CEIS TOR VERGATA Research Papers

Università di Roma Tor Vergata

ISSN 2610-931X

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 18, Issue 6, No. 493 – June 2020

Does Fake News Affect Voting Behaviour?

Michele Cantarella, Nicolò Fraccaroli and Roberto Volpe

Does fake news affect voting behaviour?*

Michele Cantarella[1], Nicolò Fraccaroli[2], and Roberto Volpe[3]

[1]University of Modena and Reggio Emilia and University of Helsinki
[2]University of Rome Tor Vergata
[3]LUISS Guido Carli

14th May 2020

Abstract

We study the impact of fake news on votes for populist parties in the Italian elections of 2018. Our empirical strategy exploits the presence of Italian- 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. We find that exposure to fake news is positively correlated with vote for populist parties, but that less than half of this correlation is causal. Our findings support the view that exposure to fake news (i) favours populist parties, but also that (ii) it is positively correlated with prior support for populist parties, suggesting a self-selection mechanism.

Keywords: *Fake News, Political Economy, Electoral Outcomes, Populism*
JEL codes: C26, D72, P16.

*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 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' and to the 13th CESifo Workshop on Political Economy, as well as to participants to the seminars at the University of Rome Tor Vergata, the IRVAPP of the Bruno Kessler Foundation in Trento, and the University of Modena and Reggio Emilia for their comments. All remaining errors are ours.

1 Introduction

Over the last years, elections in Western democracies have been characterised by the proliferation of online fake news – intentionally fabricated misinformation with politically-charged content – and the electoral success of populist candidates. Fake news has proliferated in the months preceding the 2016 US Presidential election which resulted in the victory of Donald Trump, whose communication style and electoral campaign have been consistently characterised by the use of a populist rhetoric (Norris and Inglehart, 2019). Allcott and Gentzkow (2017) estimate that the average adult American saw and remembered 1.14 hoaxes, and that those in favour of Trump were shared 30 million times on Facebook against those in favour of Clinton, his opponent, which were shared 8 million times.

Misinformation circulated during the campaign for the 2017 French Presidential election too. In this case, misinformation took again an unambiguous stance in favour of the populist candidate, Marine Le Pen, who for the first time reached the runoff.[1] Hoaxes are also believed to have supported the electoral growth of the extremist party Alternative Für Deutschland in the 2017 German election (Zimmermann and Kohring, 2020), when it turned from being unrepresented in Parliament to becoming the third most voted party in the country. Fake news also spread before the 2018 Italian election, which resulted in the first populist majority in Western Europe (D'Alimonte, 2019). In this context, fact-checkers reported that in the three months before the election the second most shared news online was false and directed against the incumbent party.[2]

As a result, fake news has been held responsible for the electoral growth of populism.[3] Concerned by the influence of disinformation on electoral behaviour, public institutions and traditional media financed a number of initiatives to counter the diffusion of misinformation.[4] Yet, empirical evidence on the positive effect of fake news on the electoral support for populism and its magnitude is scant.

The seminal work by Allcott and Gentzkow (2017) highlight a strong positive correlation

[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). The article is based on data from the fact-checking website Buzzsumo.

[3] See Parkinson, H. J. (2016); Click and elect: how fake news helped Donald Trump win a real election; The Guardian; available at: https://www.theguardian.com/commentisfree/2016/nov/14/fake-news-donald-trump-election-alt-right-social-media-tech-companies; last accessed: 15 February 2019

[4] The European Parliament included 'Fighting Disinformation' as one of its priorities and produced studies and informative material to counter their diffusion: https://www.europarl.europa.eu/news/en/headlines/priorities/disinformation (last accessed: 30 April 2020). In view of the 2019 European elections, the European Commission launched an 'Action Plan against Disinformation' and an 'Election Package' to fight disinformation during the campaign. In 2016 the German Bundestag (Parliament) published a report on the legal situations of fake news and a reform proposal: https://www.bundestag.de/resource/blob/494418/4321d229204080dce488ebd0356b5db2/wd-10-067-16-pdf-data.pdf

between fake news and support for populism in the 2016 US Presidential election. Based on data from web-browsing, fact-checking and an online survey, they show that online fake news was heavily biased in favour of populist candidate Trump. Other studies based on surveys reach similar conclusions. Gunther et al. (2019) undertook a post-election online survey on 585 respondents who voted for Obama in 2012 and showed that those that shifted from Obama to Trump in 2016 were more prone to believe fake news. Similarly, using survey data on German voters, Zimmermann and Kohring (2020) find that, among the supporters of the incumbent CDU/CSU party, those who believe fake news were more likely to refrain their preferences in favour of the populist party Alternative Für Deutschland.

However, these works do not account for the selection bias linked to fake news exposure. Voters of populist parties could self-select into misinformation bubbles and consume fake news that confirms their prior preferences for populist platforms.

Indeed, Allcott and Gentzkow (2017) also observe that voters were 15% more likely to believe news in favour of their preferred candidate and that this inference was stronger for individuals with ideologically segregated social media networks. Similarly, results from Swire et al. (2017) show that supporters of Trump believed more information attributed to Trump than information without attribution. Guess et al. (2018) show that Conservatives were more likely to share articles from fake news domains than liberal or moderate voters. Liberini et al. (2020) find that online political advertisements on Facebook targeted specifically male individuals without college education who already supported the Republican party to vote for Donald Trump in 2016. The self-selection bias finds further support in the psychology literature: Boutyline and Willer (2016) show that more conservative and extreme users on Twitter are more inclined to associate with users with similar views to theirs.

These findings highlight that the correlation between fake news and populist vote is not sufficient. In order to understand whether fake news actually affects electoral support for populists, it is necessary to identify a direction of causation.

This is addressed in Barrera Rodriguez et al. (2017). Based on a randomised online experiment on voting-age French inhabitants before the 2017 Presidential elections, this study finds that individuals exposed to false statements from populist candidate Marine Le Pen moved their voting intentions in her favour. However, as in most random experiments, they only estimate the impact of misinformation on voting intentions and not on actual voting. On the one hand, voting intentions can be relevant predictors of actual voting, especially if expressed closely to the occurrence of an election. On the other hand, previous works show that the effects of voting intentions are larger than actual voting (Gerber et al., 2009; Gerber et al., 2011; Chiang and Knight, 2011), which is a limitation also acknowledged in Barrera Rodriguez et al. (2017).

The aim of this paper is to fill these gaps and investigate whether fake news has a tangible impact on populist voting. To this aim, we analyse the case of the 2018 Italian election

in the region of Trentino Alto-Adige South Tyrol, which offers a unique quasi-experimental setting to overcome the above-mentioned limitations. Located on the border with Austria, this region hosts two communities of Italian voters: one that is German-speaking and another that is Italian-speaking. Since only a small portion of the local population is proficient in both languages (Ebner, 2016; Abel et al., 2012), the membership to each linguistic group serves as an ideal exogenous factor that influences the exposure of voters to fake news during the electoral campaign. Our basic assumption is that, in the months preceding the elections, fake news on the Italian election were in Italian, and not in German. The reason, which is motivated by the theoretical model of Allcott and Gentzkow (2017), is that fake news disseminators have no incentive to target a small pool of voters such as the German-speaking minority, since it represents a very small share of the electorate. Under this assumption, only the Italian-speaking group would have been exposed to fake news before the elections, while German-speaking voters were left unaffected, serving as an ideal control group.

We measure exposure to fake news by computing, for each municipality, the number of likes to Facebook pages that disseminated fake news in the months before the elections. Using the share of Italian-speaking population as an instrumental variable, 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. Using a difference-in-difference approach we compare this outcome with the one of the 2013 Italian election. This approach allows us to compare the voting trends of the two linguistic communities and hence to filter out community-specific political preferences, such as the preference in the German-speaking community for German-speaking parties (e.g. the Südtiroler Volkspartei).

We exploit the multi-party structure of the Italian political system to look at the effect of fake news on populist voting on a continuous scale. To this aim, and informed by the literature on populism, we construct a text-based indicator of populist rhetoric. We apply this measure to the population of Facebook posts of the parties and their leaders that were published three months before the 2013 and 2018 election. In this way we obtain a continuous indicator of populist rhetoric that characterise each party with different intensity. This allows us to have more precise estimates of populism than a binary classification of populist and non-populist candidates, which is generally used in other works, would allow.

We find that, while fake news has a positive effect on populist voting, this causal effect explains less than a half of the correlation between the two variables. This result suggests that, on the one hand, misinformation favoured populist candidates in the 2018 Italian election. On the other hand, it also shows that fake news mainly reached individuals that would have voted for populist parties anyway. In particular, we find that increasing the like to a Facebook page that disseminates fake news by one standard deviation results in a 0.37 standard deviations increase in the electoral support for populists. However, the causal effect of fake news amounts to a smaller increase of 0.19 standard deviations.

The main contribution of this work is that it is the first to estimate the impact of fake news on actual electoral results. Its quasi-experimental setting allows us to filter out the selection bias of fake news exposure and estimate its impact on actual voting. This is relevant also in light of previous works that focused on the impact of misinformation on voting intentions rather than actual voting.

While the main contribution of this work is that it is the first to estimate the impact of fake news on actual electoral results, our analysis improves upon the existing literature in several ways. First of all, it is the first to provide evidence on the 2018 Italian election. Italy is an interesting case study since it is the first Western European country where populist parties reached a majority (D'Alimonte, 2019). In addition, while in the other cases populism was circumscribed to a single party or candidate (Trump, Le Pen and her party, the AfD), in Italy the main populist parties are at least two. This enables to reduce party-specific confounding factors when studying the impact of misinformation on populism.

Our results bring new evidence on the influence of media 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) and television (DellaVigna and Kaplan, 2007; Gerber et al., 2011), affect voting behaviour. Previous works highlighted how media contributed to 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 TV made these voters more vulnerable to populist rhetoric. In line with this, they show that those municipalities exposed to Berlusconi's TV shifted their support to the Five Star Movement (M5s, hereafter) in 2013, a party ideologically distant from Forza Italia but that shares the same populist rhetoric. In line with this, both the Five Star Movement and Forza Italia score relatively high in our populist indicator. We contribute to this literature by studying the impact of non-traditional information sources on voting.

Moreover, we contribute to the literature on the impact of the internet on electoral politics. Findings on the effect of internet on voter turnout are mixed and highlight a positive (Jaber, 2013; Larcinese and Miner, 2012), negative (Falck et al., 2014; Gavazza et al., 2018) or insignificant impact (Miner, 2015). Campante et al. (2017) study the Italian elections from 1996 to 2013 and show that while internet had a negative effect on turnout until 2008, it was positively associated with online political participation. More interestingly for the purpose of our study, they show that in 2013 broadband internet favoured the electoral growth of the Five Star Movement, as it mobilised through the internet those voters disenchanted with establishment parties. This relationship holds when focusing on the case of the Province of Trento, which is one of the two provinces of the Trentino and South Tyrol region: Poy and Schüller (2020) show that broadband connectivity has a positive effect on voter turnout. Our findings add to this literature, showing that fake news contributed, but only marginally, to

this link by mobilising populist voters.

In exploiting the unique presence of two linguistic communities in a specific region, the effects we identify are naturally limited in their generalisability. However, the results of this paper are key to understand how individuals react to political misinformation beyond the context of a survey or an experiment. This contribution is therefore relevant for the action of public institutions in contrasting the spread of fake news.

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 journalistic[5] and institutional[6] 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 channel[7] reports that among the top 100 articles in Italian for social media engagement, five were hoaxes, while another ten were classified as 'highly problematic'.[8] While few in number, these hoaxes had a significant exposure. 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.

While 'purely false' news, as the latter example, has a predominantly 'anti-establishment' (and, by extension, 'anti-incumbent') character, many of the 'highly problematic' (i.e. misleading) news in the sample focuses on other issues, such as immigration.

[5] 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

[6] See Autorità per le Garanzie nelle Comunicazioni (2018)

[7] 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

[8] Articles defined as 'highly problematic' reported real events but out of context or omitting relevant details, providing a significantly distorted view of the real fact.

Similarly to the fake news on Sicily, non-traditional information sources have an anti-establishment bias. The report by Giglietto et al. (2018), which provides a detailed classification of news sources in the lead-up to 2018 general election based on partisanship of their news content, 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 Avaaz[9] provides evidence in support of this link, at least as far as Facebook is concerned.[10]

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.[11] 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.[12]

Figure 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 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

[9]La Repubblica (2019); Facebook chiude 23 pagine italiane con 2.4 milioni di follower: diffondevano fake news e parole d'odio; available at: https://www.repubblica.it/tecnologia/social-network/2019/05/12/news/facebook_chiude_23_pagine_italiane_con_2_4_milioni_di_follwers_diffondevano_fake_news_e_parole_d_odio-226098817/; last accessed: 13 May 2019

[10]The investigation uncovers an extensive network of Facebook pages and fake accounts that, in blatant violation of the platform's terms of use, pushed misinformation in support of Lega and M5s, as well as other anti-establishment and fringe causes. As a result, in May 2019 Facebook took down 23 of the pages reported by Avaaz. While some of these pages explicitly portrayed themselves as unofficial supporters of the two parties, others used a more subtle approach. The most popular page in this subset, 'I Valori Della Vita' (1.5m followers), ostensibly a lifestyle page, was actually part of a bigger network, sharing in a coordinated manner content from a right-wing, pro-Lega news site.

[11]The website was scraped in November 2019.

[12]The codes and data used in Figure 1 (including all text bags) are available in the online data archive.

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 are 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.

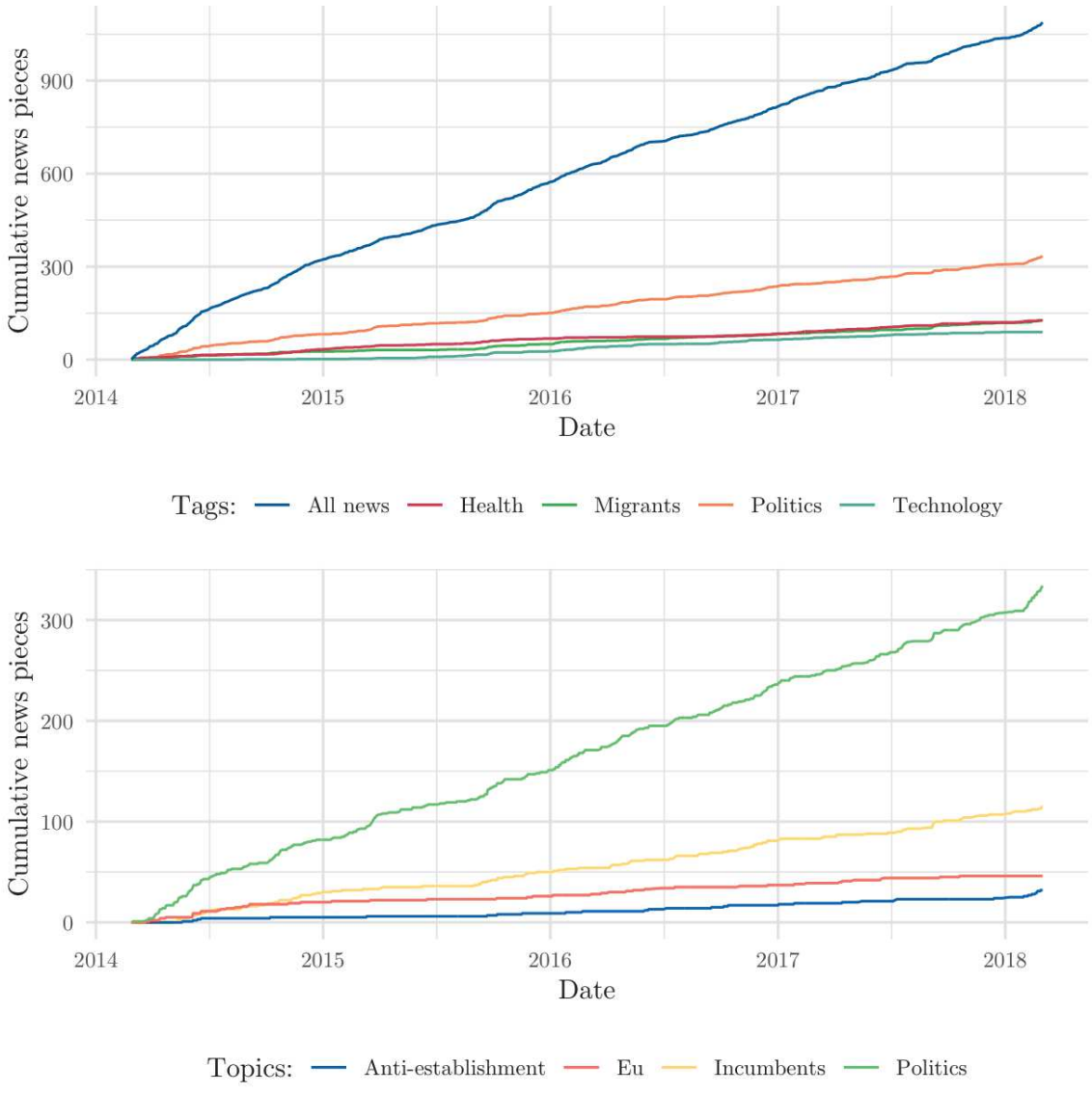

Figure 1: cumulative growth of fake news pieces reported by debunking websites from between 2013 and 2018. Frequency of fake news stories by tags and topic.

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 six among all headlines reported by Butac.it could be said to have an anti-M5s or anti-Lega slant, while two were directed against the neo-fascist parties Casapound and Forza Nuova.[13]

Moreover, the descriptive evidence presented so far shows that misinformation in support of anti-establishment parties, and in particular of the Lega and the M5s, was spread much more successfully than that of their pro-establishment counterparts *in terms of social media exposure* 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. The goal of our study is to assess whether this social media success causally translated to shifts in voting for populist parties.

2.2 Trentino and South Tyrol: political and sociolinguistic background

Trentino Alto-Adige/Südtirol is one of the five Italian regions that, since the end of World War 2, have enjoyed special legislative, administrative and financial autonomy compared to other regions. While the region forms a single electoral district for the Italian election, it is subdivided into two autonomous provinces (*province autonome*): Trento and Bolzano – Alto-Adige (in Italian) or Bozen – Südtirol (in German; 'South Tyrol' henceforth).

The two provinces have very similar characteristics. They both 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, they share many similarities. Trentino, with 540,000 inhabitants (Istat, 2018) 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,[14] scattered across hundreds of very small municipalities. Most crucially, they both enjoy a large degree of self-government compared to other Italian local authorities, with the *provincia autonoma* having significant legislative, fiscal and budgetary autonomy.[15]

[13]The four headlines with an anti-M5s stance are available at these links: https://www.butac.it/di-maio-emiliano-e-matera/; https://www.butac.it/sindaco-grillino-attacca-zanardi/; https://www.butac.it/sahaja-yoga-e-la-candidata-del-m5s/; https://www.butac.it/piccole-perle-di-facebbok-i-soldi-del-m5s/. The other 2 headlines concerning the Lega are here available: https://www.butac.it/salvini-i-giovani-padani-e-quelli-del-mezzogiorno/; https://www.butac.it/democratica-salvini-calabresi/. Last accessed: 29 May 2019

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

[15]For an overview of the competencies of the *province autonome*: Trentino: Consiglio della provincia

Moreover, due to their strong historical ties, since 1998 the two provinces form the Tyrol-South Tyrol-Trentino Euroregion together with the Austrian Federal State of Tyrol.

While similar in many regards, the two provinces are characterised by a strong long-standing linguistic divide. 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). These linguistic differences persist today. While the official languages of the region are actually five,[16] Italian and German are by far the most represented ones. The average municipality in the region has 60% of its inhabitants who speak Italian, 33.4% German, and 5% Ladin, while the remaining 1% speaks another language.

As in the past, today linguistic groups are not homogeneously distributed in the region. The German-speaking community is mostly concentrated in the North: 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 Ladin-speaking communities in the North-Eastern side. While both Italian and German are compulsory subjects for members of both language groups from age six, available data shows that effective bilingualism is not widespread. A number of studies, cited in 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 level in the Common European Framework of Reference for Languages).

This linguistic divide provides for an ideal exogenous allocation into two separate online media bubbles. This difference indeed 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,[17] 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

autonoma di Trento; "Le competenze legislative secondo lo statuto"; consiglio.provincia.tn.it; Last accessed: 25/05/2019; Available at: https://www.consiglio.provincia.tn.it/istituzione/l-autonomia/il-regime-delle-competenze-legislative-e-amministrative/Pages/il-quadro-delle-competenze-legislative-secondo-lo-statuto.aspx. South Tyrol: Amministrazione Provincia Bolzano; "Competenze e finanziamento dell'autonomia"; provincia.bz.it; Last accessed: 25/05/2019; Available at: http://www.provincia.bz.it/politicadirittorelazioni-estere/autonomia/competenze-finanziamento-autonomia.asp

[16]The five languages are Italian, German, Ladin, Cimbrian and Mòcheno. The last two languages are protected by the statute of the province of Trentino since the Italian constitutional reform of 2001. For more details see the webpage of the province of Trentino on the topic: http://www.minoranzelinguistiche.provincia.tn.it/normativa/ (last accessed 9 May 2020).

[17]Accertamenti Diffusione Stampa (ADS) (2016), *Dati medi annuali territoriali per testata: diffusione cartacea Italia*, available at: http://www.adsnotizie.it/_grafici.asp? (Last accessed August 2018)

different media bubble than the Italian-speaking ones.

The link between the linguistic and media divide in the region is supported by descriptive evidence on the diffusion of online fake news. Figure 2 displays two maps of the Trentino-Alto Adige/Südtirol region. The first map shows the share of Italian-speaking voters in each municipality.[18] The second one shows the logarithmic change in exposure to fake news disseminators from 2013 to 2018 across municipalities in the region. In particular, this second indicator captures the number of likes for each municipality to Facebook pages that disseminated fake news in 2018 (further details on the methodology used to measure this exposure are available in section 3). While large urban centres (such as Trento and Bolzano) are inevitably more exposed to fake news, the map clearly shows that across smaller municipalities, those that are Italian-speaking are more likely to be exposed to misinformation.

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. In South Tyrol, conversely, since the birth of the Italian Republic (1946), the party system has been dominated by the South Tyrolean People's Party (Südtiroler Volkspartei, 'SVP'). A catch-all party featuring both conservative and social-democratic wings, the SVP's 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, scoring landslides in several of the purely German-speaking municipalities. A remarkable exception is the Italian-majority city of Bolzano/Bozen, where the SVP has rarely exceeded 25% of the votes. 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 disenchanted voters have dropped their support for the SVP, while its supremacy has been challenged by several localist formations openly advocating for secession or reunification with Austria. In 2008 and 2013, elections that were fought with a proportional-based (PR) system, *Die Freiheitlichen* (DF), a radical-right separatist party, received 9.5% and 15% of the vote, respectively.

The DF forfeited in 2018, allowing for a landslide win for the SVP. Nonetheless, the SVP party still suffered a remarkable decrease in absolute votes cast (from 147 000 to 127 000)

[18] Source: authors' calculations from 2001 and 2011 census data; available at https://astat.provincia.bz.it/downloads/mit38_2012.pdf; and http://www.statistica.provincia.tn.it/statistiche/societa/popolazione/; last accessed: March 9, 2019

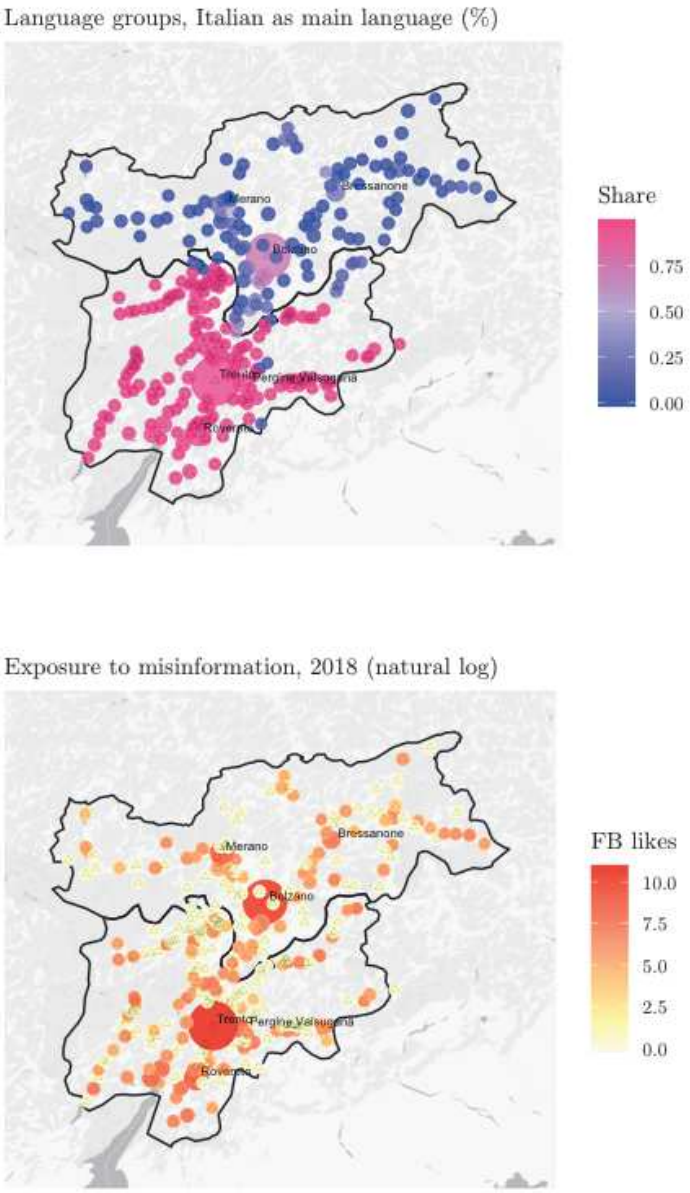

Figure 2: Language groups shares and exposure to misinformation (2018) in Trentino-Alto Adige/Südtirol. The size of the circle around each municipality indicates average electorate size between the two periods. Language groups shares are computed for the period in-between the two elections, estimated through a linear interpolation of Aitchison's (1986) log-ratios for the 2001 and 2011 census data shares. Exposure to misinformation is computed by taking the natural logarithm of the number of Facebook likes to disseminators of fake news in each municipality.

as turnout in South Tyrol fell sharply compared to 2013. In the areas with the highest representation of German-speaking voters the drop in participation was very steep, decreasing

from over 80% to slightly more than 50% of the electorate. An large number of spoilt and blank ballots was also recorded.[19] As the number of ballots cast for other parties increased only slightly, it is fair to assume that a large share of German-speaking voters opted to file a 'protest vote' by abstaining or filing an invalid ballot.

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.

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, income levels and 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.

Between the two elections, some municipalities were suppressed and new ones were formed from their merger; accordingly, a few adjustments had to be made. For all suppressed municipalities from 2013, we impute values for the corresponding 2018 entities either by summing the totals or, in the case of per capita figures, by averaging the population-weighted sum of the suppressed municipalities.

The present section provides an overview on how these sources were gathered and harmonised. Summary statistics on the core variables used in our analysis are presented in Table 1. We provide further information on how we collect and process data on misinformation in Appendix A. All data sources and codes we use to construct our final dataset and perform our analyses are available in the online data archive.

3.1 Electoral data

Official municipality-level data on general election results from the Trento and Bolzano/Bozen provinces are obtained from the Italian Ministry of the Interior,[20] where the complete election history of the Italian Republic is available. We extracted information on electoral results from the 2013 and 2018 general elections.

[19] Italian Ministry of the Interior, Archivio storico delle elezioni, Dipartimento per gli Affari Interni e Territoriali. Available at: https://elezionistorico.interno.gov.it/

[20] Italian Ministry of the Interior, Historical Archive of Italian Elections; available at: https://elezionistorico.interno.gov.it/index.php?tpel=C; last accessed: March 9, 2019

Other than indicating how many votes each party received in every election, the data set also includes a number of auxiliary variables broken down by municipality. Among these, electorate size and total number of abstentions or invalid votes were of particular interest for our research.

3.2 Socio-demographic and internet connectivity data

The size of language groups by municipality is a key variable in our identification strategy. Language group information for the region can be obtained through census data.[21] However, as census data for Italy is only released once every 10 years, the latest figures available date back to 2011. To compensate for differing trends in population growth across language group which may have affected our analysis, we perform a small and simple adjustment by interpolating the figures from 2001 and 2011 to predict the group-specific shares for the years 2013 and 2018, using Aitchison's (1986) log-ratio transformation to preserve the compositional form of the data. While the great majority of locations in Trentino are uniformly Italian-speaking, language groups shares vary significantly within South Tyrol: consequently, this indicator allows us to address the correlations between language groups and exposure to misinformation with much more precision than a binary discriminant between the two provinces would allow.

To reconstruct per capita income by municipality, we used tax data from the Italian Ministry of Economy and Finance for the years 2012 and 2017.[22] Those calculations are based on self-declared taxable income from the *Imposta sul reddito delle persone fisiche* (IRPEF), and per capita figures are already available by municipality. The connection between income and voting preference is not unambiguous, but has been discussed extensively in Galbraith and Hale (2008), Lewis-Beck and Nadeau (2011), and Hersh and Nall (2015), among others. Income differences can also arise across language groups, motivating our choice to control for them in our econometric specification.

Two different sources were used to construct statistics for broadband internet connectivity in 2013 and 2018. As mentioned previously, Campante et al. (2017) found that increases in internet connectivity have had an effect on voting preferences in Italy, suggesting that a measure for connection should be used as a control variable in our econometric model or even as an alternative to our language group instrument. Also, as Italian regions – including the two *province autonome* of Trentino and Alto Adige/Südtirol – possess considerable autonomy in the implementation of broadband infrastructures, this indicator then plays a vital role in controlling for differences in connectivity arising from staggered implementation of legislation. This control is especially important, in the light of recent evidence from Poy and Schüller

[21] Province-specific census data on language shares is available at: https://astat.provincia.bz.it/downloads/mit38_2012.pdf; and http://www.statistica.provincia.tn.it/statistiche/societa/popolazione/; last accessed: March 9, 2019

[22] Available at: https://www1.finanze.gov.it/finanze3/analisi_stat/index.php?tree=2013 and https://www1.finanze.gov.it/finanze3/analisi_stat/index.php?tree=2018; last accessed: March 31, 2019

(2020), who specifically exploit this staggered differences in connectivity in the Trento province to study the effects of broadband connectivity on electoral behaviour.

The 'digital divide' in Italy, defined as the share of households not covered by broadband connection, was originally covered by a set of municipality-specific indicators released by the 'Agenzia per la Coesione Territoriale'.[23] This indicator, which defined any landline connection whose speed exceeded 2Mbps as broadband, was then subtracted from unity to compute the share of low latency connections by municipality. As these publications were discontinued, the collection duty was then moved to the 'Autorità per le Garanzie nelle Comunicazioni' (AgCom), which has released internet connectivity indicators for 2018.[24] The new variables released by AgCom were slightly different from the previous ones, but it provided enough information to construct a fully comparable digital divide indicator using the available information on the number of households with a landline connection faster than 2Mbps.

3.3 Social media data

Information on fake news exposure and on the social media campaigns of Italian parties was all scraped from Facebook.

We first need to build a measure that captures the diffusion of fake news in each municipality. This indicator is particularly relevant as it represents our treatment, as detailed in Section 5. To do so, we collect data on the number of likes to Facebook pages that disseminate fake news for each municipality in the region using the Facebook Audience Insights tool.[25] This tool, generally intended for advertising purposes, allows us to access information on active Facebook users and filter it by place of residence, age, gender, language and many other interests. The tool yields demographic (broken down by age, gender, relationship status, education, and employment) and online activity (page likes and social media use) information on all active individuals in the targeted audience.

To identify the pages that disseminate fake news, we combine the 'interest' filter provided by the Insight tool with the lists of Facebook pages that disseminate fake news compiled by two Italian debunking websites.[26] Using the former, we are able to filter an audience through its 'interests' that is categorised under keywords, e.g. 'Fitness and wellness', 'Technology', and so on. By selecting the type of interest, the tool displays a list of pages which correlate with these interests amongst the selected audience, as well as the number of audience-specific likes for each of these pages. We find a number of interest clusters – spanning from neutral topics, such as 'news', to more politically charged ones, such as 'sovereignty' or 'immigration'–, to

[23] Available at: http://old2018.agenziacoesione.gov.it/it/arint/OpenAreeInterne/index.html; last accessed: March 31, 2019

[24] Available at: https://maps.agcom.it/#; last accessed: March 31, 2019

[25] See www.facebook.com/ads/manager/audiences.

[26] The websites are Butac.it and Bufale.net. More details are provided in Appendix A, which also provides the full list of pages used in the process.

return information on disseminator pages that are included in the debunkers' lists (see full list in Appendix A). Of these clusters, we select the one called 'News24' for two reasons. First, it is linked to the topic of news, so we can infer that users liked this page with the intention of consuming news. This does not apply to other pages such as the ones related to migration or sovereignty. Second, the tag 'News24' reports the highest number of fake news disseminators contained in the two lists compiled by debunkers.

We collect information on the likes to Facebook pages disseminating fake news for all users aged 24+, in order to ensure that these individuals had the opportunity to vote at both 2013 and 2018 elections, registered in each municipality of the Trentino-Alto Adige/Südtirol region. While the demographic information of the audience tool is biased towards Facebook users, we are confident that page likes are an unbiased measure of exposure to misinformation, as it is expected for the utmost majority of fake news to travel through social media.

Our fake news exposure indicator shows, for each municipality and for each year of election, the estimated number of Facebook likes being held by all Facebook fan pages that are known to spread politically-charged misinformation.[27] Due to the unavailability of granular data on the spread of each piece of false information, we decided to focus our attention on their 'disseminators', measuring – in a given municipality – the effect of each unitary increase in their social media following on aggregate electoral preferences. Our estimates then include both the 'intensive' and 'extensive' margin of fake news exposure, under the implicit assumption that individuals who 'liked' these pages also play an active part in spreading misinformation. The lengthy data collection and estimation process behind the construction of this variable is reported in Appendix A.

Our research question also requires to transpose categorical voting decisions into a continuous scale measuring affinity to populist discourse by municipality. The construction of such an indicator required access to propaganda content used over the course of the two electoral campaigns preceding the 2013 and 2018 elections. We turned, again, to Facebook, from which we scraped all posts from running parties and their leaders, covering all three months preceding the elections in 2013 and 2018, and coinciding with the beginning and the end of each electoral campaign. A more detailed description of the data and calculations leading to the production of our final indicators is provided in the next section.

4 A text mining approach to measuring the populist content of parties

As outlined in Section 3, descriptive evidence suggests that fake news benefits populist parties. In order to investigate the electoral impact of fake news, the dependent variable of our empirical

[27] Likes from all pages which appeared in black lists compiled by debunking websites Butac.it and Bufale.net (as discussed in section 3.2) were estimated. See Appendix A for the full list of pages used in the process.

Table 1: Summary statistics by province and year

	(1) Bolzano/Bozen			(2) Trento		
	2013	2018	Average	2013	2018	Average
Populist score (total)	5.972	6.526	6.249	4.123	5.899	5.011
	(13.376)	(15.823)	(14.621)	(12.074)	(15.465)	(13.883)
Exposed to fake news	148.424	376.876	262.650	199.858	521.874	360.866
	(869.421)	(2366.432)	(1782.496)	(1206.976)	(3459.335)	(2592.059)
Broadband connections	2469.032	3007.955	2738.494	1987.041	2217.626	2102.334
	(7542.113)	(7661.528)	(7590.388)	(6763.404)	(7030.532)	(6889.395)
Italian speaking voters	879.493	907.630	893.562	2173.312	2226.701	2200.007
	(5321.100)	(5435.457)	(5366.947)	(6837.192)	(6963.666)	(6890.933)
Income per capita	18721.313	20692.652	19706.982	17193.012	18434.036	17813.524
	(2534.504)	(2728.883)	(2807.309)	(2188.215)	(1891.333)	(2134.696)
Electorate size	3245.284	3332.853	3289.069	2276.602	2330.716	2303.659
	(7548.803)	(7683.594)	(7600.120)	(6986.644)	(7112.244)	(7039.726)
Abstentions and invalid votes	677.569	1345.802	1011.685	524.455	649.989	587.222
	(1668.889)	(2502.587)	(2148.627)	(1499.638)	(1966.976)	(1747.624)
Observations	232			352		

*N*otes: Mean coefficients, standard errors in parentheses.

model will therefore reflect the electoral performance of populist parties.

To define populism we follow the political economy and political science literature which regard as populists those parties that (1) share an anti-elite rhetoric (Mudde, 2004; Albertazzi and McDonnell, 2008; Pauwels, 2011; Rooduijn and Pauwels, 2011; Kaltwasser et al., 2017) and that (2) tend to adopt a particularly emotional language in their campaign (Taggart, 2000; Rooduijn, 2014; Caiani and Graziani, 2016; Bischof and Senninger, 2018). These two elements are generally complementary and do not exclude one another. As found by Guiso et al. (2017), the supply of populist parties tends to be higher where disappointment with traditional parties is greater. This suggests that the rise of populist parties goes hand in hand with an emotional narrative against incumbent politicians perceived as part of the establishment. Most importantly, this narrative is coherent with the one perpetuated in fake news, as shown in the previous section and in the literature on fake news.

To capture these two aspects of populist rhetoric, we create two indexes that measure the degree of populist rhetoric of each party based on the text of their Facebook posts. In particular, we create two dictionaries that capture (1) an anti-establishment rhetoric, and (2) a more emotional or assertive tone, and match them with the text of the Facebook posts published by all the Italian parties that ran in the 2013 and 2018 elections and their leaders.[28] The first dictionary contains 23 words which capture a populist tone in the Italian political

[28]We intentionally focus on populist rhetoric rather than on party-specific policy stances. This choice is motivated by the fact that, while on the one hand defining populism as linked to a specific policy stance is controversial (Caiani and Graziani, 2016), on the other hand, populism can be more easily identified by its forms of communication rather than by specific ideological stances (de Vreese et al., 2018).

language; for example, the words 'establishment' or 'caste'.[29] The second dictionary simply counts the number of exclamation marks as a proxy of an emotional tone.[30] While previous works applied dictionary techniques to measure populism on party manifestos (Rooduijn and Pauwels, 2011; Pauwels, 2011) and parliamentary speeches (Decadri and Boussalis, 2019), this paper is the first to our knowledge that uses them to measure the degree of populism in political communication on social media.

We then web-scrape the text of all the Facebook posts of Italian parties (12.159 posts) and their political leaders (8.164 posts) that posted three months before the Italian elections in 2013 and 2018. After pre-processing the text,[31] we first match the words of our two dictionaries against all Facebook posts collected. Formally, we compute the following indicator for each text bag:

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

where *Count* is the number of words matched for each text bag in each Facebook post $d = \{1, ..., D\}$ published by each party/leader $k = \{1, ..., K\}$ three months ahead of each election $t \in [2013, 2018]$. σ is an arbitrary re-scaling factor (which we set at 10). T is the total number of words contained in each Facebook post d, and *Engagement* is the sum of all Facebook shares and likes a given post has received. This last operation allows us to assign higher scores to those posts that were more visible on Facebook before the elections. In this way our index does not simply describe the supply of populism on social media, which is captured by the count of populist words or exclamation marks, but also those parties that were 'visibly' more populists online, which is captured by the number of likes and shares given to each post.

In other words, we sum the number of matches (*Count*) and weight them by the total number of words[32] (T) and by the number of interactions (*Engagement*) for each party. Values of *Popscore* for each party and elections are displayed in Fig. 3 for both the dictionary (panel on the left) and the exclamation marks (right). To obtain an indicator of populism at the municipal level, we multiply the populist score of party k in election t ($Popscore_{tw}$) by the

[29] As reported by the encyclopedia Treccani, the term 'caste' (in Italian 'casta') has been used in the public debate in Italy since 2007 to denote a privileged and at the same time corrupt political class. Link to the article by Treccani: http://www.treccani.it/magazine/lingua_italiana/speciali/crollo/Grassi.html (last accessed on 6 May 2020).

[30] Previous works in computer science, such as Kumar and Sebastian (2012), find exclamation marks to be a good predictor of emotional statements on social media.

[31] In order to allow for matching, we tokenise the text, remove Italian stopwords, punctuation (except exclamation marks in the second dictionary), numbers and white spaces and transform all terms in lower case. We also translate each term of the text-bag in German to compute the scores also for the German-speaking parties SVP and DF.

[32] This practice aims to avoid that long posts inflate the populist score. Intuitively, posts containing larger number of words are more likely to contain a higher number of words that are matched.

votes gained by the same party p in municipality i during election t. Formally, we aggregate the scores as follows:

$$Y_{ti} = \sum_{k=1}^{K} Popscore_{tk} * Votes\_received_{tki} \tag{4.2}$$

where $Votes\_received$ is the number of votes received by each party k in each election t in each municipality i. Y_{ti} will be our indicator of populism at municipal level and our main dependent variable. Given these transformations, our scores should not be assessed by looking at their absolute values, which are necessarily low, but rather by focussing on the relative distance in the scores between parties.

The two indexes display similar trends, further confirming the belief that anti-establishment discourses and emotional tones are generally correlated in political communication. In particular, Lega clearly features as the most populist party in the election of 2018 in both indexes, as a result of a steep increase in populist language from 2013. On the other hand, M5s shows a lower degree of populist language in 2018 than in 2013, for both cases. In line with our expectations, the incumbent party, PD, displays lower levels of populism. The ability of our methodologies to proxy for a populist language is further reinforced by the proximity of fringe parties, which are generally anti-establishment, to Lega and M5s.

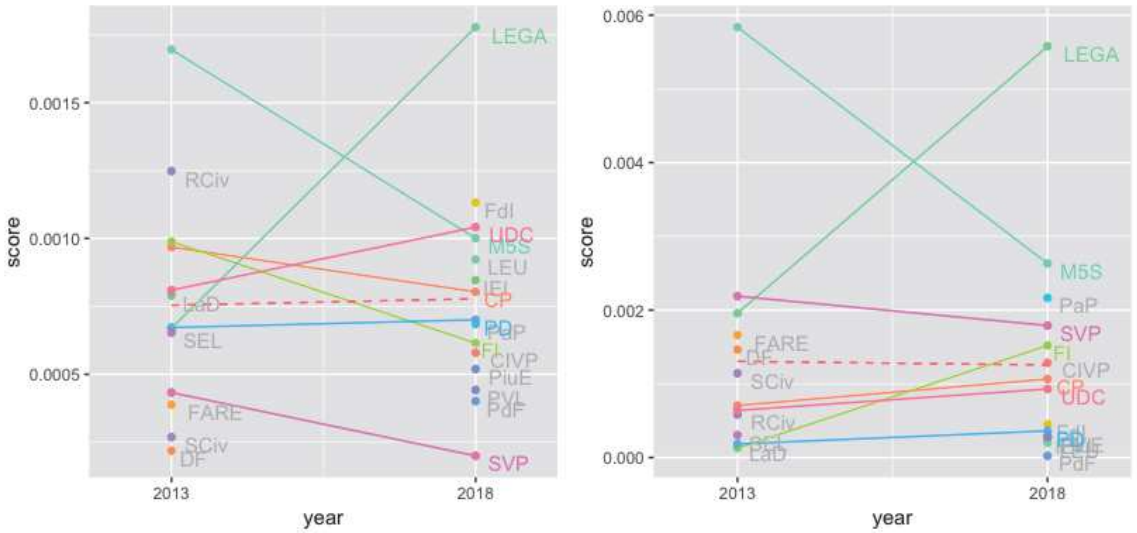

Figure 3: Text analysis scores of social media posts from parties and their leaders during the 2013 and 2018 elections campaigns. The left figure refers to the scores obtained with the populist text-bag, while the right figure computes the same score as the frequency of exclamation marks in the text. Parties in grey have only took part in one of the two elections. The red dashed line refers to the election specific average.

However, such proximity is not constant, particularly for minor forces. In a few cases, counter-intuitive results might find an explanation through a deeper study of political dynamics. The clearest example is the small Christian-democratic formation UDC: this party

presents low scores in 2013, when it run within a centrist, pro-incumbent coalition, and high in 2018, when it was in a coalition with centre-right parties, then in 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 much more prominently in 2018.

Despite their relative limitations, these results confirm the widespread categorisation of parties such as Lega and M5s as populist. As a further test, we compared our results to the values assigned by political experts in the Chapel Hill Expert Survey (Polk et al., 2017)[33] 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, in particular displaying Lega and M5s among the most populist parties in Italy (see Figures C.1 and C.2 in Appendix C).

As our index combines a number of elements including the votes gained by each party and its engagement on Facebook, Table 2 provides further details on the composition of this index by party. The table shows that from 2013 to 2018 most parties enhanced their presence of social media, increasing the number of posts published and the number of likes and shares received by their followers.

The parties that published the most on Facebook in 2018 are by far the Lega and the M5s. This is in line with the findings of Giglietto et al. (2018), for which the social media engagement sources close to these two parties outperformed the ones close to other parties. The Democratic Party ranks third in number of Facebook posts, but has a very low score in terms of populist language compared to the Lega and M5s, reflecting its moderate stances and position as incumbent. The low engagement of Forza Italia, which was the third largest party in 2013 in terms of votes, is not surprising. As documented by Giglietto et al. (2018), the party reported a very low social media presence with its leader, Silvio Berlusconi, opening his Twitter account just in October 2017.

On the contrary, the high scores of a fringe party such as Casapound, should not be surprising. While this party received a share of votes significantly lower than Forza Italia (in 2018 Casapound received 0.95% votes against the 14% of Forza Italia), the former has been particularly active on Facebook. As of 29 March 2020, the Facebook page of Casapound counts 277,948 likes, against the 216,273 of Forza Italia. This is just below the number of likes of the Facebook page of the Partito Democratico (331,809 likes), but largely inferior with respect to the pages of the Lega (670.157; named 'Lega - Salvini premier' as of March 2020) and of the M5s (1,481,152). However the high engagement of the last two parties is mostly driven by the Facebook pages of their leaders: the Facebook page of Matteo Salvini, leader of the Lega, has 4,122,945 likes, while the one of Luigi Di Maio, who was the party candidate for prime minister of the M5s in 2018, has 2,272,858 likes.

[33]Data are available at the following link: https://www.chesdata.eu/.

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

Party	Votes received at national level (%)		Populist Score (text bag 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	6,570.66	3,557.16	1,798	2,020	4,247,119	12,739,192
Partito Democratico (PD)	25.43	18.76	263.01	424.22	719	1,062	559,794	3,889,217
Lega	4.09	17.35	1,618.56	6,723.754	615	1,994	37,211	14,322,120
Forza Italia* (FI)	21.56	14.00	0.00	1,679.96	197	668	968,151	1,730,078
Fratelli d'Italia (FdI)	1.96	4.35	N/A	530.02	N/A	1,580	N/A	4,570,784
Liberi e Uguali (LeU)	N/A	3.39	N/A	192.65	N/A	356	N/A	546,399
Più Europa (PiuE)	N/A	2.56	N/A	106.26	N/A	500	N/A	640,781
Unione di Centro (UDC)	1.79	1.30	876.46	1,243.324	304	396	103,722	194,736
Potere al Popolo (PaP)	N/A	1.13	N/A	2,314.47	N/A	992	N/A	650,732
Casapound (CP)	0.14	0.95	1,140.35	903.43	753	835	146,040	1,930,393
Italia Europa Insieme (IEI)	N/A	0.58	N/A	12.71	N/A	541	N/A	100,496
Civica Popolare (CIVP)	N/A	0.45	N/A	1,420.63	N/A	429	N/A	184,056
Südtiroler Volkspartei (SVP)	0.43	0.41	1,444.57	1,231.21	385	373	3,371	10,368
Il Popolo della Famiglia (PdF)	N/A	0.67	N/A	24.64	N/A	420	N/A	160,875
Partito Valore Umano (PVL)	N/A	0.15	N/A	302.99	N/A	41	N/A	13,388
Die Freiheitlichen (DF)	0.14	N/A	1,423.76	N/A	270	N/A	2,200	N/A
Scelta Civica (SC)	8.30	N/A	1,229.78	N/A	494	N/A	420,434	N/A
Sinistra, Ecologia e Libertà (SEL)	3.20	N/A	530.35	N/A	829	N/A	847,083	N/A
Rivoluzione Civile (RCiv)	2.25	N/A	694.08	N/A	607	N/A	319,163	N/A
Fare per Fermare il Declino (FARE)	1.12	N/A	1,825.95	N/A	526	N/A	570,311	N/A
La Destra (LaD)	0.65	N/A	333.32	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 that specific election. Engagement is the sum of the likes, shares and comments received on each posts by the party and leader Facebook page. *Forza Italia ran as Popolo delle Libertà in 2013. The party leader was Silvio Berlusconi in both cases.

5 Empirical framework

Our aim is to understand if the diffusion of fake news in 2018 is responsible for the electoral growth of populist parties we described in the previous section. The hypotheses we intend to test is whether fake news exposure increased the votes for populist parties. The null hypothesis is that fake news exposure has no effect on votes for populist parties. The null hypothesis is motivated by the self-selection mechanism (described in the Introduction) for which populist voters tend to be more exposed to fake news.

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:

(5.1) $Y_{ti} = \alpha + (F_{ti} \times P_t)\lambda + F_{ti}\delta + P_t\zeta + X_{ti}^{'}\gamma + e_{ti}$

The outcome variable is Y_{ti}, 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 Equation 4.2, which combines the number of votes received by each party in each municipality in 2013 and 2018 with party- and election-specific populist scores, as described in detail in Section 4. We will replace this indicator with different measures of populist voting to test the robustness of our results. F captures exposure to fake news measured as the number of likes to Facebook pages that disseminate fake news. P is 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.

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, then we would expect $\lambda > 0$.

Equation 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. In this context, we need to compare λ with an alternative estimate that isolates the effect of fake news from the selection bias.

Based on our discussion on the linguistic differences in the region of Trentino Alto-Adige, we exploit the 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:

(5.2.1) $F_{ti} = \alpha_1 + (Z_{ti} \times P_t)\lambda_1 + Z_{ti}\delta_1 + P_t\zeta_1 + X_{ti}^{'}\gamma_1 + e_{1ti}$

(5.2.2) $(F_{ti} \times P_t) = \alpha_2 + (Z_{ti} \times P_t)\lambda_2 + Z_{ti}\delta_2 + P_t\zeta_2 + X_{ti}^{'}\gamma_2 + e_{2ti}$

(5.2.3) $Y_{ti} = \alpha_3 + (\widehat{F_{ti} \times P_t})\lambda_3 + \hat{F}_{ti}\delta_3 + P_t\zeta_3 + X_{ti}^{'}\gamma_3 + e_{3ti}$

Exposure to fake news is captured by F in the first-stage regressions (5.2.1) and (5.2.2), where its value – and the value of its interaction with the year of election in (5.2.2) – is predicted by the instrument Z (indicating, again, assignment to the Italian linguistic group) and other covariates. The fitted values $\hat{F}$ and $\widehat{F_i \times P_t}$ from first stage regressions (5.2.1) and (5.2.2) are then plugged into equation (5.2.3) to predict populist preferences Y.

The coefficient of interest is λ_3, which captures the causal effect of fake news exposure on electoral support for populism. More precisely, λ_3 indicates the increase in the electoral support for populist party for each additional like to a Facebook page that disseminates fake news in each municipality.

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 stages of our model (5.2.1) and (5.2.2), 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_3 will capture the causal effect of fake news exposure on electoral support for populism.

Although in the first-stages we test whether the German-speaking population is unexposed to fake news on the Italian election, our model is unable to control for the possibility that German-speaking voters are exposed to fake news in German that could affect their voting behaviour. While this is an important caveat that applies to our model, it is unlikely to affect our estimates for two main reasons motivated by the findings of Allcott and Gentzkow (2017).

First, Italian fake news disseminators have no incentive to produce fake news on the Italian election in German. Allcott and Gentzkow (2017) show that producers of fake news are motivated by either profit – i.e. to maximise the advertising revenue of their Facebook page –, or ideology, i.e. to increase the electoral gain of their party. Given the small size of the German-speaking community in the region relatively to the general Italian population, the marginal return in terms of economic revenue and electoral gains of a piece of misinformation concerning Italian politics is very small since misinformation is only accessible to this specific population. Disseminators are hence better off producing fake news in Italian.

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 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.[34] It is therefore very unlikely that these hoaxes could have traction on both foreign German-speaking and Italian voters.

Moreover, owing to the way the dependent variable is constructed, the model is robust to the potential occurrence of an exogenous shock that increases simultaneously the supply and demand of populism in one linguistic community and not in the other. This is relevant since variations in the electoral outcome of a party could be related to this unobserved change rather than to fake news exposure.

Our text-based indicator controls for this heterogeneity both on the supply- and demand-side of populism. As shown in Section 4, on the one hand, our indicator captures 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,

[34]This is a commonality with the fake news database of Allcott and Gentzkow (2017), where most fake news targeted specific politicians – mostly the two candidates – and mentioned their names in the title.

it controls for variations in the demand for populism – or, more precisely, in the demand for populist rhetoric.

It is important to notice that, since our indicator of populism is continuous, it allows us to control for changes in the supply of populism in-between the two elections. In this way, we can account for party-specific changes between 2013 and 2018. Votes for the same party across the two elections should not be weighted equally if, across the years, this party has revised its populist stance: undergoing, for example, a radical transformation into a populist, or an establishment, force.

6 Results

Table 3 shows the results from an initial linear model, using, as for all subsequent specifications, the natural logarithm of total populist scores by municipality as the outcome variable. This has been computed by aggregating text analysis scores obtained through the 'Assertive' dictionary and weighting them by the size of the electorate. Estimates from columns (1) to (4) indicate that the positive correlation between exposure to fake news and populist vote is robust to the addition of various controls. In order to compare the coefficients, we standardise each variable by subtracting its mean and dividing it by its standard deviation (for the results with the non-standardised coefficient see Table D.6 in Appendix D).

Column (1) presents our baseline model, including exposure to fake news, year of election (a dummy that equals 1 for 2018 and 0 for 2013), their interaction, and the number of Italian-speaking voters by municipality. As expected, populist vote increases when the year dummy is equal to 1, as 2018 coincided with the electoral success of populist platforms. Nonetheless, the presence of confounding factors makes the results difficult to interpret.

Column (2) adds the number of landline low-latency connections to the covariates. The introduction of the variable in the model increases the model precision from a 86.8% to a 99.0% R-squared statistic. The effect of exposure to fake news, given again by the interaction term, is significant, and surpasses the coefficients for language group in magnitude. The interaction coefficient is again unaffected from the introduction of controls for income per capita in Column (3).

In Column (4) we add electorate size and abstention behaviour as further controls, increasing the adjusted R-squared to 99.6%. The negative link between abstention and populist vote is not unexpected and indicates that populist platforms may have attracted a number of individuals disenchanted by traditional politics. Income per capita also turns significant, suggesting that, in higher income municipalities, voters are less prone to vote for populist parties. Most importantly, electorate size is one of the most significant predictors in the model, indicating the presence of a non-linear relationship between population density and populist vote.

Table 3: OLS estimates of the effect of misinformation on populist vote, standardised variables

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Exposed to Fake News	1.764***	-0.249***	-0.243***	-0.356***
	(0.239)	(0.060)	(0.062)	(0.064)
Exposed to Fake News × Year of election	-0.942***	0.276***	0.273***	0.371***
	(0.245)	(0.044)	(0.044)	(0.047)
Year of election	0.037***	0.006*	0.005	0.033***
	(0.014)	(0.004)	(0.004)	(0.004)
Broadband connections		1.018***	1.014***	0.096**
		(0.030)	(0.031)	(0.045)
Income per capita (natural log)			0.005	-0.005**
			(0.004)	(0.003)
Electorate size				1.072***
				(0.071)
Abstentions and invalid votes				-0.121**
				(0.048)
Observations	584	584	584	584
Adjusted R-squared	0.868	0.990	0.990	0.996

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

The effect of exposure on voting remains statistically significant, indicating that increasing the likes to a misinformation disseminator page in a given municipality by one standard deviation results in a 0.37 standard deviation increase in the populist score of that municipality. While the magnitude of this effect is relatively large, it should be noted that the size of the electorate is a stronger predictor. Moreover, the order of magnitude of the coefficients of the share of Italian speaking voters is still relatively large and significant.

Our OLS results show a correlation between populist preference and exposure to misinformation. However, these estimates may 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 Column 4, including reduced form (Column 1) and first stage (Columns 2 and 3) estimates (for the unstandardised

coefficients see D.7 in Appendix D). Here, the number of Italian-speaking voters by municipality and its interaction with year of election are instrumented to predict the exposure to fake news (Column 2) and its interaction with year (Column 3). With the exception of the instrument and the endogenous exposure variable, the specifications in Table 4 retain the same control group used in our final OLS model, as in column (4) of Table D.6.

Table 4: 2SLS estimates of the effect of misinformation on populist vote, standardised variables

	(1)	(2)	(3)	(4)
	DiD	First Stages		2SLS
VARIABLES		Exposure	Interaction	
Italian speaking voters	-0.026	0.363**	-0.025	
	(0.028)	(0.149)	(0.068)	
Italian speaking voters × Year of election	0.152***	0.595***	0.965***	
	(0.007)	(0.079)	(0.120)	
Year of election	0.015***	-0.045***	-0.076***	0.027***
	(0.002)	(0.007)	(0.011)	(0.003)
Broadband connection	-0.021	0.067	0.044	-0.025
	(0.058)	(0.246)	(0.143)	(0.073)
Electorate size	0.969***	0.069	-0.062	0.985***
	(0.050)	(0.168)	(0.171)	(0.060)
Income per capita (natural log)	0.005***	-0.009	-0.003	0.005
	(0.002)	(0.011)	(0.008)	(0.003)
Abstentions and invalid votes	-0.036	-0.039	0.056	-0.049
	(0.041)	(0.120)	(0.140)	(0.054)
Exposed to fake news				-0.058
				(0.083)
Exposed to fake news × Year of election				0.193***
				(0.038)
Observations	584	584	584	584
Adjusted R-squared	0.997	0.931	0.930	0.994
Partial R-squared		0.342	0.397	
F-Test		45.03	59.19	

Notes: IV estimates (including reduced form - DiD - and first stages) for the effect of misinformation on populist vote. Populist scores computed using the 'Assertive' text bag. F-tests for excluded instruments for the individual instrument (voters in the Italian-speaking language group) and its interaction with year of election are reported as F-Test (exposition) and F-Test (interaction), respectively. Standard errors robust to clustering by municipality in parentheses.

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

Reduced form estimates are displayed in column (1) of Table 4. As discussed earlier, this specification corresponds to the difference in differences model of equation (1), where the effect of the exogenous variable on the outcome is given by the interaction term between language

group shares and year of election.

Diff-in-diff estimates indicate a positive and statistically significant effect of the Italian language group on voting in the 2018 election. Accordingly, after controlling for all observables, the populist content of the vote in each municipality increased by 0.15 standard deviations for each additional Italian-speaking voter, during the 2018 general elections. These figures indicate a smaller coefficient compared to the one of fake news exposure in the final OLS specification (Table D.6, Column 4). Reduced form figures are proportional to the causal effect of interest, therefore the positive sign might still indicate the presence of an effect. Nonetheless, unless we assume that exposure to misinformation is completely proxied by language groups, the coefficient of the interaction term is not a substitute for the effect of exposure to fake news on voting.

Turning at our instrumental variable estimates, first stage regressions predicting exposure to fake news and its interaction with year are shown in columns (2) and (3). The positive effect of the interaction between language group and year of election confirms our assumptions on differential exposure to fake news based on linguistic grouping. Results from column (3) indicate that, for each additional Italian-speaking voter, the number of likes to fake news disseminators increased by 0.93. Both models predict exposure with relative precision, boasting, overall, a 93.1% and a 93.0% adjusted R-squared. 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, both instruments pass the F-test for excluded instruments (as suggested in Bound et al., 1995), increasing our confidence in our estimates.

Our final estimates are presented in column (4), showing that fake news had a small but statistically significant effect on contributing to the rise of populist platforms. An increase in the likes of a disseminator page of one standard deviation has the effect of a 0.19 standard deviation increase in populism in the same municipality. More importantly, while the interaction coefficient between fake news exposure and year is statistically different from zero, it is lower than the OLS model in Table D.6, dropping from 0.37 to 0.19.

These results take some credit for the success of populist parties in 2018 away from the spread of fake news, and support the hypothesis that voters self-select into misinformation bubbles and consume fake news because of their prior preference for populist platforms, and not the other way around.

The electorate size and the year of election dummy are good enough for our language group instrument not to affect our estimates. We therefore conclude that, while fake news has had an effect on voting outcomes in Trentino-Alto Adige/Südtirol over the course of the 2018 elections, over half of this effect is explained by self-selection.

6.1 Robustness Checks

6.1.1 Exclusion Restriction

An issue that needs to be solved concerns the exclusion restriction assumption underlying our estimates. This assumption requires fake news to be the only channel through which the share of Italian-speakers impact on populism in a municipality.

However, there could be alternative channels. One channel is cultural: German speakers tend strongly to support German-speaking parties, such as the SVP and the DF. We account for this issue using the DiD approach, which compares the trends of the treatment (Italian-speaking) and control (German-speaking) groups.

One time-varying channel related to this is the supply of populism. It could be argued that while the supply of populism increased among Italian-speaking parties between 2013 and 2018, it decreased among German-speaking ones in the same year. According to our indicator of populism, in 2018 the SVP became less populist. Moreover, the DF, a German-speaking party which scored relatively high in our index of populism in 2013, did not run in 2018. In this context, German-speaking voters with populist preferences could have shifted their vote, previously assigned to a 'more populist' SVP and to the DF, to Italian populist parties. If this is the case, the positive effect of Italian speakers on populism that we identify might not be driven by the diffusion of fake news among this share of the population, but by the shift of German populist votes to Italian populist parties due to the lower supply of 'German-speaking populism' in 2018.

On the other hand, given the low supply of populism, the same German-speaking voters could have decided simply not to vote. These would result in an increase of abstentions among the German-speaking population. Indeed, OLS estimates in Table 3 display a negative coefficient that links abstentions and invalid votes with populism. If this is the case, then our estimates on the effect of fake news would not be biased.

To account for this possibility we implement two new specifications similar to the ones in the previous section, but with different dependent variables and a different treatment variable. We first create a new treatment variable, G, that measures the support of German-speaking populist parties. This variable is the sum of the populist score of the two German-speaking parties, the SVP and the DF, multiplied by their electoral outcome.

In the first specification, the dependent variable is the number of abstentions in each municipality, which we denote as Abs. If the effect of the interaction between German speaking voters on abstentions is positive, then our estimates on fake news are unaffected by the decrease in the supply of German-speaking populism in the region.

In the second specification, we apply the same model but replace the dependent variable with the votes for Lega and M5s, which are the two Italian-speaking parties with the highest populist scores across the two elections. We denote this variable as $ITPopulist$. In this way,

we test whether those German-speaking voters that supported German-speaking populism shifted their vote to Italian-speaking populist parties in 2018 as a reaction to the decrease in the supply of German-speaking populism.

Formally, we run the same IV model proposed before, but replacing the treatment variable with support for German-speaking populist platforms, and replacing the dependent variable with abstentions in the first set of results and with support for the Lega and M5s in the second one. Formally, our specifications are as follows:

(6.2.1) $G_{ti} = \alpha_1 + (Z_{ti} \times P_t)\lambda_1 + Z_{ti}\delta_1 + P_t\zeta_1 + X_{ti}^{'}\gamma_1 + e_{1ti}$

(7.2.2) $(G_{ti} \times P_t) = \alpha_2 + (Z_{ti} \times P_t)\lambda_2 + Z_{ti}\delta_2 + P_t\zeta_2 + X_{ti}^{'}\gamma_2 + e_{2ti}$

(6.2.3) $Abs_{ti} = \alpha_3 + (\widehat{G_{ti} \times P_t})\lambda_3 + \hat{G}_{ti}\delta_3 + P_t\zeta_3 + X_{ti}^{'}\gamma_3 + e_{3ti}$

(6.2.4) $ITPopulist_{ti} = \alpha_4 + (\widehat{G_{ti} \times P_t})\lambda_4 + \hat{G}_{ti}\delta_4 + P_t\zeta_4 + X_{ti}^{'}\gamma_4 + e_{4ti}$

Our coefficients of interest are λ_3 and λ_4. If λ_3 is positive and significant, it means that German-speaking voters that supported populism tended to abstain in 2018. However, if λ_4 is also positive and significant, it means that many of these voters turned to Italian populist parties. For our results on fake news to be valid, we would therefore need the support for German-speaking populism to have a positive and significant effect on abstention and to be unrelated with support for the Lega and M5s.

The results for abstentions are displayed in Table 5. The results of the reduced form model (Column 1) show that the share of Italian speakers has a negative and significant impact on abstentions. This signals that the share of German-speaking inhabitants has a positive impact on abstentions. The sign turns positive when interacted with year 2018, as the number of abstentions grew among Italian-speaking voters too.

The results of the first stage show a negative coefficient of the share of Italian-speaking voters on the support for German-speaking populism. This means that as the share of German-speaking voters increases, so does the support for German populist parties, which results in a 2.64 standard deviation increase. This result is weaker when the treatment is interacted with year 2018 as the supply of German-speaking populism is lower.

Finally, Column 4 shows that the support of German-speaking populist parties has a positive impact on abstentions. This result goes in the direction of the hypothesis for which German populist voters did not shift to Italian-speaking populists. This result therefore provides evidence against the bias of our estimates due to the lower supply of German-speaking populism.

We test whether some of the support to German-speaking populism that did not result in abstentions was channelled to Italian-speaking populist parties. Therefore, in Table 6 we replace the dependent variable with votes for Lega and M5s, the parties that reported the highest scores in populism.

The reduced form model (Table 6, Column 1) shows that, as expected, the share of Italian-speaking population has a positive effect on the support for Lega and M5s. Columns (2) and (3) are identical to Table 5 as they represent the first stage regressions. The final results show that the support for German populism did not impact significantly on the growth of the two Italian populist parties.

Table 5: 2SLS estimates of the effect of German-speaking populist support on abstentions

	(1)	(2)	(3)	(4)
	DiD	First Stages		2SLS
VARIABLES		Exposure	Interaction	
Italian speaking voters	-0.459***	-2.645***	-2.158***	
	(0.040)	(0.103)	(0.170)	
Italian speaking voters × Year of election	0.214***	-0.058	0.355*	
	(0.064)	(0.087)	(0.198)	
Year of election	0.050***	-0.120***	0.272***	-0.109
	(0.008)	(0.018)	(0.038)	(0.073)
Broadband connection	0.265**	-1.504***	0.665	-0.369
	(0.112)	(0.232)	(0.517)	(0.412)
Electorate size	1.231***	4.614***	1.571***	1.271***
	(0.255)	(0.219)	(0.398)	(0.407)
Income per capita (natural log)	-0.004	0.015	0.035**	-0.019***
	(0.005)	(0.009)	(0.014)	(0.006)
Populist score	-0.224			
	(0.276)			
German-speaking populist support				-0.216
				(0.132)
German-speaking populist support × Year of election				0.476***
				(0.181)
Observations	584	584	584	584
Adjusted R-squared	0.983	0.908	0.622	0.943
Partial R-squared		0.675	0.292	
F-Test		455.23	28.16	

Notes: IV estimates (including reduced form - DiD - and first stages) for the effect of misinformation on populist vote. Populist scores computed using the 'Assertive' text bag. F-tests for excluded instruments for the individual instrument (voters in the Italian-speaking language group) and its interaction with year of election are reported as F-Test (exposition) and F-Test (interaction), respectively. Standard errors robust to clustering by municipality in parentheses.

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

Table 6: 2SLS estimates of the effect of German-spaking populist support on votes for Lega and M5s

	(1)	(2)	(3)	(4)
	DiD	First Stages		2SLS
VARIABLES		Exposure	Interaction	
Italian speaking voters	0.585***	-2.645***	-2.158***	
	(0.020)	(0.103)	(0.170)	
Italian speaking voters × Year of election	0.149***	-0.058	0.355*	
	(0.031)	(0.087)	(0.198)	
Year of election	0.004*	-0.120***	0.272***	-0.286
	(0.002)	(0.018)	(0.038)	(0.211)
Broadband connection	0.167***	-1.504***	0.665	-1.624
	(0.034)	(0.232)	(0.517)	(1.092)
Electorate size	-1.029***	4.614***	1.571***	2.703**
	(0.106)	(0.219)	(0.398)	(1.072)
Income per capita (natural log)	0.005***	0.015	0.035**	-0.003
	(0.002)	(0.009)	(0.014)	(0.013)
Populist score	1.151***			
	(0.112)			
German-speaking populist support				-0.834**
				(0.371)
German-speaking populist support × Year of election				0.756
				(0.523)
Observations	584	584	584	584
Adjusted R-squared	0.997	0.908	0.622	0.563
Partial R-squared		0.675	0.292	
F-Test		455.23	28.16	

Notes: IV estimates (including reduced form - DiD - and first stages) for the effect of misinformation on populist vote. Populist scores computed using the 'Assertive' text bag. F-tests for excluded instruments for the individual instrument (voters in the Italian-speaking language group) and its interaction with year of election are reported as F-Test (exposition) and F-Test (interaction), respectively. Standard errors robust to clustering by municipality in parentheses.

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

6.1.2 Alternative measure of populism

For robustness, we present alternative estimates in the Appendix using the alternative populist score based on the 'Anti-establishment/aggressive' text bag (as exposed in Section 4). The two tables replicate the ones presented in this section, with Table B.2 showing OLS estimates and Table B.3 displaying IV estimates using language group as instrument. Our results and interpretations remain mostly unchanged. Table B.2 shows that a positive correlation between

populist preference and exposure is still present. Again, these results are overturned with our IV estimates in Tables B.3, showing that exposure had a negative and significant, yet negligible, effect on voting when treatment is assigned through random variations in language groups.[35]

6.1.3 The impact of fake news on individual parties

Table B.4 in the Appendix shows estimates under different outcome variables, investigating whether the spread of misinformation has generated electoral gains for a specific group of parties, and how it has affected turnout. In this specification, we leave populist scores aside and estimate the direct effect of misinformation on the electoral outcomes of the major anti-establishment party platforms in the region (and in the country).

This model retains the same specifications used for Table 4, but focuses on the empirical connections – discussed in Section 2 – between misinformation and the votes for Lega and M5s, testing whether these parties have benefited from any electoral gain by the introduction of fake news. This analysis also investigates whether misinformation has led to electoral gains for Lega at the expenses of M5S, and also discusses whether the incumbent party, PD, lost any votes because of fake news.

In all cases, the IV results display a smaller effect than the OLS estimates and, in column (2), the language group instrument even shows a -0.45 reduction in votes to Lega and M5s for each additional like to a disseminator.

Studying the effects of misinformation on the electoral outcomes of Lega and M5s is, however, not sufficient in the context of the region. As discussed earlier, due to the differences in voting history between language groups the German-speaking population might be less inclined to vote for Italian-speaking parties, explaining the negative coefficient.

The use of a continuous indicator for populism allowed us to bypass this issue, serving as an effective control for the supply of populism in the region. Now that we are looking at specific electoral outcomes, an anti-establishment choice for the German-speaking population needs to be available. As discussed in Section 2, many German-speaking voters have opted to abstain or to vote for the Die Freiheitlichen (DF) party. We then measure the impact of misinformation on vote to Lega, M5s, DF or abstentions in columns (3) and (4). Our IV results point at a significant 0.57 increase in anti-establishment turnout for each additional like to a fake news disseminator in a given municipality. A self-selection mechanism is, again, noted, as this estimate is lower that the 0.79 estimate obtained with OLS, supporting our prior results obtained under the populist index.

Furthermore, in columns (5) and (6) we also investigate whether misinformation has impacted voting behaviour within the anti-establishment platforms. To do so, we estimate the impact of misinformation on votes to Lega, holding votes to M5s, DF and abstensions fixed.

[35] First stages are reproduced for clarity purposes but, for obvious reasons, they are identical to the ones shown in Table 4.

While the OLS results point at positive gains for Lega, the IV coefficient indicates that misinformation has not benefited a specific party within the anti-establishment platform.

Finally, in Table B.5 we provide the results of our model using using three alternative classifications of populist parties proposed in Guiso et al. (2017) and based on recent works in the literature on populism. We replace the dependent variable with the sum of votes received by those parties considered as populist by (1) the *Popu-List* database[36], (2) VanKessel (2015) and (3) Norris and Inglehart (2019). In all three cases fake news has a positive effect on populist voting and, in line with our baseline results, OLS coefficients are larger than the ones of the IV.

The large effect of fake news on populist vote of these specifications compared to the baseline model can be explained by two main reasons. First, these indicators derive from studies that compare populism across countries and therefore limit the sample to large parties. For this reason, the three classifications include parties such as M5s, Lega, Forza Italia and Fratelli d'Italia, but exclude smaller fringe parties with high populist scores such as the extreme right-wing Casapound and the extreme left-wing Potere al Popolo. As small populist parties are excluded from these classification, the estimates of the impact of fake news on populist voting are inflated.

Second, these indicators are less precise since they are time-invariant and based on a binary classification. These two limitations make the estimates unable to take into account within-party and between-party variability in populism. More precisely, the time-invariant feature of these indicators makes them unable to account for changes in each party's populist stance from one election to the other. Their binary feature makes them unable to distinguish between different degrees of populism across parties.

7 Conclusions

The measurement of the influence of fake news on electoral behaviour has, so far, escaped empirical assessment. With this study, we set to fill this void, identifying both the presence of this influence and its magnitude.

In order to identify the direction of causation between fake news and populism, we rely on a quasi-experimental approach. Gathering municipality-level data on electoral outcomes, demographics and social media usage, we exploit the bilingual context of the Trentino-Alto Adige/Südtirol in the 2018 Italian elections to randomise exposure to fake news.

We show that misinformation had a positive effect on the populist vote in Trentino and South Tyrol during the Italian 2018 general elections. However, our results indicate that exposure to fake news is largely dictated by self-selection in misinformation bubbles, meaning

[36] Available at http://www.popu-list.org/.

that the causal channel between voting and fake news goes both ways, with many individuals being exposed to misinformation because of their pre-existing political preferences.

This does not mean that fake news is less problematic, but only that the fundamental causes of the populist shift in voting have to be found elsewhere, and that misinformation still thrives within these filter bubbles. At the same time, the rise of negative externalities related to the spread of misinformation is not to be excluded. In cases such as, for example, public health crises, misinformation might generate cascading effects where the behaviour of a misinformed single individual might negatively affect others.

In terms of electoral preferences, however, fake news might stand as the embodiment of shared narrations within groups of voters which are further reinforced by confirmation bias and the increasing personalisation of social media echo chambers. It may be possible that, once entered a misinformation bubble, partisan opinions are being reinforced by the presence of fake news, which might then ensure continued support for partisan beliefs as reputable sources of information are progressively removed and discredited in favour of 'alternative facts'. In the presence of personalised filter bubbles, preference predates facts and not the other way around; social media plays a role in as much as it provides individuals with the information they want to believe. Cascading effects might then be contained within groups of like-minded individuals.

Our final notes address the validity of our results and suggest future pathways for research, in the hope that our work spurs further empirical and theoretical contributions on the role misinformation has on voting.

The exploitation as a natural experiment in 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. A similar methodology to ours could be applied to different contexts and with different units of observation, as language groups and broadband penetration have proven as good predictors for access to fake news. Survey data might also shed more light on individual preferences and social media behaviours.

Our estimates are also robust to a single definition of populism. Due to the liquid nature of this phenomenon, it is possible for our figures to change under different conceptions of populism, such as its more ethno-nationalistic departures. Our methodology could then be replicated using different text-bags as shown in Appendix B. In this sense, further work is certainly needed to address competition between party platforms sharing contiguous filter bubbles and investigate whether misinformation can affect voting preferences between populist platforms, favouring certain versions of populism over other ones.

Our work provides new evidence light on the relationship between the spread of fake news and populist echo chambers, contributing to the literature on these topics. Moreover, it introduces a new indicator of populist content based on text mining and applied to the text from social media posts of parties and their leaders during their electoral campaigns. This

represents a new tool applicable to study populism as a phenomenon that eschews the political dimensions of left and right. In its current form, our dictionaries are applicable to all political texts in Italian and German, but could also be translated and extended to other languages.

Finally, as argued earlier, we believe that future research should focus on the relationship between misinformation and echo chambers. As filter bubble access is already determined by prior preferences and individual characteristics, the attention of researchers and policy makers should rather focus on the socio-economic determinants for access into misinformation bubbles. If changes in voting behaviour are unaffected by increases in exposure to fake news, then, we believe, the personalisation of information streams in social media is ultimately the most important and socially poignant factor that should be addressed when studying misinformation and the rise of populism.

Supplementary material

Datasets and codes used in our estimation can be found in the online data archive, available at the url address: https://sites.google.com/site/michelecantarella1992/data-archive-by-paper.

References

Abel, A., Vettori, C., and Forer, D. (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. In for Minority Issues & European Academy Bolzano/Bozen, E. C., editor, *European Yearbook of Minority Issues, vol. 9*, pages 271–304. Brill Academic Publishers, Leiden.

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):139–177.

Albertazzi, D. and McDonnell, D. (2008). *Twenty-First Century Populism. The Spectre of Western European Democracy*. Palgrave Macmillan.

Allcott, H. and Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 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.

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.

Barrera Rodriguez, O., Guriev, S. M., Henry, E., and Zhuravskaya, E. (2017). Facts, alternative facts, and fact checking in times of post-truth politics. Cepr discussion paper no. dp12220, CEPR.

Bischof, D. and Senninger, R. (2018). Simple politics for the people? complexity in campaign messages and political knowledge. *European Journal of Political Research*, 57(2):473–495.

Boelitz, O. (1930). *Das Grenz- und Auslanddeutschtum, seine Geschichte und seine Bedeutung*. Oldenbourg.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430):443–450.

Boutyline, A. and Willer, R. (2016). The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political Psychology*, 38(3):551–569.

Caiani, M. and Graziani, P. R. (2016). Varieties of populism: insights from the italian case. *Italian Political Science Review*, 46(2):243–267.

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*, 16(4):1094–1136.

Chiang, C.-F. and Knight, B. (2011). Media Bias and Influence: Evidence from Newspaper Endorsements. *The Review of Economic Studies*, 78(3):795–820.

D'Alimonte, R. (2019). How the Populists Won in Italy. *Journal of Democracy*, 30(1):114–27.

de Vreese, C. H., Esser, F., Aalberg, T., Reinemann, C., and Stanyer, J. (2018). Populism as an expression of political communication content and style: A new perspective. *International Journal of Press/Politics*, 23(4):423–438.

Decadri, S. and Boussalis, C. (2019). Populism, party membership, and language complexity in the italian chamber of deputies. *Journal of Elections, Public Opinion and Parties*, 0(0):1–20.

DellaVigna, S. and Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting*. *The Quarterly Journal of Economics*, 122(3):1187–1234.

Drago, F., Nannicini, T., and Sobbrio, F. (2014). Meet the Press: How Voters and Politicians Respond to Newspaper Entry and Exit. *American Economic Journal: Applied Economics*, 6(3):159–88.

Durante, R., Pinotti, P., and Tesei, A. (2019). The political legacy of entertainment TV. *American Economic Review*, 109(7):2497–2530.

Ebner, C. V. (2016). The long way to bilingualism: the peculiar case of multilingual south tyrol. *International Journal for 21st Century Education*, 3(2):25.

Falck, O., Gold, R., and Heblich, S. (2014). E-lections: Voting behavior and the internet. *American Economic Review*, 104(7):2238–65.

Galbraith, J. K. and Hale, J. T. (2008). State income inequality and presidential election turnout and outcomes. *Social Science Quarterly*, 89(4):887–901.

Gavazza, A., Nardotto, M., and Valletti, T. (2018). Internet and Politics: Evidence from U.K. Local Elections and Local Government Policies. *The Review of Economic Studies*, 86(5):2092–2135.

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. *The American Political Science Review*, 105(1):135–150.

Gerber, A. S., Karlan, D., and Bergan, D. (2009). Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics*, 1(2):35–52.

Giglietto, F., Iannelli, L., Rossi, L., Valeriani, A., Righetti, N., Carabini, F., Marino, G., Usai, S., and Zurovac, E. (2018). Mapping italian news media political coverage in the lead-up of 2018 general election. *SSRN Electronic Journal*.

Guess, A., Nyhan, B., and Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 us presidential campaign. *European Research Council*.

Guiso, L., Herrera, H., Morelli, M., and Sonno, T. (2017). Populism: Demand and supply. *Center for Economic Policy Research Discussion Paper*, 11871.

Gunther, R., Beck, P. A., and Nisbet, E. C. (2019). "Fake news" and the defection of 2012 Obama voters in the 2016 presidential election. *Electoral Studies*, 61:102030.

Hersh, E. D. and Nall, C. (2015). The primacy of race in the geography of income-based voting: New evidence from public voting records. *American Journal of Political Science*, 60(2):289–303.

Jaber, A. (2013). Individual Campaign Contributions, Electoral Participation, and Broadband Network Availability. Available at ssrn: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2201960, SSRN.

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

Kumar, A. and Sebastian, T. M. (2012). Sentiment analysis on twitter. *International Journal of Computer Science Issues*, 9(4):372–378.

Larcinese, V. and Miner, L. (2012). The Political Impact of the Internet on US Presidential Elections. Working paper, London School of Economics.

Lewis-Beck, M. S. and Nadeau, R. (2011). Economic voting theory: Testing new dimensions. *Electoral Studies*, 30(2):288–294.

Liberini, F., Redoano, M., Russo, A., Cuevas, A., and Cuevas, R. (2020). Politics in the Facebook Era Evidence from the 2016 US Presidential Elections. CESifo Working Paper 8235, CESifo.

Miner, L. (2015). The unintended consequences of internet diffusion: Evidence from malaysia. *Journal of Public Economics*, 132:66 – 78.

Mudde, C. (2004). The Populist Zeitgeist. *Government and Opposition*, 39(4):541–563.

Norris, P. and Inglehart, R. (2019). *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. Cambridge University Press.

Pauwels, T. (2011). Measuring populism: A quantitative text analysis of party literature in belgium. *Journal of Elections, Public Opinion and Parties*, 21(1):97–119.

Polk, J., Rovny, J., Bakker, R., Edwards, E., Hooghe, L., Jolly, S., Koedam, J., Kostelka, F., Marks, G., Schumacher, G., 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. *Research & Politics*, 4(1):2053168016686915.

Poy, S. and Schüller, S. (2020). Internet and voting in the social media era: Evidence from a local broadband policy. *Research Policy*, 49(1):103861.

Rooduijn, M. (2014). The nucleus of populism: In search of the lowest common denominator. *Government and Opposition*, 49(4):572–598.

Rooduijn, M. and Pauwels, T. (2011). Measuring populism: Comparing two methods of content analysis. *West European Politics*, 34(6):1272–1283.

Swire, B., Berinsky, A. J., Lewandowsky, S., and Ecker, U. K. H. (2017). Processing political misinformation: comprehending the trump phenomenon. *Royal Society Open Science*, 4(3):160802.

Taggart, P. A. (2000). *Populism*. Open University Press.

VanKessel, S. (2015). *Populist Parties in Europe. Agents of Discontent?* PalgraveMacMillan, London.

Zimmermann, F. and Kohring, M. (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 Communication*, 37(2):215–237.

A Measuring exposure to fake news

The problem of measuring exposure to fake news is far from a trivial one. Other studies have relied on different methods, such as individual level survey data or the total social media shares for each piece of misinformation. The structure of our data, however, imposes to obtain a proxy for exposure that can be broken down at municipality level and that is available for all units of analysis. There is no way to reconstruct how each fake news has been shared across social media, and who has been exposed to it. It is, however, possible to construct an approximate measure of the following each Facebook page which disseminate fake news has in a given municipality.

This information can only be acquired through a single source: the Facebook Audience Insight Tools (see www.facebook.com/ads/manager/audiences). This tool, generally intended for advertising purposes, allows to access information on active Facebook users, who can be filtered by their place of residence, age, gender, language and many other interests. The tool yields demographic (broken down by age, gender, relationship status, education, and employment) and online activity (page likes and social media use) information on all active individuals in the targeted audience. While it is not possible to filter an audience by their appreciation for a particular Facebook page, it is possible to target an audience through 'interests', and then be presented with a list of pages which correlate with these interests amongst the selected audience, where the number of audience-specific likes for each of these pages is also displayed.

We found a number of interest to be correlated with appreciation of fake news disseminators in Italy. Keywords (English translation in parenthesis) such as 'Immigrazione' (immigration), 'Stato sovrano' (sovereign state), 'Scienza di confine' (border science), 'Illuminati' (the enlightened), 'Medicina alternativa' (alternative medicine), 'Casta' (caste), 'Teoria del complotto' (conspiracy theory), 'Notizie Incredibili' (incredible news), 'Popolo' (the people), 'News24', 'Sovranità' (sovereignty), 'La casta' (the caste), 'Massoneria' (Freemasonry), and 'Notizie.it' (news.it), all returned information on many disseminators included in 'Black lists' compiled by debunking websites.

While the demographic information of the audience tool is biased towards Facebook users, we are confident that page likes which is information obtained from these tools presents an unbiased figure of exposure to misinformation, as it is expected for the the utmost majority of fake news to travel through social media.

We then used the Facebook Audience Insight Tools to extract a sample of pages likes on Facebook, using the 'News24' interest. We collected this information for each Municipality in the Trentino-Alto Adige/Südtirol region, for all users aged 24+, in order to ensure that these individuals had the opportunity to vote at both 2013 and 2018 elections.

The collection process required a few precautions to be taken. First, as the tool will not

display figure on a location if the number of users in that area is below a certain (unspecified) threshold, we collected all exposure information for pairs of municipalities $Trento + Municipality_i$, using Trento (the largest city in Trentino-Alto Adige/Südtirol) as the baseline to be subtracted at the end of the collection. This method ensured that: (1) information could be collected for all locations, and (2) that the tool always returned the same list of pages.

A second issue is generated by the 'interest' targeting options, as the tool only returns information on the number of users belonging to a given 'interest' audience who also liked that page. This means that the number of observed likes through interest targeting will be a fraction – approaching unity, for certain interests – of the total audience, and that exposure figures will be under-estimated. To adjust for this issue, we make two further assumptions: (1) that the total audience of these pages is composed solely of people residing in Italy and (2) that this fraction is the same for both Trentino Alto Adige/Südtirol and Italy. More formally, we assume that $obslikes_{IT}/totlikes_{IT} = obslikes_{TA}/totlikes_{TA}$, where the subscript IT and refer to users located in Italy, TA to those in Trentino Alto Adige/Südtirol. This means that the ratio between observed likes and total likes is equal across the country. Under this assumption, we divide for each page the total number of likes by the audience figures obtained in the Italian territory. We then adjust the shares collected in Trentino Alto Adige/Südtirol by this scalar.

The inability to target specific pages through the audience tool implies that not all fake news disseminators will be captured by this methodology, as some are – perhaps purposely – unreachable through interest targeting.

To account for this issue, we develop a simple yet powerful predictive model that we use to impute the social media following for each of these missing pages, following from the intuition that the number of individuals liking a page in a given municipality will be proportional to the total Facebook likes of said page, holding specific effects from each location, such as its size and the language group, as fixed. If we assume that all page likes follow a similar functional form, we can then estimate the parameters from this function for observed pages, and then use the model to impute the following of unobserved pages. In the model:

$$y_i = \alpha + totlikes_i\beta + municip_i'\gamma + totlikes_i \times municip_i'\delta + collegio_i'\zeta + e_i$$

y_i indicates the total number of Facebook likes each observed page has in each municipality, for $i = \{1, 2, ..., M \times P\}$, where M indicates the total number of municipalities, and P the total number of observed pages. $Totlikes$, instead, indicates the total number of likes each of these pages has on Facebook, while $Municip$ and $Collegio$ are column vectors of dummies for municipality and constituency, respectively. The pages used in our estimation are presented in Table A.1, where observed and modelled pages are labelled as 'donor' and 'recipient'

respectively.[37]

The model allows to predict the number of likes for an observed page in a given location as a function of the total number of likes of said page, allowing for different slopes and intercepts between municipalities.[38] With 3504 observations and a multiple R-squared of 0.9237, the model achieves a satisfying fit and its parameters are extracted and used to impute, for each municipality, the number of people liking the other pages appearing in 'black lists' using the total number of likes of these pages only. After these values have been imputed, they are summed with the observed values in each location to obtain an estimate of the total number of social media likes to fake news disseminators in each municipality.

While it is reasonable to argue that the fake news phenomenon has risen to mainstream attention only in the eve of the 2016 referendum and 2018 elections, it should be noted that a few of these pages already existed before 2013. To account for this issue, we adopt a 'conservative' methodology and, for 2013, exposure to fake news is still computed using pages which existed in the date of the 2013 elections, keeping the total number of likes for each of these pages unaffected. While the results shown in Section 7 make use of these adjusted figures, our estimates are nearly unaffected by the use of a less conservative indicator where exposure to fake news is set at zero for 2013.

It should be remarked that this imputation method is not stochastic, but deterministic. However, as the final variable will consist in the sum of these estimates, we are generally uninterested in correctly simulating within-municipality variation, while we feel that between-municipality variation is a second-order problem considering the good fit of the model. Also, we remark that the purpose of this imputation is to improve the figures of exposure so that the effect of each additional like to a disinformation disseminator can be quantified with more precision. In any case, even in the presence of over or under-estimation of our exposure figures, the sign, and the statistical presence of an effect of fake news on voting should remain unchanged, as the imputation mostly scales the total number of likes in a municipality upwards. Indeed, we constructed another alternative variable for exposure to misinformation using only information from the pages we managed to observe with the Facebook Audience Insight Tool, and our final figures are again unaffected by these changes.

Last but not least, our results are only partially reproducible, as the Facebook API suffers from severe transparency issues that affect the possibility to perfectly replicate this method. It is well possible, then, that exposure figures may vary slightly when collected in a different moment in time. Also, some pages may have changed their name or have been unpublished in the meanwhile, complicating, again, the reproduction of our results. Indeed, many of the pages

[37] A very similar imputation model was also used to correct figures for the 'Chedonna.it' page. As the exposure figures are rounded by the nearest hundred after a page reaches 1,000 likes in a given location, figures for this have been adjusted accordingly using the variation in the other observed pages.

[38] The coefficient vector γ will produce fixed effects specific to the municipality, while the vector of the interaction coefficients δ will allow for different slopes.

reported in Table A.1 were closed by Facebook in May 2019 following a flagging campaing from the non-governmental organisation AVAAZ.[39] In any case, the original data set used in our calculations is made available in the online data archive.

[39]Ibid. 9

Table A.1: Fake news disseminators

Page	likes	followers	founded in	type	Page	likes	followers	founded in	type
Adesso Basta	508971	522664	06-Apr-16	donor	Notizie live	11309	11246	03-Nov-14	recipient
Chedonna.it	1996969	1905609	03-Dec-13	donor	Questa è l'Italia	3982	3987	26-Jan-12	recipient
Citazioni che ispirano	322977	324125	02-Nov-13	donor	Questa è l'Italia di oggi	19674	20961	01-Aug-16	recipient
DiariodelWeb.it	506535	512444	19-Dec-14	donor	Rimani informato	44758	53156	03-Jul-17	recipient
Dimissioni e tutti a casa	784505	749631	19-Aug-13	donor	Rothschild: la Bestia che domina il mondo	19172	19163	21-Jul-11	recipient
Giornale Interattivo	525468	525468	03-Mar-14	donor	Scienza di confine	81788	81060	11-Feb-12	recipient
Italia Patria Mia	375094	369117	29-May-15	donor	Segreto di Stato	6918	6934	07-Dec-16	recipient
Mister Link	436253	404946	11-Apr-11	donor	Se ti fai un'altra birra la prima non s'incazza	205564	204537	11-Nov-12	recipient
Quello che i TG non dicono	369617	377979	23-Oct-16	donor	SocialTv Network (websocialtv)	99181	107729	05-Mar-16	recipient
Silenzi e falsità della stampa italiana	853552	861987	06-Nov-13	donor	Sono senza parole	938033	905287	29-Jun-14	recipient
24H Italia News	171561	173703	07-Dec-16	recipient	Sovranità Monetaria & Debito Pubblico	2128	2158	16-Nov-11	recipient
Aprite gli occhi	148648	149245	08-Feb-11	recipient	Stop alle scie chimiche	32729	32471	19-Oct-09	recipient
Autismo e Vaccini	22391	23016	22-Jul-13	recipient	Ultimissime	2889	2876	29-Apr-14	recipient
Avvistamenti di Creature Mitologiche	181711	179820	20-Mar-12	recipient	Un male tutto italiano, questo regime è una vera piaga sociale	10912	10686	10-Jun-11	recipient
Banda Bassotti	32666	32776	29-Jan-13	recipient	Vaccini Basta	27314	28314	21-Jan-12	recipient
Catena Umana	34464	34372	16-May-14	recipient	Zona grigia	66149	65974	29-Apr-16	recipient
Come i treni a vapore	132131	136164	07-Feb-14	recipient	Stop Invasione	18197	18718	23-Oct-15	recipient
Contro i poteri forti	11441	11591	23-May-15	recipient	Un caffe al giorno	235717	236577	21-Jul-16	recipient
Cose che nessuno ti dirà di nocensura.com	1666222	1612016	16-Aug-10	recipient	Non cielo dicono	153986	160074	25-Jun-16	recipient
CrimeNews (mafiacapitale.info) (Catena Umana 3)	11804	11813	13-Oct-15	recipient	Roby	1425198	1414607	06-Apr-11	recipient
CSSC - Cieli Senza Scie Chimiche	10531	10619	08-Dec-11	recipient	Sardegna Today	5428	5550	04-Sep-18	recipient
Eco(R)esistenza	85473	84418	02-Jun-11	recipient	Dangerous News	5385	5275	07-Jul-13	recipient
Informati (informatitalia)	211461	207818	07-Apr-13	recipient	Il Matto Quotidiano	12313	11952	02-Apr-15	recipient
IO VI Spengo	481267	550151	12-Apr-15	recipient	Imola Oggi	9188	9403	16-Sep-18	recipient
Italia Malata	96781	96935	02-Sep-15	recipient	Italiani Uniti per la Patria	96382	98380	28-Apr-16	recipient
Killuminati Soldiers	2406	2434	23-Jan-16	recipient	Scenari Economici	39224	42555	06-Mar-13	recipient
John Koenig Mb Koenig	2585	2670	16-Feb-14	recipient	Informare X Resistere	1086691	1020283	16-Jul-09	recipient
L'alimentazione e gli illuminati	10032	10095	10-Aug-11	recipient	Piovegovernoladro	55839	54235	19-Dec-13	recipient
L'antipolitica	14517	14519	29-Apr-12	recipient	Jeda News	372364	357423	02-Jan-12	recipient
La Vera Italia	3742	3722	16-May-15	recipient	Breaknotizie	81656	80762	10-Sep-12	recipient
La Verità ci Rende Liberi	68857	72846	20-Jun-18	recipient	Controinformo	27095	26917	13-Jan-16	recipient
Le notizie che non ti aspetti	133697	133185	26-Jun-14	recipient	DirettaNews24	130449	130453	09-Oct-16	recipient
Mafia Capitales (whatsappecultura)	99487	98732	30-Dec-08	recipient	FascinAzione	1866	1910	03-Apr-16	recipient
Movimento anti NWO	12324	12274	20-Feb-12	recipient	TG 5 Stelle	16934	16216	26-Jun-13	recipient
Neovitruvian	6742	6763	09-Dec-10	recipient	CVDiariodelPollineo	1569	1572	14-Oct-16	recipient
NO alla dittatoriale Unione Europea e al Nuovo Ordine Mondiale	2629	2714	11-Nov-11	recipient	Notizie in Movimento	116951	118686	29-Sep-14	recipient
Notiziario (notiziario.face)	22882	22943	15-Nov-12	recipient					

*N*otes: Donor and recipient pages used in the estimation of total page likes by municipality. Previous page names are shown in parentheses.

B Alternative definitions of populism

Table B.2: OLS estimates of the effect of misinformation on populist vote using the anti-establishment text bag

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Exposed to Fake News	1.886***	0.236***	0.216***	-0.017
	(0.228)	(0.072)	(0.072)	(0.087)
Exposed to Fake News × Year of election	-1.065***	-0.066	-0.055	0.148**
	(0.230)	(0.047)	(0.047)	(0.062)
Year of election	0.023**	-0.002	0.003	0.031***
	(0.011)	(0.004)	(0.005)	(0.004)
Broadband connection		0.834***	0.848***	0.466***
		(0.037)	(0.038)	(0.080)
Income per capita (natural log)			-0.017***	-0.021***
			(0.005)	(0.005)
Electorate size				0.760***
				(0.114)
Abstentions and invalid votes				-0.321***
				(0.057)
Observations	584	584	584	584
Adjusted R-squared	0.906	0.988	0.988	0.991

*N*otes: OLS estimates for the effect of misinformation on populist vote. Populist scores computed using the 'Anti-establishment/aggressive' text bag. Standard errors robust to clustering by municipality in parentheses.

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

Table B.3: 2SLS estimates of the effect of misinformation on populist vote using the anti-establishment text bag

	(1)	(2)	(3)	(4)
	DiD	First Stages		2SLS
VARIABLES		Exposure	Interaction	
Italian speaking voters	0.309***	0.363**	-0.025	
	(0.052)	(0.149)	(0.068)	
Italian speaking voters × Year of election	0.104***	0.595***	0.965***	
	(0.007)	(0.079)	(0.120)	
Year of election	0.007***	-0.045***	-0.076***	0.013***
	(0.003)	(0.007)	(0.011)	(0.004)
Broadband connection	0.212***	0.067	0.044	0.174
	(0.073)	(0.246)	(0.143)	(0.212)
Electorate size	0.498***	0.069	-0.062	0.415***
	(0.043)	(0.168)	(0.171)	(0.082)
Income per capita (natural log)	-0.003	-0.009	-0.003	0.004
	(0.003)	(0.011)	(0.008)	(0.010)
Abstentions and invalid votes	-0.088**	-0.039	0.056	-0.033
	(0.044)	(0.120)	(0.140)	(0.090)
Exposed to Fake News				0.824**
				(0.363)
Exposed to Fake News × Year of election				-0.400*
				(0.216)
Observations	584	584	584	584
Adjusted R-squared	0.998	0.931	0.930	0.978
Partial R-squared		0.342	0.397	
F-Test		45.03	59.19	

Notes: IV estimates (including reduced form - DiD - and first stages) for the effect of misinformation on populist vote. Populist scores computed using the 'Anti-establishment/aggressive' text bag. F-tests for excluded instruments for the individual instrument (voters in the Italian-speaking language group) and its interaction with year of election are reported as F-Test (exposition) and F-Test (interaction), respectively. Standard errors robust to clustering by municipality in parentheses.

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

Table B.4: OLS and 2SLS estimates of the effect of misinformation on votes to specific parties

	(1)	(2)	(3)	(4)	(5)	(6)
	Lega/M5s		Anti-est./Abst.		Lega	
VARIABLES	OLS	2SLS	OLS	2SLS	OLS	2SLS
Exposed to fake news	0.102	1.127***	-0.628***	-0.232*	-0.181	0.241
	(0.147)	(0.417)	(0.111)	(0.137)	(0.122)	(0.184)
Exposed to fake news × Year of election	0.176*	-0.449**	0.791***	0.571***	0.332***	0.045
	(0.105)	(0.224)	(0.113)	(0.086)	(0.098)	(0.140)
Year of election	177.411***	168.265***	256.740***	244.125***	142.678***	117.182***
	(26.625)	(23.305)	(26.422)	(23.083)	(15.375)	(17.464)
Broadband connections	0.154***	0.021	0.080**	0.020	-0.056**	-0.078**
	(0.047)	(0.087)	(0.035)	(0.037)	(0.028)	(0.035)
Electorate size	0.023	-0.000	0.435***	0.429***	-0.023	-0.123*
	(0.047)	(0.053)	(0.031)	(0.028)	(0.034)	(0.072)
Income per capita (natural log)	-390.447**	271.342	-184.842**	92.679	-41.254	91.579
	(153.846)	(200.852)	(84.399)	(90.996)	(61.455)	(78.921)
Votes to M5s					0.764***	0.789***
					(0.121)	(0.215)
Votes to DF/Abstentions					0.085	0.319*
					(0.092)	(0.182)
Observations	584	584	584	584	584	584
Adjusted R-squared	0.957	0.909	0.994	0.992	0.946	0.920
F-Test (Exposition)		307.3		307.3		17.34
F-Test (Interaction)		1158		1158		16.44

*N*otes: Robust standard errors in parentheses. IV estimates (including reduced form - DiD - and first stages) for the effect of misinformation on vote to specific parties. F-tests for excluded instruments for the individual instrument and its interaction with year of election are reported as F-Test (instrument) and F-Test (interaction), respectively.

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

Table B.5: OLS and 2SLS estimates of the effect of misinformation on populist vote using different indicators of populism, unstandardised variables

VARIABLES	(1) OLS PL	(2) OLS vK	(3) OLS N&I	(4) 2SLS PL	(5) 2SLS vK	(6) 2SLS N&I
Exposed to Fake News	-1.123***	-1.128***	-1.135***	-0.131	-0.120	-0.088
	(0.335)	(0.336)	(0.343)	(0.169)	(0.171)	(0.181)
Exposed to Fake News × Year of election	1.597***	1.601***	1.617***	1.048***	1.042***	1.034***
	(0.251)	(0.252)	(0.256)	(0.152)	(0.149)	(0.142)
Year of election	0.109***	0.113***	0.116***	0.092***	0.095***	0.098***
	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)
Broadband connections	0.505**	0.496**	0.485**	0.057	0.042	0.016
	(0.208)	(0.207)	(0.207)	(0.159)	(0.160)	(0.165)
Income per capita (natural log)	-0.034***	-0.033***	-0.033***	0.006	0.007	0.008
	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)
Electorate size	0.018	0.040	0.103	-0.182	-0.166	-0.117
	(0.224)	(0.224)	(0.222)	(0.227)	(0.225)	(0.219)
Abstentions and invalid votes	0.063	0.050	-0.009	0.228	0.219	0.173
	(0.177)	(0.176)	(0.171)	(0.210)	(0.210)	(0.202)
Observations	584	584	584	584	584	584
Adjusted R-squared	0.939	0.938	0.937	0.910	0.908	0.905

Notes: Robust standard errors in parentheses. The dependent variable is composed by the sum of votes of those parties coded as populist according to three sources: the PopuList database (PL), the work of VanKessel (2015) (vK) and the work of Norris and Inglehart (2019) (N&I).

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

C Correlation between text-based populist scores and CHES data

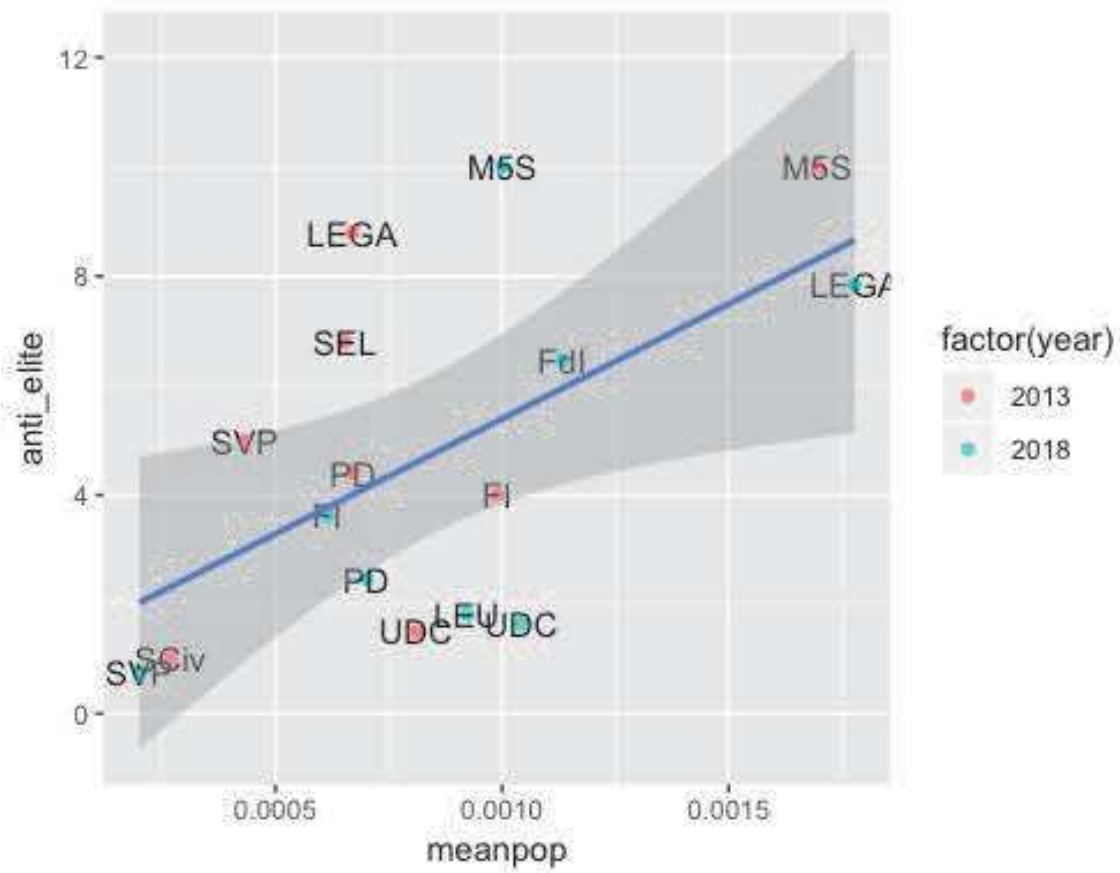

Figure C.1: Text analysis scores of social media posts from parties and their leaders during the 2013 and 2018 elections campaigns (x-axis) and their relationship with the scores on the variable 'People vs the Elites' from the Chapel Hill Expert Survey for year 2014 (y-axis). Higher values on the y-axis correspond to higher salience of anti-establishment and anti-elite rhetoric on a scale from 0 to 10. Scores are computed using the 'Anti-establishment/ aggressive' text bag.

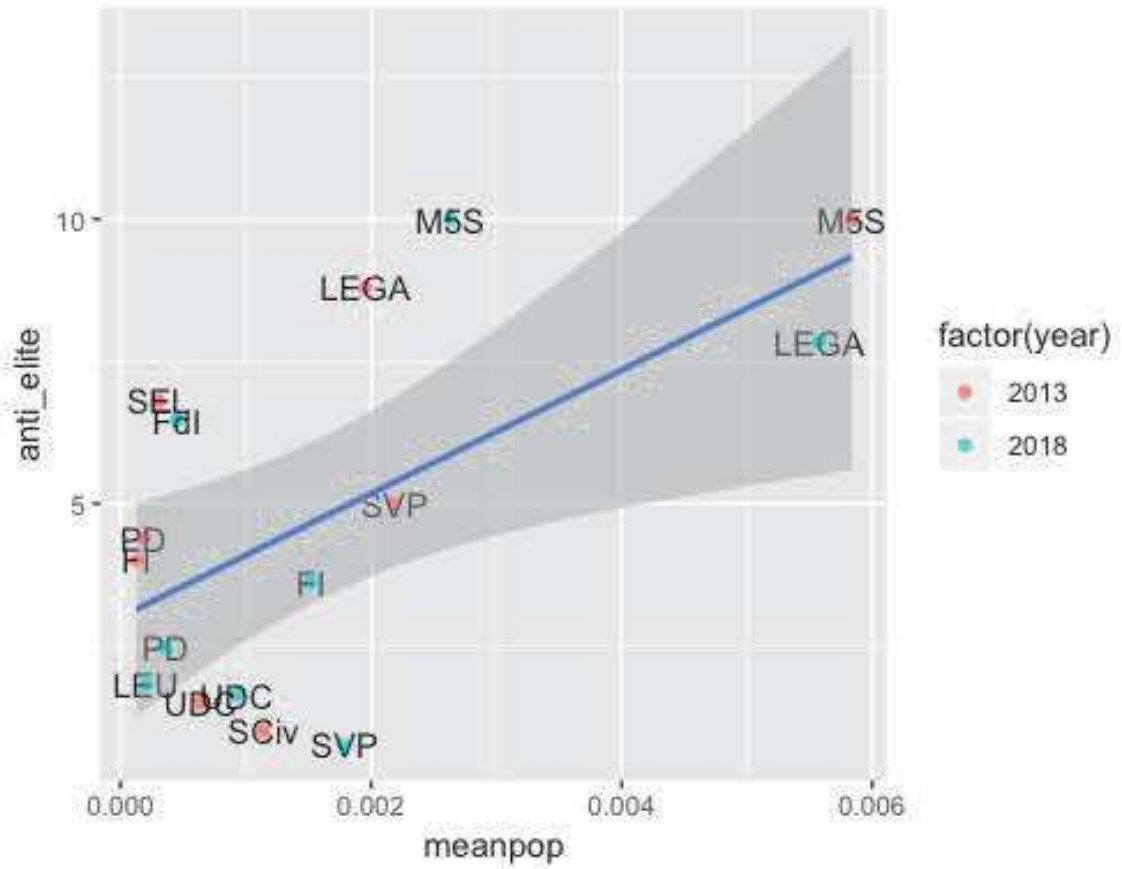

Figure C.2: Text analysis scores of social media posts from parties and their leaders during the 2013 and 2018 elections campaigns (x-axis) and their relationship with the scores on the variable 'People vs the Elites' from the Chapel Hill Expert Survey for year 2014 (y-axis). Higher values on the y-axis correspond to higher salience of anti-establishment and anti-elite rhetoric on a scale from 0 to 10. Scores are computed using the 'Assertive' text bag.

D Results with non-standardised coefficients

Table D.6: OLS estimates of the effect of misinformation on populist vote, unstandardised variables

	(1)	(2)	(3)	(4)
VARIABLES	OLS	OLS	OLS	OLS
Exposed to fake news	0.011***	-0.002***	-0.001***	-0.002***
	(0.001)	(0.000)	(0.000)	(0.000)
Exposed to fake news × Year of election	-0.006***	0.002***	0.002***	0.002***
	(0.001)	(0.000)	(0.000)	(0.000)
Year of election	1.042***	0.182*	0.143	0.936***
	(0.093)	(0.104)	(0.114)	(0.099)
Broadband connections		0.002***	0.002***	0.000*
		(0.000)	(0.000)	(0.000)
Income per capita (natural log)			0.492	-0.533*
			(0.406)	(0.288)
Electorate size				0.002***
				(0.000)
Abstentions and invalid votes				-0.001***
				(0.000)
Observations	584	584	584	584
Adjusted R-squared	0.868	0.990	0.990	0.996

*N*otes: 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

Table D.7: 2SLS estimates of the effect of misinformation on populist vote, unstandardised variables

	(1)	(2)	(3)	(4)
	DiD	First Stages		2SLS
VARIABLES		Exposure	Interaction	
Italian speaking voters	-0.000	0.131**	-0.009	
	(0.000)	(0.054)	(0.023)	
Italian speaking voters × Year of election	0.000***	0.297***	0.455***	
	(0.000)	(0.039)	(0.057)	
Year of election	0.434***	-207.135***	-332.732***	0.778***
	(0.069)	(33.100)	(46.819)	(0.098)
Broadband connections	-0.000	0.021	0.013	-0.000
	(0.000)	(0.079)	(0.043)	(0.000)
Electorate size	0.002***	0.022	-0.019	0.002***
	(0.000)	(0.053)	(0.051)	(0.000)
Income per capita (natural log)	0.495***	-143.422	-37.960	0.491
	(0.191)	(171.636)	(124.660)	(0.299)
Abstentions and invalid votes	-0.000	-0.047	0.063	-0.000
	(0.000)	(0.143)	(0.158)	(0.000)
Exposed to fake news				-0.000
				(0.001)
Exposed to fake news × Year of election				0.001***
				(0.000)
Observations	584	584	584	584
Adjusted R-squared	0.997	0.931	0.930	0.994
Partial R-squared		0.342	0.397	
F-Test		45.03	59.19	

Notes: IV estimates (including reduced form - DiD - and first stages) for the effect of misinformation on populist vote. Populist scores computed using the 'Assertive' text bag. F-tests for excluded instruments for the individual instrument (voters in the Italian-speaking language group) and its interaction with year of election are reported as F-Test (exposition) and F-Test (interaction), respectively. Standard errors robust to clustering by municipality in parentheses.

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

RECENT PUBLICATIONS BY *CEIS Tor Vergata*

Peaks, Gaps, and Time Reversibility of Economic Time Series
Tommaso Proietti
CEIS Research Paper, 492, June 2020

Microdata for Macro Models: the Distributional Effects of Monetary Policy.
Luisa Corrado and Daniela Fantozzi
CEIS Research Paper, 491, June 2020

Suboptimality of probability matching - A formal proof, a graphical analysis and an impulse balance interpretation
Vittorio Larocca and Luca Panaccione
CEIS Research Paper, 490, June 2020

Nowcasting GDP and its Components in a Data-rich Environment: the Merits of the Indirect Approach
Alessandro Giovannelli, Tommaso Proietti, Ambra Citton, Ottavio Ricchi, Cristian Tegami and Cristina Tinti
CEIS Research Paper, 489, May 2020

Elementary Facts About Immigration in Italy. What Do We Know About Immigration and Its Impact
Rama Dasi Mariani, Alessandra Pasquini and Furio Camillo Rosati
CEIS Research Paper, 488, May 2020

Financial Fragility across Europe and the US: The Role of Portfolio Choices, Household Features and Economic-institutional Setup
Marianna Brunetti, Elena Giarda and Costanza Torricelli
CEIS Research Paper, 487, May 2020

Competition Between Infinitely Many Fund Managers and Investor's Welfare
Enrico Lupi
CEIS Research Paper, 486, May 2020

Optimal Factor Taxation in A Scale Free Model of Vertical Innovation
Barbara Annicchiarico, Valentina Antonaroli and Alessandra Pelloni
CEIS Research Paper, 485, May 2020

Congestion and Incentives in the Age of Driverless Cars
Federico Boffa, Alessandro Fedele and Alberto Iozzi
CEIS Research Paper, 484, May 2020

Complexity and Growth
Alberto Bucci, Lorenzo Carbonari, Pedro Mazeda Gil and Giovanni Trovato
CEIS Research Paper, 483, May 2020

DISTRIBUTION

Our publications are available online at www.ceistorvergata.it

DISCLAIMER

The opinions expressed in these publications are the authors' alone and therefore do not necessarily reflect the opinions of the supporters, staff, or boards of CEIS Tor Vergata.

COPYRIGHT

Copyright © 2020 by authors. All rights reserved. No part of this publication may be reproduced in any manner whatsoever without written permission except in the case of brief passages quoted in critical articles and reviews.

MEDIA INQUIRIES AND INFORMATION

For media inquiries, please contact Barbara Piazzi at +39 06 72595652/01 or by e-mail at piazzi@ceis.uniroma2.it. Our web site, www.ceistorvergata.it, contains more information about Center's events, publications, and staff.

DEVELOPMENT AND SUPPORT

For information about contributing to CEIS Tor Vergata, please contact at +39 06 72595601 or by e-mail at segr.ceis@economia.uniroma2.it