exposure to fake news measured as the number of likes to Facebook pages that disseminate fake news. As discussed earlier, fake news virtually did not exist before the 2013, so time-heterogeneity is captured by the treatment indicator and the interaction with time, so that $F_{it} = F_i \times P_t$.

Our coefficient of interest is λ , which captures the variation of the electoral support for populism for one additional like to a Facebook page disseminating fake news in 2018 in municipality i . If exposure

to fake news increases the electoral support for populism beyond each individual it reached, then we would expect $\lambda > 1$.

Eq. (5.1) is informative about the correlation between fake news and populism, but not on the direction of causation. Nonetheless, the magnitude of λ is still relevant to our research question as it captures both the potential effect of fake news on populist voting and the selection bias of populist voters into misinformation bubbles. We therefore need to compare λ with an alternative estimate that isolates the effect of fake news from the selection bias.

Based on the linguistic differences in Trentino Alto-Adige, we exploit the historical share of Italian-speaking voters in each municipality, denoted by Z , as an instrument for exposure to fake news on Italian politics. Formally, we estimate the following two-stage specification:

$$(F_i \times P_t) = \alpha_1 + (Z_i \times P_t)\lambda_1 + P_t\zeta_1 + X'_{it}\gamma_1 + \delta_{1i} + e_{1it} \tag{5.2}$$

$$Y_{it} = \alpha_2 + (\widehat{F_i \times P_t})\lambda_2 + P_t\zeta_2 + X'_{it}\gamma_2 + \delta_{2i} + e_{2it} \tag{5.3}$$

Exposure to fake news is captured by F in the first-stage regression (5.2), where its value – and the value of its interaction with the year of election in (5.2) – is predicted by the instrument Z and other covariates. The fitted values $\widehat{F_i \times P_t}$ from the first stage regression are then plugged into Eq. (5.3) to predict populist preferences Y .

Essentially, we develop a diff-in-diff model where treatment – exposure to fake news – is predicted by the exogenous assignment to the language group. In this way, the relationship between linguistic groups and exposure to fake news can be tested in the first stage of our model (5.2), rather than naively assuming that the German-speaking population is completely unexposed to fake news. If randomisation is achieved through assignment to a linguistic community, then the coefficient λ_2 will capture the causal effect of fake news exposure on electoral support for populism. Also, it is worth stressing out that Z_i is a continuous indicator pointing at the (historically-adjusted) number of Italian-speaking citizens, and not a dummy. In this way, we can control for the presence of municipalities with mixed population, and allow for contamination across linguistic communities.

The usage of instrumented diff-in-diff is motivated by the different electoral patterns across the two linguistic groups. This is important because the SVP has consistently proven to be the most popular voting choice across the German-speaking population, our estimates would

suffer from upward bias when not controlling for previous elections through municipality and time fixed effects. The idea here is that holding pre-exposure levels in the outcome as fixed we can focus on how the fake news shock has affected changes in the outcome. The exposure effect will then be consistently estimated as long as this effect is stable and homogeneous (de Chaisemartin and D’Haultfoeuille, 2018).

For what concerns the choice of the instrument, linguistic groupings provide for an ideal source of exogenous allocation. The intuition is that native languages are assigned at birth, and they necessarily predate exposure to fake news or electoral outcomes. This is not true for other variables that may seem good candidates as instruments, such as access to broadband internet. While the presence of internet infrastructure certainly influences the exposure to online (mis)information, it is not necessarily independent from electoral outcomes. Since internet diffusion is mostly privately funded, it is likely correlated with certain socioeconomic characteristics, such as income, age and education. As explained by Poy and Schüller (2020), these characteristics are likely correlated with voting behaviour, creating a potential bias in the estimates of the effect of internet availability on electoral outcomes. For this reason, a number of works that study the impact of internet on voting behaviour treat broadband connectivity as an explanatory variable, rather than as an instrument (Campante et al., 2017; Falck et al., 2014; Larcinese and Miner, 2012).

The exogeneity of language grouping is based on the assumption that the share of Italian-speaking population is independent from the variation in populist preference within each municipality and from the other covariates included in the model. In other words, the linguistic groups should not be retroactively affected by the voting intentions in 2013 and 2018 or any other unobservable factor. This assumption might be difficult to defend when using population data at the time of the elections as the decision to live in a particular area by Italian/German speakers might not be random at all and could retroactively affect voting intentions through a self-selection mechanism. While it is easy to show that there is little time-variation in these groups, we prefer to avoid making this assumption and adopt a set of historically-adjusted instruments which are less likely to be reflect the endogeneity in these preferences.

More precisely, for each municipality we collect data on the proportion of Italian-speaking population in the years 1971, 1981, 1991, and 2001,¹⁴ and use them separately as instruments in alternative specifications. These historically-adjusted linguistic grouping instruments allows us to better capture the intention to treat than an instrument based on linguistic groups at the time of the elections, while holding total population as fixed. The only caveat that applies is that the share of Italian-speaking population is now time-invariant, but this is not an issue since we exploit its time-specific effect on exposure in 2018 alone (i.e., the interaction $Z_i \times P_t$). Since the 1971-adjusted instrument is the less likely to suffer from endogeneity, we adopt it as our preferred instrumental variable. For robustness, we will also show estimates with 1981, 1991, and 2001-adjusted instruments.

Turning to the relevance condition, the historical assignment to the Italian linguistic community should increase the voters’ likelihood of exposure to misinformation concerning the Italian election. We test this assumption in the first-stages (Eq. (5.2)). It is important to recall that our assumption is not that German-speakers are completely oblivious to fake news nor that they are unable to understand it, but simply that fake news has a lower probability of reaching them due to the different language-induced media bubbles.¹⁵ The main implication is

¹⁴ These historical proportions are still at the municipal level, see Appendix A for more details.

¹⁵ It should be noted that German-speaking voters could potentially use online tools to translate fake news written in Italian that appears on their social media feed. Nevertheless, in general German-speaking voters are less likely to

that our estimate is the Local Average Treatment Effect of fake news on populism, which might not necessarily correspond to the Average Treatment Effect of fake news. Our estimates are internally valid for the group of compliers only, and measure the effect of fake news on Italian-speaking municipalities that were more exposed to fake news against German-speaking municipalities that were less exposed. The results are then local in the sense that they can only be extended to non-compliers (so, always-takers and never-takers) under the assumption of homogeneity of treatment effects.¹⁶ A weak first stage would suggest that the proportion of non-compliers is particularly high, but as discussed later in Section 6 this does not apply to our results.

However, our model is unable to control for the possibility that German-speaking voters are exposed to fake news target to German-speaking population only. This can instead lead to measurement error if fake news were produced specifically for the German-speaking population in the region or if fake from other German-speaking countries directly affected voting decisions in Italy. While this is an important caveat that applies to our model, it is unlikely to affect our estimates for two main reasons.

First, fake news disseminators have no incentive to produce fake news on the Italian election in German (or in Ladin, Mocheno or Cimbrian). Following Allcott and Gentzkow (2017), producers of fake news are motivated by either profit – maximising advertising revenue or sell value of their Facebook page – or ideology – increasing electoral gains of their party. Given the small size of these non-Italian-speaking communities relatively to the population of Italian voters,¹⁷ the marginal return in terms of economic revenue and electoral gains of a piece of misinformation concerning Italian politics is very small. In the context of the Italian elections disseminators are hence better off producing fake news in Italian. Accordingly, to the best of our knowledge, there are no documented instances of fake news in German directly targeted to the German-speaking Italian population.

Second, producers of fake news do not attempt to build a long-run reputation and maximise short-run profits around the timing of an election. The probability that producers of fake news in German who are set in Germany and Austria “export” misinformation in German to Italy is therefore very small. This is further motivated by the fact that our fake news database is dominated by hoaxes targeted to specific Italian politicians, as described in Section 2, rather than to broader topics such as migration or the European Union.¹⁸

A relevant discussion concerns the exclusion restriction and its relationship with parallel trends. It could be argued that linguistic

receive online news in Italian in their feed than Italian-speaking voters. This is because their social media feed is more likely to report content in their native language rather than in a language they do not master. Another factor that reduces this likelihood is that fake news tend to target Italian-speaking politicians rather than broader topics that could be appealing to voters that do not speak Italian, such as migration or health issues, as shown in Fig. 1.

Although it is unlikely that German-speaking voters are exposed to fake news in Italian for these reasons, our empirical framework takes into account this possibility when estimating the effect of misinformation. Since we are estimating a local average treatment effect, German-speaking municipalities that are exposed to fake news fall in the category of ‘always-takers’, and hence do not contribute to the estimation of the treatment effect. Therefore, while fake news aimed at affecting Italian-speaking individuals may indirectly affect a small pool of German-speaking voters too, such effect does not influence our estimates.

¹⁶ For example, German-speaking municipalities with high fake news exposure are treated as “always-takers” and as such they contribute less to the estimation of the local average treatment effect, and the same goes for “never-takers” Italian-speaking municipalities with low exposure.

¹⁷ A very small minority of the Italian population speaks German (0.6% of the whole Italian population), Ladin (0.06%), Mocheno (0.002%), and Cimbrian (0.001%). Source: Astat (2020) and Ispat (2012).

¹⁸ Allcott and Gentzkow (2017) similarly found that fake news targeted specific politicians.

groupings could affect populist preference through channels other than exposure to misinformation. For this reason, a cross-sectional setting would violate the exclusion restriction, since it would assume that the instrument does not affect the levels of the outcome. In the two-periods setting of our specification, however, we study the effect of the instrument on *changes* in populist preferences (rather than on *levels*), which are unaffected by the instrument and therefore allow for the exclusion restriction to hold.¹⁹ This exclusion restriction is then intrinsically related with the assumption of parallel trends.

A potential criticism to this specification is that municipalities which have historically been Italian-speaking are more susceptible to populism than German-speaking municipalities, so much that differing trends in policy preferences across the years are correlated to ancestral cultural differences. We find this argument hard to defend.

On the demand side, populism is hardly a phenomenon exclusive to the Italian-speaking population. Recent studies, such as [Schaub and Morisi \(2019\)](#), have shown how German and Italian-speaking populations are equally susceptible to populist rhetoric. Our underlying assumption is that, after checking for different voting trends through the difference in differences setting, Italian-speaking and German-speaking Italians are equally susceptible to populism. We find difficult to assume otherwise without falling into cultural stereotyping or ignoring these peoples' recent shared history with authoritarianism and populism.

Differences between the linguistic groups could still be attributed to other time-varying socio-demographic factors. This is particularly relevant since municipal-level information on educational attainment and other confounders is unavailable,²⁰ leading to the omission of a potential source of variation in our model. However, as municipality fixed effects capture time-invariant heterogeneity, unobserved socio-demographic differences become problematic only if the 2013–2018 variation between linguistic groups is significant. This does not seem to be the case. Observed socio-demographic trends remain parallel, while language groups affiliations have also remained stable over the last decades (as discussed in the next section). It follows that the province is characterised by remarkable demographic stability.²¹ To further corroborate this statement, in the next section we will check for the amount of within variation captured by municipality fixed effects and the model fit after adding the year effect. If these factors are time invariant or changing at a constant rate, then the fixed effects alone should absorb most of the variation in the model, with the time effect absorbing much of the remaining variation.

Finally, a related issue concerns the fact that variations in the electoral outcome of a party could be related to a potential unobserved change in the supply of populism between linguistic groups rather than to fake news exposure. Our text-based indicator controls for this heterogeneity on both the supply- and demand-side of populism at national level. As shown in Section 4, on the one hand, our indicator captures changes in the supply of populism by measuring party-level changes in the populist rhetoric of Facebook posts. On the other hand, by weighting the populist rhetoric of each Facebook post by its engagement, it controls for variations in the change in demand for populism at a national level. This means that we can study variation in demand for

¹⁹ This is a reasoning that similarly applies to Bartik-like instruments in general ([Goldsmith-Pinkham et al., 2020](#)).

²⁰ As a matter of fact, most of this information, if released at a municipal level, is only available every ten years together with census data and as such is not particularly useful for our analysis.

²¹ Looking for example at educational achievement alone, Fig. AF.2 in Appendix F shows that tertiary educational attainment trends across the two provinces have proceeded in parallel during the last decade (2011–2019). This is rationalised by the presence of renowned universities in the capitals of both provinces, ensuring that individuals from both linguistic groups have access to tertiary education in their preferred language. As both provinces followed remarkably similar trends in educational attainment across the two periods, the municipality-year fixed effects will absorb educational effects.

populism at a local level, net of aggregate supply and demand factors which are allowed to vary between language groups.

Exclusion restrictions notoriously cannot be tested and our empirical framework is no exception. In our case, an informal test is to check for parallel trends in the outcome variable during pre-treatment years. Nevertheless, this strategy is not feasible for several reasons.

The earlier national election before 2013 dates back to 2008. Major populist parties which ran for elections in 2013 and 2018 either did not exist in 2008 (as M5S and Brothers of Italy) or could simply not be considered as populist (as LEGA, which was still called LEGA NORD). The seismic shock of modern populism in Italy cannot simply be replicated in a earlier time-frame. In addition, many of the parties running for election in 2008 barely had any social media presence: looking at the major parties running for election in 2008, the earliest appearance on Facebook was 27 April 2009 for Partito Democratico²² and 6 May 2010 for Silvio Berlusconi's (at the time leader of Popolo della Libertà),²³ meaning that it would not be possible for us to accurately derive the same Facebook based populist indicator (described in the previous section) for these other elections. Against this backdrop, even if we opt to use a binary indicator, it would be difficult to categorise pre-2013 parties as populist without a certain degree of discretionary judgement.

So far, our checks for the predictive power of municipal and time fixed effects, along the fact that all time-varying exogenous variables seem to evolve in parallel between the two regions reasonably point at the presence of parallel trends. Nonetheless, we feel it is important to spell out the limitations of our approach clearly.

6. Results

6.1. Main results

[Table 3](#) shows the results from the linear model of Eq. (5.1), using, as for all subsequent specifications, the total populist scores by municipality as the outcome variable. This has been computed by aggregating text analysis scores obtained through the assertive dictionary (for the results with the anti-establishment dictionary, see the next section), standardising these scores to the score of the LEGA party in 2018, and weighting them by the size of the electorate. Based on this standardisation, a unitary increase in the total populist score for a municipality corresponds to a vote to a party which can be qualified as approximately as populist as the LEGA, which scored as the most populist in 2018.

Column (1) presents a baseline model including exposure to fake news and the year of election dummy alone. The model greatly overestimates the impact of misinformation. When we add municipality fixed effects in column (2), the results are more closely aligned to our expectations. As expected, populist vote increases when the year dummy is equal to 1, as 2018 coincided with the electoral success of populist platforms. The coefficient of fake news exposure in 2018 is positive and significant at the 0.001 level, suggesting that populists gained 0.166 votes for each additional like to the observed disseminator pages. Fake news exposure yet explains only a residual amount of variance in the model, as the coefficient on the year of election suggests that the greatest component of growth in populist voting is to be attributed to other factors.

The increase of model fit from 0.52 to 0.99 after the introduction of year fixed effects, suggesting that most of the unobserved municipality characteristics were, indeed, time-invariant, as we hypothesised. This also suggests that the observed time-varying changes in populist preference captured by the year dummy are unrelated to municipal

²² Source: https://www.facebook.com/partitodemocratico/about_profile_transparency. Last accessed: 23 June 2022.

²³ Source: https://www.facebook.com/SilvioBerlusconi/about_profile_transparency. Last accessed: 23 June 2022.

Table 3
OLS estimates of the effect of misinformation on populist vote.

Variables	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Exposed to fake news × Year of election	0.848*** (0.108)	0.166*** (0.010)	0.166*** (0.009)	0.163*** (0.009)	0.124*** (0.008)
Year of election	−162.405 (108.402)	153.993*** (13.739)	148.801*** (15.042)	236.154*** (25.475)	158.539*** (20.852)
Broadband connections			0.016 (0.047)	0.019 (0.046)	−0.046* (0.025)
Electorate size					1.310*** (0.262)
Income per capita (natural log)				−1,039.756*** (283.242)	−675.253*** (205.742)
Observations	584	584	584	584	584
Adjusted R-squared	0.523	0.995	0.995	0.995	0.997
Within R-squared		0.833	0.833	0.841	0.896
Municipality FE	NO	YES	YES	YES	YES

Notes: OLS estimates for the effect of misinformation on populist vote. Populist scores computed using the ‘Assertive’ text bag. Standard errors robust to clustering by municipality in parentheses.

*p < .05; **p < .01; ***p < .001

characteristics and rather reflect a regional-wide (and nation-wide) populist shock, as they alone explain 0.83% of the within-variation in populism in the region.

Column (2) adds the number of landline low-latency connections to the covariates. As mentioned, several studies find a significant relationship between broadband connectivity and populist preference (Campanante et al., 2017; Schaub and Morisi, 2019). Our estimates are, however, only marginally affected, suggesting that broadband connections did not grow sufficiently between the two years to explain the variation in exposure to misinformation in 2018, nor the growth in populism. Estimates are again marginally affected by the introduction of income per capita in column (4) which, while significant and large in its effect, has little impact on the effect of fake news exposure.

Electorate size controls affect the coefficient pointing some non-linearities between electorate size and populist preference, with more populated areas having experienced the greatest growth in populism, and the correlation with income being lowered by a non-negligible degree.²⁴ As a result, the effect of exposure on voting remains statistically significant, but its effect is lowered to 0.124.

Our OLS results show a correlation between populist preference and exposure to misinformation. However, these estimates might still suffer from bias, as they give no information on the direction of causation. Indeed, we do not know whether the significance of the coefficient is actually showing the impact of fake news on voting or, on the contrary, that access to misinformation bubbles is linked to individual characteristics that may already determine a populist preference, either through self-selection or online recruitment. These endogeneity issues cannot be addressed by simple correlations, and motivate our instrumental variable design.

Results from our two-stage-least-squares model are shown in Table 4. For each specification the table presents the first stage (columns 1, 3, 5 and 7) and second stage estimates (columns 2, 4, 6, and 8). Here, the interactions between year and the historically-adjusted number of Italian-speaking voters by municipality are instrumented to predict the exposure to fake news in 2018. In each pair of column we use a different the historically-adjusted instrument, starting with 1971 and ending with 2001. The controls are the same we used in our final OLS specification in column (5) of Table 3.

Looking at the first stages, we find that the historical share of Italian-speakers is a strong predictor for fake news exposure, regardless of the reference year. On average, one additional Italian-speaking voter in a municipality increases the number of likes to a disseminator page by a factor of ~0.5. The lack of variance in the coefficients reminds

²⁴ We also tested for the interaction of electorate size and income, but the estimated coefficient was not significant at the 0.05 level.

us that these linguistic groups have remained stable over time, and that most endogeneity concerns in for the instrument can be kept at bay. Partial R-squared statistics also indicate that, in both cases, most of the variation unexplained by the control covariates is captured by the instrument. Most importantly, all instruments pass the F-test for excluded instruments (Bound et al., 1995), increasing the robustness of our estimates. These F-tests abundantly pass the conventional threshold of $F > 10$ and are also robust (with the only exception of 1981) to the more rigorous $F > 104.7$ threshold recently suggested by Lee et al. (2021).

Turning to our second stage estimates, we still find that a significant and positive effect of fake news exposure on populist voting. Results from column (2), which use the 1971-adjusted instrument, indicate that for each additional like to a disseminator page populist parties gained 0.145 votes. The estimated coefficient is unaffected from the use of the other historically-adjusted instruments, and is very close to the OLS estimate of 0.124, suggesting that fake news effectively succeeded in nudging voters towards more populist electoral picks.

6.2. Robustness checks

For robustness, we present alternative estimates using the populist score based on the anti-establishment text bag (described in Section 4) in Section C of the Appendix. The two tables replicate the baseline specifications, with Table A2 showing OLS estimates and Table A3 displaying IV estimates using language group as instrument. Our results and interpretations remain mostly unchanged.

In addition, Section C of the Appendix presents and discusses estimates of the direct effect of fake news on the electoral outcomes of major anti-establishment parties, leaving the populist score aside.

7. Conclusions

The influence of fake news on electoral results has, so far, escaped empirical assessment. With this study, we aim to fill this void and identify both the presence and the magnitude of the effect of fake news exposure on voting behaviour.

We find that linguistic membership in Trentino-Alto Adige/Südtirol significantly affects exposure to fake news, as Italian-speaking individuals are more likely to follow social media accounts disseminating fake news on Italian politics when compared to other linguistic minorities in the region. This relationship is robust to the inclusion of relevant controls such as income, population size, and broadband internet coverage. By exploiting the presence of language-biased exposure to misinformation in the context of the 2018 Italian elections, and controlling for time-invariant factors via municipality fixed effects, we identify the effect of fake news on populist voting at municipality level.

Table 4
2SLS estimates of the effect of misinformation on populist vote.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1st Stage 1971	2SLS 1971	1st Stage 1981	2SLS 1981	1st Stage 1991	2SLS 1991	1st Stage 2001	2SLS 2001
Historically-adjusted IT-speaking pop. in 2018	0.538*** (0.046)		0.524*** (0.052)		0.533*** (0.046)		0.532*** (0.046)	
Exposed to fake news × Year of election		0.145*** (0.015)		0.145*** (0.015)		0.146*** (0.015)		0.146*** (0.015)
Year of election	−434.906*** (74.160)	169.063*** (20.485)	−417.727*** (77.251)	169.054*** (20.454)	−435.354*** (75.235)	169.318*** (20.480)	−433.860*** (75.477)	169.232*** (20.476)
Broadband connections	0.113 (0.086)	−0.049** (0.025)	0.128 (0.089)	−0.049** (0.025)	0.112 (0.089)	−0.049** (0.025)	0.109 (0.089)	−0.049** (0.025)
Electorate size	−4.149*** (1.320)	1.033*** (0.238)	−3.903*** (1.375)	1.033*** (0.236)	−3.789*** (1.328)	1.026*** (0.237)	−3.770*** (1.338)	1.029*** (0.237)
Income per capita (natural log)	2,499.944*** (714.212)	−683.468*** (207.589)	2,369.199*** (707.900)	−683.461*** (207.559)	2,521.600*** (720.071)	−683.667*** (207.678)	2,505.227*** (717.984)	−683.600*** (207.655)
Observations	584	584	584	584	584	584	584	584
Adjusted R-squared	0.948	0.890	0.945	0.890	0.946	0.890	0.946	0.890
Municipality FE	YES	YES	YES	YES	YES	YES	YES	YES
F-Test	135.3		102.4		134.7		131.3	
Partial R-Squared	0.906		0.900		0.902		0.902	

Notes: IV estimates for the effect of misinformation on populist vote. Populist scores computed using the 'Assertive' text bag. Standard errors robust to clustering by municipality in parentheses.

*p < .05; **p < .01; ***p < .001

We show that misinformation had a positive effect on the electoral support for populist parties. An additional like to a Facebook page that disseminated fake news led, on average, to an electoral gain of 0.145 votes for populist parties. At the same time, our results indicate that most of the growth in populist vote cannot be explained exposure to fake news nor by pre-existing voting preferences. This does not mean that misinformation is less problematic, but only that fake news cannot be accounted as the fundamental cause of the growth of populism. Moreover, the spread of misinformation can still have important negative externalities that go beyond the impact on voting behaviour, as in the case, for example, of public health crises.

Our final notes address the validity of our results and suggest future pathways for research. Focusing on the region of Trentino-Alto Adige/Südtirol imposes a constraint on the external validity of our results, as the relationship between misinformation and voting might differ in other regional and national contexts. However, a similar methodology could be applied to other contexts where language groups have proven good predictors for access to fake news. Survey data might also shed more light on individual preferences and social media behaviours.

Our results are also robust to a specific definition of populism based on rhetoric. Our methodology could be extended using different text-bags that touch on different aspects of populism, such as its ethno-nationalistic features. Further work is certainly needed to address competition between party platforms sharing contiguous filter bubbles and investigate whether misinformation favours certain versions of populism over other ones.

More broadly, the new indicator of populist rhetoric we introduce is useful to study populism as a phenomenon that eschews the political dimensions of left and right. In its current form, our dictionaries are applicable to other types of political texts in Italian and German, and could also be translated in other languages. The greatest benefit of using a continuous indicator for the populist content of parties is that it allows the researcher to study their electoral success net of their idiosyncratic supply of populism.

Our estimates rest on the assumption that linguistic groups, while allowed to affect the levels of electoral success of populism, do not affect the change in populism if not through filtering access to specific forms of information. In the case of these 2018 elections, fake news constituted this kind of information. This exclusion restriction cannot be tested and our findings call for further research to study whether linguistic groups can trigger heterogeneous responses in populism.

Finally, future research could focus on the relationship between misinformation and echo chambers. Given the important role of prior

preferences and individual characteristics to determine access to filter bubble, researchers and policy makers could explore which socio-economic characteristics are associated with such users. Since, as we find, fake news has only a marginal impact on voting behaviour, then the study of information streams in social media is relevant to enhance our understanding of the link between misinformation and populism.

CRediT authorship contribution statement

Michele Cantarella: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing – original draft, Writing – review & editing. **Nicolò Fraccaroli:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing – original draft, Writing – review & editing. **Roberto Volpe:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.respol.2022.104628>.

References

- Abel, Andrea, Vettori, Chiara, Forer, Doris, 2012. Learning the neighbour's language: The many challenges in achieving a real multilingual society. The case of second language acquisition in the minority-majority context of South Tyrol. *Eur. Yearb. Minor. Issues Online* 9 (1), 271–303.
- Albertazzi, Daniele, McDonnell, Duncan, 2008. Twenty-First Century Populism. *The Spectre of Western European Democracy*. Palgrave Macmillan.
- Allcott, Hunt, Gentzkow, Matthew, 2017. Social media and fake news in the 2016 election. *J. Econ. Perspect.* 31 (2), 211–236.
- Astat, 2011. Censimento della popolazione 2011: Determinazione della consistenza dei tre gruppi linguistici della Provincia Autonoma di Bolzano-Alto Adige. Astat Informazioni.

- Astat, 2020. South Tyrol in Figures. Report, Astat.
- Autorità per le Garanzie nelle Comunicazioni, 2018. News vs Fake nel sistema dell'informazione, Interim report nell'ambito dell'indagine conoscitiva di cui alla delibera n. 309/16/CONS. Interim Report, Agcom.
- Baggio, S., 2020. Il Trentino, modello di terra multilingue e multiculturale. In: Davide, Astori (Ed.), Dieci anni (2009–2019) di Premio “G. Canuto” per tesi di laurea in Interlinguistica ed Esperantologia: Con cinque interventi linguistici a Parma. Athenaeum Edizioni Universitarie.
- Barrera, Oscar, Guriev, Sergei, Henry, Emeric, Zhuravskaya, Ekaterina, 2020. Facts, alternative facts, and fact checking in times of post-truth politics. *J. Public Econ.* 182, 104123.
- Bischof, Daniel, Senninger, Roman, 2018. Simple politics for the people? Complexity in campaign messages and political knowledge. *Eur. J. Political Res.* 57 (2), 473–495.
- Boelitz, Otto, 1930. Das Grenz- und Auslandsdeutschum, seine Geschichte und seine Bedeutung. Oldenbourg.
- Bound, John, Jaeger, David A., Baker, Regina M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *J. Amer. Statist. Assoc.* 90 (430), 443–450.
- Boutyline, Andrei, Willer, Robb, 2016. The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political Psychol.* 38 (3), 551–569.
- Caiani, Manuela, Graziani, Paolo R., 2016. Varieties of populism: Insights from the Italian case. *Ital. Political Sci. Rev.* 46 (2), 243–267.
- Campante, Filipe, Durante, Ruben, Sobbrío, Francesco, 2017. Politics 2.0: The multifaceted effect of broadband internet on political participation. *J. Eur. Econom. Assoc.* 16 (4), 1094–1136.
- de Chaisemartin, Clément, D'Haultfoeuille, Xavier, 2018. Fuzzy differences-in-differences. *Rev. Econom. Stud.* 85 (2), 999–1028.
- Chiang, Chun-Fang, Knight, Brian, 2011. Media bias and influence: Evidence from newspaper endorsements. *Rev. Econom. Stud.* 78 (3), 795–820.
- D'Alimonte, Roberto, 2019. How the populists won in Italy. *J. Democr.* 30 (1), 114–127.
- Decadri, Silvia, Boussalis, Constantine, 2019. Populism, party membership, and language complexity in the Italian chamber of deputies. *J. Elections Public Opin. Parties* 1–20.
- DellaVigna, Stefano, Kaplan, Ethan, 2007. The Fox News effect: Media bias and voting. *Q. J. Econ.* 122 (3), 1187–1234.
- Drago, Francesco, Nannicini, Tommaso, Sobbrío, Francesco, 2014. Meet the press: How voters and politicians respond to newspaper entry and exit. *Am. Econ. J. Appl. Econ.* 6 (3), 159–188.
- Durante, Ruben, Pinotti, Paolo, Tesei, Andrea, 2019. The political legacy of entertainment TV. *Amer. Econ. Rev.* 109 (7), 2497–2530.
- Ebner, Chiara Vettori, 2016. The long way to bilingualism: The peculiar case of multilingual South Tyrol. *Int. J. 21st Century Educ.* 3 (2), 25.
- Ernst, Nicole, Engesser, Sven, Büchel, Florin, Blassnig, Sina, Esser, Frank, 2017. Extreme parties and populism: An analysis of Facebook and Twitter across six countries. *Inf. Commun. Soc.* 20 (9), 1347–1364.
- Falck, Oliver, Gold, Robert, Heblich, Stephan, 2014. E-lections: Voting behavior and the internet. *Amer. Econ. Rev.* 104 (7), 2238–2265.
- Gavazza, Alessandro, Nardotto, Mattia, Valletti, Tommaso, 2018. Internet and politics: Evidence from U.K. local elections and local government policies. *Rev. Econom. Stud.* 86 (5), 2092–2135.
- Gerber, Alan S., Gimpel, James G., Green, Donald P., Shaw, Daron R., 2011. How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment. *Am. Political Sci. Rev.* 105 (1), 135–150.
- Gerber, Alan S., Karlan, Dean, Bergan, Daniel, 2009. Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions. *Am. Econ. J. Appl. Econ.* 1 (2), 35–52.
- Giglietto, Fabio, Iannelli, Laura, Rossi, Luca, Valeriani, Augusto, Righetti, Nicola, Carabini, Francesca, Marino, Giada, Usai, Stefano, Zurovac, Elisabetta, 2018. Mapping Italian news media political coverage in the lead-up of 2018 general election. *SSRN Electron. J.*
- Goldsmith-Pinkham, Paul, Sorkin, Isaac, Swift, Henry, 2020. Bartik instruments: What, when, why, and how. *Amer. Econ. Rev.* 110 (8), 2586–2624.
- Guess, Andrew, Nyhan, Brendan, Reifler, Jason, 2018. Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *Eur. Res. Council.*
- Guiso, Luigi, Herrera, Helios, Morelli, Massimo, Sonno, Tommaso, 2017. Populism: Demand and Supply. Center for Economic Policy Research Discussion Paper, 11871.
- Gunther, Richard, Beck, Paul A., Nisbet, Erik C., 2019. “Fake news” and the defection of 2012 Obama voters in the 2016 presidential election. *Elect. Stud.* 61, 102030.
- Guriev, Sergei, 2018. Economic drivers of populism. In: *AEA Papers and Proceedings*, Vol. 108, pp. 200–203.
- Guriev, Sergei, Papaioannou, Elias, 2020. The political economy of populism. *J. Econ. Lit.* (forthcoming).
- Hopkin, Jonathan, 2020. Anti-System Politics: The Crisis of Market Liberalism in Rich Democracies. Oxford University Press.
- Ispat, 2012. 15esimo Censimento della popolazione e delle abitazioni. Rilevazione sulla consistenza e la dislocazione territoriale degli appartenenti alle popolazioni di lingua ladina, mòchena e cimbra. Report, Ispat.
- Jaber, Ahmed, 2013. Individual Campaign Contributions, Electoral Participation, and Broadband Network Availability. SSRN, Available at SSRN: URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2201960.
- Kaltwasser, Cristóbal Rovira, Taggart, Paul A., Espejo, Paulina Ochoa, Ostiguy, Pierre, 2017. *The Oxford Handbook of Populism*. Oxford University Press.
- Kumar, Akshi, Sebastian, Teeja Mary, 2012. Sentiment analysis on Twitter. *Int. J. Comput. Sci. Issues* 9 (4), 372–378.
- Larcinese, Valentino, Miner, Luke, 2012. The Political Impact of the Internet on US Presidential Elections. Working paper, London School of Economics.
- Lee, David S, McCrary, Justin, Moreira, Marcelo J, Porter, Jack R, 2021. Valid T-Ratio Inference for IV. Working Paper Series, 29124, National Bureau of Economic Research.
- Lelkes, Yphtach, Sood, Gaurav, Iyengar, Shanto, 2017. The hostile audience: The effect of access to broadband internet on partisan affect. *Am. J. Political Sci.* 61 (1), 5–20.
- Liberini, Federica, Redoano, Michela, Russo, Antonio, Cuevas, Angel, Cuevas, Ruben, 2020. Politics in the Facebook Era Evidence from the 2016 US Presidential Elections. CESifo Working Paper 8235, CESifo.
- Miner, Luke, 2015. The unintended consequences of internet diffusion: Evidence from Malaysia. *J. Public Econ.* 132, 66–78.
- Mudde, Cas, 2004. The Populist Zeitgeist. *Gov. Oppos.* 39 (4), 541–563.
- Munger, Kevin, Egan, Patrick J., Nagler, Jonathan, Ronen, Jonathan, Tucker, Joshua, 2020. Political knowledge and misinformation in the era of social media: Evidence from the 2015 UK election. *Br. J. Political Sci.* 1–21.
- Norris, Pippa, 2019. Varieties of populist parties. *Philos. Soc. Crit.* 45 (9–10), 981–1012.
- Norris, Pippa, Inglehart, Ronald, 2019. *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. Cambridge University Press.
- Pauwels, Teun, 2011. Measuring populism: A quantitative text analysis of party literature in Belgium. *J. Elections Public Opin. Parties* 21 (1), 97–119.
- Peisakhin, Leonid, Rozenas, Arturas, 2018. Electoral effects of biased media: Russian television in Ukraine. *Am. J. Political Sci.* 62 (3), 535–550.
- Polk, Jonathan, Rovny, Jan, Bakker, Ryan, Edwards, Erica, Hooghe, Liesbet, Jolly, Seth, Koedam, Jelle, Kostelka, Filip, Marks, Gary, Schumacher, Gijis, et al., 2017. Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data. *Res. Politics* 4 (1), 2053168016686915.
- Poy, Samuele, Schüller, Simone, 2020. Internet and voting in the social media era: Evidence from a local broadband policy. *Res. Policy* 49 (1), 103861.
- Rooduijn, Matthijs, 2014. The nucleus of populism: In search of the lowest common denominator. *Gov. Oppos.* 49 (4), 572–598.
- Rooduijn, Matthijs, Pauwels, Teun, 2011. Measuring populism: Comparing two methods of content analysis. *West Eur. Politics* 34 (6), 1272–1283.
- Schaub, Max, Morisi, Davide, 2019. Voter mobilisation in the echo chamber: Broadband internet and the rise of populism in Europe. *Eur. J. Political Res.*
- Sunstein, Cass R., 2018. *#Republic*. Princeton University Press.
- Swire, Briony, Berinsky, Adam J., Lewandowsky, Stephan, Ecker, Ullrich K.H., 2017. Processing political misinformation: Comprehending the Trump phenomenon. *R. Soc. Open Sci.* 4 (3), 160802.
- Taggart, Paul A., 2000. *Populism*. Open University Press.
- Zimmermann, Fabian, Kohring, Matthias, 2020. Mistrust, disinforming news, and vote choice: A panel survey on the origins and consequences of believing disinformation in the 2017 German Parliamentary Election. *Political Commun.* 37 (2), 215–237.