Journal of Urban Economics 141 (2024) 103613

Contents lists available at ScienceDirect

Journal of Urban Economics

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

Spatial and historical drivers of fake news diffusion: Evidence from anti-Muslim discrimination in India

Samira S. Abraham [a,1], Gianandrea Lanzara [b,1], Sara Lazzaroni [a,1], Paolo Masella [a,1], Mara P. Squicciarini [b,*,1,2]

[a] University of Bologna, Italy
[b] Bocconi University, Italy

ARTICLE INFO

JEL classification:
J15
Z12

Keywords:
Discrimination
Fake news
Religion
Covid
India

ABSTRACT

What drives the propagation of discriminatory fake news? To answer this question, this paper focuses on India at the onset of the COVID-19 pandemic: on March 30, a Muslim convention (the *Tablighi Jamaat*) in New Delhi became publicly recognized as a COVID hotspot. Using Twitter data, we build a comprehensive novel dataset of georeferenced tweets to identify anti-Muslim fake news. First, we document that fake news about Muslims intentionally spreading the virus spiked after March 30. Then, we investigate the geographical and historical determinants of the spread of fake news in a difference-in-difference setting. We find that the diffusion of anti-Muslim false stories was more pronounced (i) in districts closer to New Delhi, suggesting that fake news spread spatially; and (ii) in districts exposed to historical attacks by Muslim groups, suggesting that the propensity to disseminate fake news has deep-rooted historical origins.

"It Was Already Dangerous to Be Muslim in India. Then Came the Coronavirus" (Perrigo, 2020).

1. Introduction

Fake news, sometimes referred to as false stories, may likely have widespread, long-lasting political, social, and economic consequences for societies (Zhuravskaya et al., 2020). In today's interconnected world through media and social platforms, fake news may spread even more rapidly and reach larger audiences than factual news, thereby distorting the opinions of those who receive it.[3] At the same time, the propensity to believe and propagate false stories is likely to vary across space depending on a range of local factors, some of which may have deep historical roots.

Of particular concern is the use of fake news as a tool to target vulnerable groups, including ethnic and cultural minorities. However, while scholars have extensively focused on the role of misinformation in political contexts, there has been limited study of fake news as a vehicle for disseminating hate and discriminatory sentiments towards minorities. As a result, we still have a limited understanding of the factors that enhance the diffusion of discriminatory false stories. In a word of mass global communication, do false stories still spread spatially? Are certain people primed to disseminate misinformation because of local events that happened far back in history?

Answering these questions presents several empirical challenges. It is necessary to have measures of fake news at a fine geographical level and, more generally, to find a suitable empirical context for studying their spatial diffusion. To partially bypass these challenges, this paper focuses on the social media platform Twitter and analyzes the spread of anti-Muslim fake news in India at the onset of the coronavirus pandemic, in 2020.[4] In particular, we evaluate two main hypotheses,

* Corresponding author.
E-mail address: mara.squicciarini@unibocconi.it (M.P. Squicciarini).

[1] Contributed equally to the manuscript.

[2] We thank Mathieu Couttenier, Sophie Hatte, and Paolo Vanin for helpful comments. We thank James Fenske for sharing data on the location of precolonial conflicts in India. This paper has been supported by the Italian Ministerial grant PRIN 2017 – Linea B "Religious and Racial Discrimination Attitudes: Evidence from a Contemporary and a Historical Context", Prot. 2017ATLJHB_001 (CUP J34I19003780001) and by the Fondazione Romeo ed Enrica Invernizzi – Bocconi Junior Researchers' Grant, Italy.

[3] Sunstein (2018) suggests that social media may threaten democracy through the diffusion of false information. More formally, using data covering about 126,000 stories posted by roughly 3 million people on Twitter, Vosoughi et al. (2018) study the diffusion of true and false stories during the 2006–2017 period. The authors show that false stories spread faster and more broadly than true stories across several categories of information, but especially in the case of political news. By contrast, focusing on the context of the 2016 American Presidential elections, Guess et al. (2018) and Grinberg et al. (2019) find that fake news were shared by only a small fraction of social media users.

[4] In line with Guess et al. (2019), we define fake news as "false or misleading content intentionally dressed up to look like news articles".

https://doi.org/10.1016/j.jue.2023.103613
Received 31 May 2022; Received in revised form 7 November 2023
Available online 7 December 2023

i.e, whether geographical and historical factors determine the diffusion of discriminatory false stories.

First, we explore whether fake news spread spatially. Previous research in urban economics has shown that idea and knowledge flows, as measured through patent citations, decline with geographic distance (Jaffe et al., 1993; Henderson et al., 2005; Thompson and Fox-Kean, 2005; Figueiredo et al., 2015). In addition, Bailey et al. (2018) have uncovered a similar distance decay for social connections on Facebook. By drawing upon these findings, we propose and test the hypothesis that discriminatory fake news on Twitter spreads spatially, that is, its diffusion is stronger in locations that are closer to its origin. To the best of our knowledge, we are the first to document spatial spillover effects for viral ideas concerning discriminatory false stories.

Second, we investigate whether the spreading of fake news has historical origins. A large literature in economics has documented the persistent effects of historical events on contemporary attitudes and behaviors (for a review, see Voth, 2021). For instance, Voigtländer and Voth (2012) show that Jewish pogroms during the Black Death strongly predict antisemitic violence in Nazi Germany and votes for the Nazi party. Similarly, Ochsner and Roesel (2019) find that Austrian municipalities attacked by Turkish troops in early modern times displayed stronger anti-Muslim sentiment and cast more votes for the far-right party, after it began to recall past Turkish atrocities during its 2005 campaign. By drawing upon these findings, we propose and test the hypothesis that local events occurred in the past influence the propensity to disseminate discriminatory false stories. To the best of our knowledge, we are the first to document that historical events may prime certain populations to disseminate misinformation.

India is an ideal setting to test these hypotheses for three main reasons. First, India is an extremely diverse country with a history of tensions among religious communities, and its Muslim minority has been one of the communities most discriminated against.[5] Second, discussion about fake news and misinformation in India has grown louder in recent years, following the explosive growth in the number of active Internet users—Nielsen (2022) estimates that India had 646 million active Internet users in December 2021, and Internet penetration continues to increase in rural and in urban areas.[6] Third, at the onset of the coronavirus pandemic, a Muslim religious gathering – the *Tablighi Jamaat* convention in New Delhi – was identified as a national COVID-19 hotspot: suddenly, about 9000 congregants and related primary contacts were sent to quarantine facilities or hospitals in a country that had seen only a handful of reported cases beforehand (Times of India, 2020). These events led part of the public opinion to establish a connection between the Muslim minority and the diffusion of COVID-19 in India.

In this context, we analyze a comprehensive novel dataset of georeferenced tweets covering the period from December 1, 2019, to April 30, 2020. To identify tweets containing discriminatory false stories against Muslims, we compiled a list of keywords based on the false stories reported by the popular Indian fact-checking website AltNews in March and April 2020. The vast majority of false stories in our sample claimed that Muslim individuals were deliberately infecting other people, often drawing an analogy between the spread of the virus and the spread of religion, and asserting that these allegedly intentional infections were a form of jihad.

Our empirical strategy exploits the timing of a sequence of events tied to the aforementioned religious gathering. In fact, though the Muslim convention started already from the beginning of March, its alleged connection with the pandemic did not become salient until the evening of March 30, when multiple deadly COVID-19 cases were reported among *Tablighi* participants and when several institutional actors blamed the convention for spreading COVID in India. The hashtag #coronajihad began trending strongly on March 31, 2020. For ease of discussion, we will henceforth refer to this bundle of events on March 30 as the "*Tablighi* shock".

In line with the literature on the scapegoating of minorities during epidemic outbreaks (Jedwab et al., 2019), we document a large spike in anti-Muslim fake news on March 31: the share of tweets reporting anti-Muslim false stories – around 0.1% of total tweets in the week preceding the *Tablighi* shock – jumped to 2.8% on March 31, then peaked at 3.3% a day later.[7]

We then exploit the *Tablighi* shock and, using a difference-in-difference estimation strategy at the daily district level, we test our two hypotheses on the influence of geography and history on the diffusion of anti-Muslim discriminatory fake news. First, in line with our first hypothesis, we find that the spread of fake news is more pronounced in districts located closer to New Delhi, where the *Tablighi* event took place. Specifically, our difference-in-difference coefficient suggests that moving one standard deviation closer to New Delhi (roughly 558 km) increased the daily number of tweets with anti-Muslim fake news by 1.09 tweets. This is a sizeable difference, compared to the daily district average number of fake news tweets in the week after the shock (3.38)—suggesting that anti-Muslim fake news propagated in space from New Delhi following the *Tablighi* shock. This result is robust to considering alternative measures of proximity to New Delhi, to accounting for spatial autocorrelation, and to the use of different definitions of the dependent variable (such as the share of fake news tweets over the total number of tweets). We then explore the mechanisms behind this result. We find suggestive evidence that the spatial diffusion from New Delhi is partly due to personal ties being stronger in districts closer to New Delhi. By contrast, the effect of distance to New Delhi is only slightly reduced when we account for the intensity of Twitter connections although we still find that more intense Twitter connections are associated with a more pronounced spread of fake news. All in all, this evidence indicates that physical distance is a key determinant of the diffusion of fake news on Twitter.

Next, we test our second hypothesis and investigate whether the observed differences in the diffusion of fake news can be traced back to deeply rooted characteristics of Indian society. To do so, we exploit a comprehensive dataset of historical conflicts on the Indian subcontinent; we focus on land-based attacks initiated by Muslim entities in precolonial times (1000–1757). For the sake of brevity, we will henceforth refer to these events as "Muslim attacks". We find that the spread of fake news is more pronounced in districts that were exposed to Muslim attacks. Specifically, our difference-in-difference coefficient suggests that the historical occurrence of a Muslim attack increased the daily number of tweets with anti-Muslim fake news by 1 tweet, compared to districts that never experienced any attack. These results suggest that the historical experience of conflict against Muslims predicts current anti-Muslim discriminatory behavior, as captured by the diffusion of anti-Muslim fake news.

In addition, our findings are robust to alternative definitions of exposure to Muslim attacks and to controlling for other potentially confounding factors (interacted by a *Post-March30* dummy), including further geographical controls, historical state capacity, local exposure to colonizers, measures of linguistic and ethnic fractionalization, and availability of transportation routes.

[5] According to a 2019–2020 Pew Research Center survey of religion across India, one in five Muslims say they have recently faced religious discrimination (Pew Research Center, 2021).

[6] According to the 2019 Reuters Institute survey about digital news consumption in India, 57% of the respondents were concerned whether the online news they came across was real or fake (Aneez et al., 2019). Moreover, according to the National Crime Records Bureau, the number of cases filed in India against people who "circulate fake/false news/rumors" rose from 486 in 2019 to 1527 cases in 2020, that is by a sharp 214 percent (The Quint, 2021).

[7] In 2019, Equity Labs conducted an analysis of over 1000 hate speech posts in six different languages on Facebook in India. They found that almost 40% of the hate speech and disinformation fell under the category of "Islamophobia" (Soundararajan et al., 2019).

Literature. This paper is related to four strands of literature.

First, our finding that discriminatory false stories spread spatially contributes to the urban literature on the diffusion of ideas in geographic space. A number of papers have used patent citations to measure knowledge spillovers across U.S. counties, finding that physical distance hampers knowledge diffusion (Jaffe et al., 1993; Henderson et al., 2005; Thompson and Fox-Kean, 2005; Figueiredo et al., 2015). Building on these studies, Bailey et al. (2018) provide evidence that also social media connections play a role in facilitating the spread of patent citations. Similarly, Comin et al. (2012) show that technology adoption follows a negative distance gradient from adoption leaders. With respect to these studies, we focus on the diffusion of ideas of a different kind, i.e. harmful ideas such as discriminatory false stories, and exploit the timing and location of a health shock, i.e. a COVID hotspot in India, to measure the spatial diffusion of anti-Muslim fake news. Moreover, our assessment of the role played by Twitter connections as a potential driver of fake news diffusion contributes to a body of research in urban economics studying the spatial consequences of new communication technologies, and in particular of social media.[8]

Second, our finding that past Muslim attacks shape the diffusion of anti-Muslim fake news today contributes to the literature on the persistent effects of historical events, institutions, norms, and values on various contemporary outcomes (for a review, see Voth, 2021). In particular, a number of studies have documented that discrimination against minorities has deep-rooted origins (see, for instance, Voigtländer and Voth (2012) on antisemitic violence in Germany, and Ochsner and Roesel (2019) on anti-Muslim sentiments in Austria). Zooming into the Indian context, Jha (2013) shows that stronger medieval Hindu–Muslim trade relationships in port locations is associated with fewer Hindu–Muslim riots from 1850 to 1995. We contribute to this literature by showing that the proliferation of anti-Muslim fake news at the onset of the coronavirus pandemic was greater in districts with a history of precolonial Muslim attacks. In so doing, we also contribute to the literature analyzing precolonial determinants of present-day outcomes in India, a context in which research has largely been focused on the legacy of colonialism (see, e.g., Banerjee and Iyer (2005), Iyer (2010), Castelló-Climent et al. (2018), Bharadwaj and Mirza (2019) and Chaudhary et al. (2020)).

Third, by focusing on the rise of anti-Muslim fake news after the coronavirus outbreak, this paper relates to a recent literature in economics showing that anti-minority behavior increases during economic and epidemic crises. In the context of economic crises, Doerr et al. (2021) show an increase in anti-Jewish persecutions after Germany's 1931 banking crisis, while Anderson et al. (2017, 2020) show, respectively, an increase in anti-Jewish persecutions after colder growing seasons in the 1000–1800 period and an increase in anti-Black discrimination during the Great Recession (2007–2009). In the context of epidemics, Jedwab et al. (2019) find that the Black Death of the 1350s triggered anti-Jewish persecutions in towns where people believed antisemitic allegations but not in towns where the activities of Jews were complementary to the local economy. Bartoš et al. (2021), Dipoppa et al. (2021), Lu and Sheng (2022), and Lanzara et al. (2023) show that discrimination against the Chinese minority has increased after the coronavirus pandemic. We contribute to this literature by focusing on the dissemination of false stories as a specific means to target minorities.

Fourth, and more broadly, this paper relates to the literature studying the effects of media on discrimination. Regarding traditional media, DellaVigna et al. (2014), Yanagizawa-Drott (2014), Adena et al. (2015), and Couttenier et al. (2021) show that propaganda and distorted news coverage can contribute to ethnic violence against immigrant minorities and increase votes for far-right parties. Regarding digital media, Müller and Schwarz (2020, 2021) show, respectively, how Twitter, Facebook and other social media can activate hatred of minorities in the contexts of Donald Trump's political rise in the United States and of the refugee crisis in Germany. We contribute to this literature by investigating the determinants of diffusion of fake news that targets a minority, a topic rather unexplored by the literature.

2. Background

2.1. COVID in India and fake news against the Muslim minority

2.1.1. COVID and the Tablighi Jamaat

COVID-19 was first identified in December 2019 in the Chinese city of Wuhan, where a major local outbreak quickly escalated into a global public health emergency. In India, the first cases of contagions were reported in late January 2020, but they were identified and contained in hopes of avoiding a mass outbreak. Despite its geographical proximity to China, India had close to no cases in February; by March, however, daily cases of contagions were being reported. As the situation worsened, regional and national authorities enacted several restrictive policies that canceled domestic and international flights, suspended railway services, created social distancing measures, restricted public gatherings and dine-in restaurants, closed nonessential businesses, and made it compulsory to wear a mask. These policies culminated in a nationwide lockdown that began on March 25.

An Islamic missionary movement called the *Tablighi Jamaat* had scheduled a religious congregation in March 2020 at the *Nizamuddin Markaz* mosque in the Indian capital city of New Delhi. Thousands of worshipers from across India and abroad poured into the *Markaz* in the first weeks of March. From March 13, the Delhi state government began issuing public notices to avoid gatherings. Direct letters were sent to the *Tablighi Jamaat* asking them to disassemble the congregation. The *Tablighi Jamaat* responded that it had suspended all activities and had managed to send home some of the attendees, but that it was having difficulty making arrangements for attendees who were still at the mosque and now stranded by the railway closures. The Delhi state government announced a weeklong shutdown of the capital city beginning March 23, further restricting the worshipers from road travel. On March 29, the district police and health officials started sending some of the worshipers to hospitals and quarantine facilities after learning that many of them had tested positive for COVID (Outlook, 2020).

The back and forth between local officials and the *Tablighi Jamaat*, as well as the displacement of attendees and the increasing number of COVID cases from the *Markaz* led to a blame game between the *Tablighi Jamaat* and state officials. By the end of March, social media and news channels were filled with discussions on the government's late response in shutting down the event and on the *Tablighi Jamaat*'s irresponsibility in going through with it. Fake news and discriminatory hashtags blaming Muslims for spreading coronavirus on purpose—including videos showing Muslims licking utensils and a Muslim vendor spitting on fruit to spread the virus—started popping up.

[8] Ioannides et al. (2008) discuss the repercussions of improvements in ICT (Information and Communication Technologies) on agglomeration effects and on the dispersion of economic activity across cities. More closely to our work, Gaspar and Glaeser (1998) study how new communication technologies (e.g., phones, computers, and social media) affect in-person interactions, leading to a change in the size of cities, considered as places where face-to-face interactions are more likely to occur. Several recent studies exploring patterns of connectedness through email and social media also emphasize that the strength of internet relationships negatively correlates with physical distance. For instance, Buchel and Ehrlich (2020) find that travel distance is negatively related to social interactions by mobile phone, pointing again to complementarities between electronic and face-to-face interactions. See also Tranos and Nijkamp (2013) who focus on distance and IP links, Levy and Goldenberg (2014) and Mok et al. (2010) who report evidence using data on email messages, with Mok et al. (2010) showing evidence also exploiting phone data. Finally, Levy and Goldenberg (2014) and Bailey et al. (2018) explore online friendship in Facebook and Takhteyev et al. (2012) and Scellato et al. (2010) look at the structure of Twitter links.

2.1.2. The evening of March 30

On the evening of March 30, several events surrounding the *Nizamuddin Markaz* transformed the once-sporadic sharing of fake news and Islamophobic tweets into a full-blown crisis. First, the Delhi Police cordoned off the entire area around the *Nizamuddin Markaz*. Drones were also deployed to scan the streets in the area for lockdown violators. Then, the Delhi government announced its intention to file a case against the Maulana, the religious head of the *Tablighi Jamaat* (Press Trust of India, 2020), and two politicians from the ruling party of the Central Indian government tweeted that the rise in COVID cases in India was linked to the *Nizamuddin Markaz*. Late that night, the Telangana State Chief Minister's office tweeted that six *Markaz* attendees had succumbed to the coronavirus and died in Telangana. Even if a *Tablighi Jamaat* attendee had died from COVID the week before, this was the first time multiple deaths were linked to the *Tablighi* event.

Though discriminatory anti-Muslim hashtags and fake news had been shared before (see Appendix Table A1), March 30 marked a watershed moment. The next day, #coronajihad topped Twitter trends in India, and from then on, the floodgates opened for widespread misinformation and fake news about Muslims deliberately spreading coronavirus (Ritika Jain, Article14, 2020).

On Twitter, hashtags such as #Nizammudin or #NizammudinMarkaz – referring to the mosque – went from being tweeted 10,200 times at 6 p.m. on March 30 to more than 114,000 times at 10 p.m. on March 31. Similarly, the hashtag #TablighiJamaat had been tweeted about 51,100 times by the end of March 31, with blame extending to all Muslims. This escalation peaked two days later, when members of India's central government blamed the *Tablighi Jamaat* for the sudden spike in COVID cases in India (The Week, 2020). Hashtags such as #coronajihad, #NizamuddinIdiots, #TablighiJamatVirus, and #muslimvirus were widely tweeted and retweeted.[9]

Tweets of the same flavor were also rife in local languages. Alongside the Islamophobic hashtags, these tweets mostly shared old, unrelated videos suggesting that Muslims were trying to purposely spread coronavirus. One widely circulated video involved an elderly fruit vendor being accused of sprinkling urine on the fruits he was selling. Another old video of a Sufi ritual, with Muslims purposely sneezing to spread the virus, went viral. Many other false, convoluted video and audio clips depicting Muslims licking utensils, scattering currency notes, and spitting on food to spread the coronavirus were created and shared during this time (Pooja Chaudari, Alt News, 2020).[10] On April 2, numerous religious gatherings took place across India to celebrate the Hindu festival of *Ram Navami*, but the same blame and discrimination pattern did not ensue; Islamophobic tweets proliferated on these days too.

A notable aspect is that some of the fake news shared during this time recalled precolonial events involving Muslims. The Hindu journalist Samer Halarnkar wrote in April 2020: "Last week, I exited a family WhatsApp group after listening to a particularly inaccurate, agitated and rambling rant [...] about how Hindus have been 'humiliated, subjugated and massacred', [...] and how we were ruled by [Muslim] 'marauders' for 1000 years" (Halarnkar, 2020). Similarly, some tweets accused Muslim NGOs and charities of denying free supplies and meals to poor Hindus and Christians during the lockdown, asserting that "Mughal rulers too denied food supplies to Hindu subjects during famines. They are only continuing the legacy". Other tweets pointed to the *Tablighi* convention as a tool for spreading Islam, converting Hindus, and evolving into a "Corona-Jihad", suggesting that coronavirus could be used as a weapon for a renewed violent attack on India.[11]

In the next section, we briefly discuss historical accounts of precolonial Muslim rule in India and how it has shaped the collective memory of non-Muslims.

2.2. Historical conflict in India

In the precolonial period, the Indian subcontinent was divided into many independent and politically fragmented states that, throughout the centuries, have often been in conflict with each other. These conflicts, motivated by aims of territorial expansion mixed with religious motives, have been shaping the long-term development patterns of the country (Dincecco et al., 2022).

The earliest Muslim invasions of India can be traced back to the seventh and eighth centuries. Arabs reached the Bombay coast in 636 AD, and the Umayyud campaigns took place across the present-day Pakistan-India border between 712 and 740. However, not until the campaigns of Mahmud of Ghazni, beginning in 1001 with the Battle of Peshawar, did the banner of Islam reached the heart of India (Britannica, 2022). Ghazni sacked and conquered several cities, including the Hindu temple city of Somnath. His empire was overthrown by the (Muslim) Ghurid dynasty in 1186, which was in turn succeeded by the (Muslim) Delhi Sultanate in 1206.

Despite the difficulty in clearly defining "Hinduism" or "Islam" at the time, the close contact between Arabs, Persians, and Turks, and the people of the Indian subcontinent through war and trade possibly led to an antagonistic religious identification (Mukhia et al., 2017). As early as the 11th century, the prominent Muslim scholar Al Biruni wrote *"They (the Hindus) totally differ from us in religion, as we believe in nothing in which they believe and vice versa (...). Their fanaticism is directed against those who do not belong to them—against all foreigners. They call them mleccha, ie. impure, and forbid having any connection with them, be it by marriage or any other kind of relationship, or by sitting, eating, drinking with them, because thereby, they think they would be polluted. (...) The Hindus claim to differ from us, and to be something better than we, as we on our side, of course, do vice versa!"* (Sachau, 2013, p.17).

Tensions among different religious identities continued to rise until the 16th century, when several conflicts between major historical rival states – the Delhi Sultanate, the Deccan Sultanates, the Rajput states, and the Vijayanagara Empire – took place.[12] While conquests were not exclusively motivated by religion, rulers were particularly harsh in their treatment of people from different religious affiliations. Desecrations of Hindu temples by Muslims were clear manifestations of these conflicts, trying to undermine and destroy Hindu religious identity. In particular, the Delhi Sultanate (1206–1526) tried to build a Muslim state and society in Northern India, using selective temple desecration to delegitimize and extirpate Indian ruling houses (Eaton, 2000).

The Delhi Sultanate succumbed to another Muslim entity, the Mughal Empire, which became one of the most powerful states on the Indian subcontinent. The reign of the Mughal ruler Aurangzeb (1658–1707) featured temple desecrations, differential customs duty based on religion, replacement of Hindu headclerks and accountants by their Muslim counterparts, an additional tax on nonbelievers "to spread Islam and put down the practice of infidelity" (the "*jaziya*"),

[9] As of April 1 at 7 p.m. the fake news hashtags #NizamuddinIdiots and #TablighiJamatVirus were posted in 53,900, and 27,500 tweets, respectively. These are sizeable number. As a comparison, as of December 31 at 7 p.m., the hashtags #HappyNewYear (relating to the start of 2020) which should be, in principle, shared by a very large portion of the population, received about 551,700 tweets. Source: https://getdaytrends.com/india/.

[10] Yet others depicted Muslims groping and misbehaving with nurses at hospitals, beating up doctors, and serving food mixed with human feces.

[11] The full text of the tweet is: "The main purpose of establishing this organization is to spread Islam. It has so far converted people from other religion, mainly Hindus, and trained them in Islamic religious matters and it is spread in 150 countries. ... Are they Human beings? #Corona_Jihad".

[12] In the Delhi and Deccan Sultanates, Islam was professed, while in the Vijayanagara Empire, Hinduism was the main religion. The Rajput states were mostly Hindu; only a few professed Islam.

and rewards for Muslim conversions (Sarkar, 1930). Religion was also a key factor during the Mughal–Maratha and Mughal–Sikh battles.[13]

The nature of interstate conflict changed after the 1757 Battle of Plassey and the victory of the British East India Company, which established itself as the major player in the political landscape, gradually defeating other states and local rulers. Following the Battle of Plassey, Indians of all religious faiths saw themselves increasingly allied against a common enemy, the European colonial powers.

The history of past Muslim attacks has left a permanent wound in the memory of other communities in India, shaping beliefs and perceptions towards Muslims. Wolpert (2004) argues that attempts to unify India before the British colonization had always been difficult and short-lived, with religious influences having divisive effects. Along the same lines, the 1947 partition of British India between Hindu-majority India and Muslim-majority Pakistan and Bangladesh triggered a massive population transfer: Hindus and Sikhs left Pakistan and Bangladesh for India, while Muslims left India to reach the territories where their religious community was larger (Bharadwaj et al., 2015).[14] More recently, during both the World Value Survey for India in 1990 and a major 2019–2020 Pew Research Center survey of religion across India, around 30% of the surveyed population reported they were not willing to have Muslim neighbors (Inglehart et al., 2018; Pew Research Center, 2021).[15]

Even today, non-Muslim communities often follow practices that recollect and remind themselves about past Muslim atrocities. For example, the Rajput festival Jauhar Mela is celebrated every year in Chittorgarh, in Rajasthan, to remember the "*jauhar,*" a mass self-immolation custom performed by Hindu Rajput women to avoid capture, enslavement, and rape by foreign invaders. Locals from Chittorgarh believe *jauhars* were performed three times during history, each time as a consequence of an invasion by a different Muslim ruler. Another example is the chant of the Sikh community in their holy places (*gurudwaras*) that refers to an episode in history where Sikh women were jailed by Mughals and forced to grind flour with heavier-than-normal millstones. The chant[16] venerates those women who chose not to convert to Islam despite the kidnapping and killing of their children.

3. Data and descriptive evidence

We have assembled a rich dataset from several primary and secondary sources. In this section, we describe the dataset, introduce the main variables used to test our hypotheses, and provide some preliminary evidence on the geographic and historical determinants of anti-Muslim fake news. Online Appendix A provides further details on some of the steps we undertook to build the main dependent variable and shows summary statistics for all variables used in the empirical analysis.

We carried out our analysis at the district level. India is currently divided into 773 districts across 28 states and 8 union territories. However, since our analysis employs several variables from the 2011 census, we focus on the administrative divisions of India based on this census year; the resulting sample contains 626 districts and 34 states or union territories.[17]

3.1. Twitter data and main dependent variable

At the core of this project lies the collection of a novel dataset of georeferenced text data posted on the social media platform Twitter. In the following subsections, we describe (i) how we assembled our main sample of tweets; (ii) how we constructed our main dependent variable by identifying anti-Muslim fake news in our sample.

Main sample

We obtained all georeferenced tweets published in India from December 1, 2019, to April 30, 2020 (Twitter API for academic research). This set of queries, conducted at the beginning of 2022, returned 16,967,380 tweets. Each tweet comes with either precise coordinates or a place identifier internal to Twitter.[18] In the latter cases, we also obtained the geographic coordinates of all place identifiers (Twitter API for academic research). For our analysis, we focus on tweets whose place identifier is below the state level (these represent roughly 84% of the original sample). In our main exercise, we use the subset of tweets posted from March 24 to April 6, i.e., from one week before to one week after the *Tablighi* shock on March 30. This sample comprises 1,863,349 tweets published by 187,787 distinct users in 21,977 locations.

Anti-Muslim fake news

We assembled an extensive list of English-language hashtags and keywords related to anti-Muslim fake news in India. We started our search for keywords by reading the common fake news stories gathered by AltNews, one of the main Indian fact-checking websites, in March and April 2020. In particular, we concentrated on fake news that contains the keyword "Muslim" in their titles or excerpts.[19] This list suggests two sets of relevant keywords. The more prominent set identifies Muslims as intentionally spreading the coronavirus through "sneezing", "spitting", "licking", or "peeing" on people or food. The second set associates Muslims with terrorism, by mentioning, for instance, "Islamic Jihadis" or "Corona Jihadis". We first searched the tweets in our dataset containing these initial keywords (and their English-language variations) and the keyword "Tablighi", then we adopted a snowball approach and collected additional relevant keywords that further identify fake news. Appendix Table A2 reports the final list of all keywords we gathered.

Then, we translated our keyword list into 11 Indian languages: Hindi, Kannada, Malayalam. Gujarati, Marathi, Punjabi, Odia, Tamil, Telugu, Urdu, and Bengali. We selected these languages because (i)

[13] The Mughal–Sikh battles started with the killing and jailing of Sikh leaders by Mughal emperors (who were intent on halting the expansion of Sikhism), and the clashes continued for more than a century. Similarly, the Mughal–Maratha battles involved Muslims and Hindus. Shivaji, the leader of the Marathas, opposed the tax on non-Muslims and revived Hindu traditions in his new empire, taking up Sanskrit and Marathi and abandoning Persian as the court language.

[14] Verghese (2021) emphasizes the relevance of divisions and clashes between Hindu and Muslims before colonialism as a potential driver of Hindu–Muslim communal violence after independence.

[15] The India Value Survey of 1990 was based on 2500 interviews, while the Pew Research Center survey of religion in India conducted nearly 30,000 face-to-face interviews of adults in 17 languages. The latter survey was carried out before the COVID-19 pandemic. The exact wording of the question in 1990 was: "On this list are various groups of people. Could you please sort out any that you would not like to have as neighbors?" ("Muslims" was one of the options in the list), while in 2019 the question was: "Would you be willing to accept a Muslim as a neighbor?"

[16] *"Singhnian jinna ne sawa sawa mann de pisne peese, bachiye de tota galean vich pavaye, par Dharm na haariya",* which translates to "The lioness like Sikh women who grinded a ton of grain, who wore their torn kids around their necks, but did not give up on religion".

[17] There are 641 districts, 28 states and 7 union territories in the 2011 Indian census. We drop four districts associated with the Lakshadweep, Nicobar, and Andaman islands. Because Twitter's georeferencing service often assigns tweets within the Delhi state to its centroid, we group together its nine districts and use the Delhi state as a unique observation in our analysis. Finally, some control variables are missing for three districts. Therefore our final sample comprises 626 districts.

[18] See https://developer.twitter.com/en/docs/twitter-api/v1/data-dictionary/object-model/geo for more information on geolocation in Twitter (last accessed: the main sample of tweets was retrieved between January 28 and March 2, 2022).

[19] Appendix Table A1 reports all fake news we retrieved from the website.

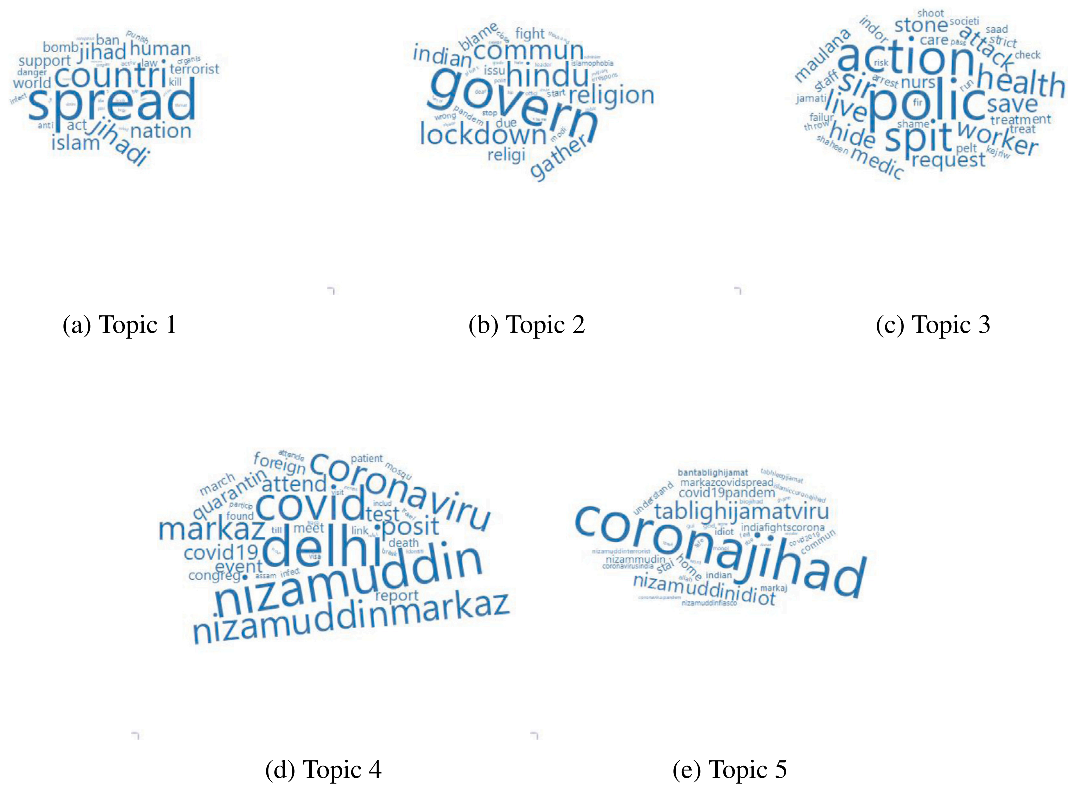


Fig. 1. Wordclouds for the LDA topics in the sample of Anti-Muslim tweets posted in English between March 24 and April 4, 2020.

they are the 11 most-spoken languages in India (excluding English), and (ii) they are included in the list of languages that Twitter allows users to choose as their preferred language. Based on the Twitter language classification, about 40% of the tweets in our working dataset are published in English, 30% percent in Hindi, and 14% in other languages (none of which exceeds 2.5% of the tweets). For the remaining 16%, Twitter was not able to identify the language of writing.

To validate the ability of our set of keywords to correctly capture Muslim-related fake news stories and illustrate their content, we perform the Latent Dirichlet Allocation (LDA) algorithm on the subset of tweets written in English and containing at least one anti-Muslim keyword.[20] The LDA is an unsupervised machine-learning algorithm that can detect latent topics in a corpus of documents from the co-occurrence of patterns of words (Blei et al., 2003). A topic is defined as a probability distribution over the entire set of words present in the corpus, and the number of topics is a parameter left to the choice of the researcher. In our baseline exercise, we set the number of topics equal to five, while the other free parameters (governing the shape of the underlying probability distributions) were set according to the standard values in the literature.[21] As is standard in text analysis (and in particular for text analysis of Twitter data), we preprocessed the text to remove punctuation, numbers, URLs, and mentions; removed stopwords; and reduced the remaining words to their English stems, so that, for instance, “spreader”, “spreading”, and “spreaded” were all replaced with “spread” in the analysis. Fig. 1 shows the wordclouds for each of the five topics. Each wordcloud is composed of the words with the largest probability weights in the corresponding topic (provided they pass a minimum cutoff, for better readability), such that the size of the word is proportional to the weight. The topics isolate different aspects of the tweets’ anti-Muslim content.

Topics 1, 3, and 5 seem to be the ones most likely to contain anti-Muslim false stories. Indeed, Topic 1 centers around the accusation that Muslims are responsible for spreading the virus in India, often drawing an analogy between spreading Islam and spreading the virus. Here are two sample tweets: “All this while I thought Islam isn’t a Religion, it’s a cult, I was wrong, - Islam is a disease”.; and “The main purpose of establishing this organization is to spread Islam. [...] #Corona_Jihad”. Topic 3 captures a specific type of fake news that spread in the aftermath of the *Tablighi* gathering, accusing *Tablighi* participants of intentionally infecting other people and other obscene behaviors. Here are two sample tweets: “Tablighi Jamaat members in quarantine are walking around without trousers on, listening to vulgar songs, asking for bidi cigarette from nurse and staff and making obscene gestures towards nurses. Asks police to restrain them”; and “Are the docs are lying that ur orthodox Muslim men would be naked in front of women medical staff, spit on them, in some place pelt stone [...]”. Topic 5 portrays the Tablighi Jamaat convention as a form of terrorism, using expressions like “corona jihad” and “human bombs”. On the other hand, Topic 2 focuses on the responsibilities of the local government and of the police in managing the *Tablighi Jamat* episode and Topic 4 identifies the *Tablighi Jamaat* convention as a COVID hotspot and blames the event for the spread of COVID from New Delhi to other parts of India.

Fig. 2 reports the share of tweets containing anti-Muslim fake news over the total number of tweets posted in a given day. Interestingly, the share of discriminatory tweets is virtually zero in the weeks preceding the *Tablighi* shock, represented in the graph by the solid red vertical line. Consistent with the narratives discussed in Section 2.1, an abrupt spike occurs on March 31, after the events of March 30 established a link in the public’s mind between the *Tablighi Jamaat* convention and the COVID-19 crisis. The share of anti-Muslim fake news peaked above 3% of total tweets on April 1, then steadily declined—though it

[20] We rely on the language classification provided by Twitter to detect tweets written in English. These tweets may still contain some non-English characters.

[21] To maximize interpretability, we set the number of topics equal to five, the parameter governing the shape of the topic distribution equal to 50/number of topics = 10, and the parameter governing the shape of the term distribution equal to 0.09.

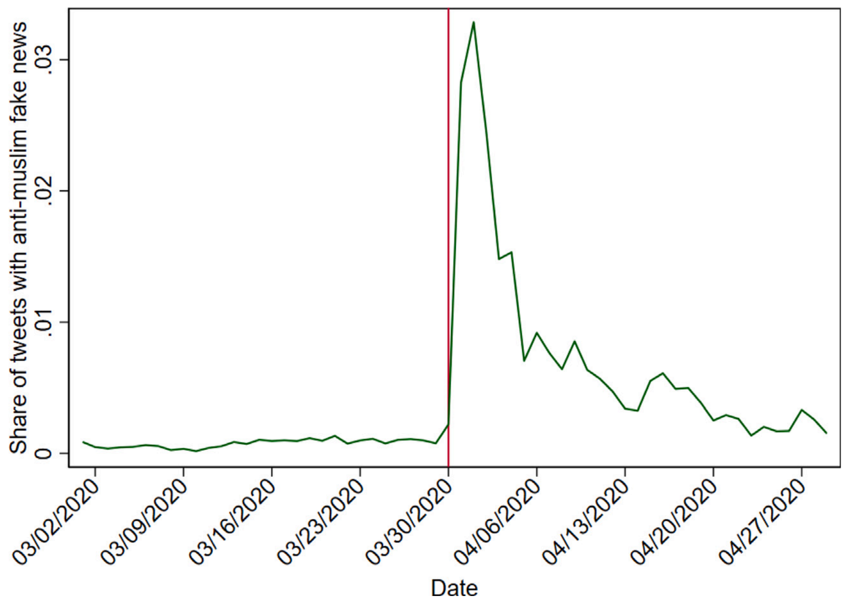


Fig. 2. Time series of Anti-Muslim fake news. *Notes*: The green line represents the share of tweets containing anti-Muslim fake news keywords over the total number of tweets in our sample in each day from March 1 to April 29. The solid red line highlights the date of the shock, March 30. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

remained above its preshock level one month later. In our empirical analysis, the event-study exercise focuses on the full months of March and April. However, as most of the increase in anti-Muslim fake news occurs in the first week after the shock, in the main regression Tables we zoom in on the two weeks centered around March 30, i.e., from March 24 to April 6.

3.2. Main explanatory variables

Distance to New Delhi

To test whether discriminatory fake news spread spatially, we focus on *Distance to New Delhi*, that is, the location where the *Tablighi* shock took place. Specifically, we reference the centroid of each district and calculate its distance (in units of 1000 km) from the coordinates of New Delhi. In robustness checks, we also consider a dummy equal to one for the New Delhi Region, comprising the Delhi state and the surrounding districts.

Muslim attacks

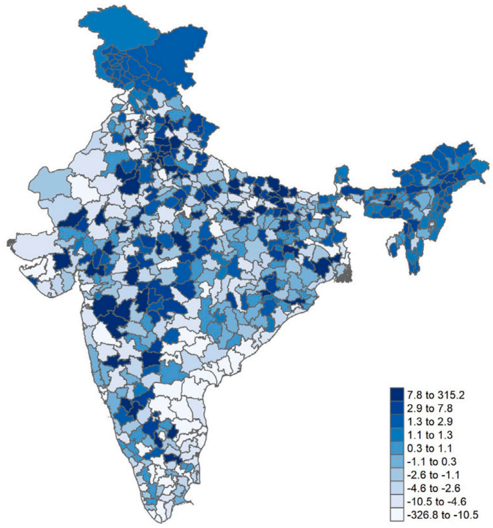
To test whether the propensity to spread fake news against Muslims has deeply rooted historical origins, we focus on precolonial conflicts in which a Muslim group or entity participated as an aggressor. As discussed in Section 2, this class of conflicts, which we refer to as Muslim attacks, characterized the history of the Indian subcontinent for several centuries. Our measure of exposure to Muslim attacks relies on the database assembled by Dincecco et al. (2022)—based on Jaques (2007)—which contains information on the universe of conflicts recorded on the Indian subcontinent between the years 1000 and 2000.[22] For each conflict, we have the date, the geographical coordinates, and a short description of the event. Starting from these records, we manually selected conflicts initiated by a Muslim group or entity. As in Dincecco et al. (2022), we focus on land battles that occurred in precolonial times (i.e., before 1757).[23] Our baseline district-level measure of historical anti-Muslim sentiments is, thus, the dummy variable *Muslim Attack* taking the value one if at least one Muslim land-based attack is observed in a district in the 1000–1757 period. In our robustness checks, we also present results with perturbed versions of this measure that rely on different time periods and with a distance-based measure of exposure to Muslim attacks (along the lines of Dincecco et al., 2022).

3.3. Preliminary evidence on the main hypotheses

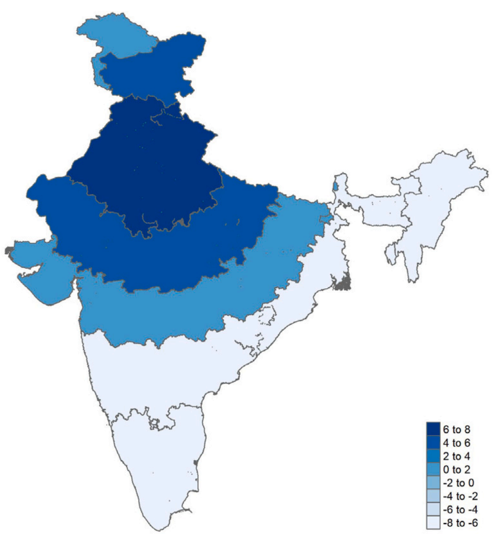
Fig. 3(a) shows the geographical distribution of anti-Muslim fake news across Indian districts in the week following March 30. To account for heterogeneous levels of Twitter activity across districts, the map reports the residuals of an OLS regression of the number of tweets with anti-Muslim fake news posted in the week after the shock on the total number of tweets posted in the week preceding the shock. The color scale corresponds to the different deciles of the distribution of residuals, ranging from low discrimination (white) to high discrimination (dark blue). There is substantial spatial variation in the intensity of anti-Muslim discrimination in the aftermath of the shock. Discrimination is higher in the area around New Delhi, along the Ganges Valley in the northeast, and along the western coast, whereas it is lower, for instance, in the southwest, in the area corresponding to the state of Andra Pradesh. A second feature that emerges from the figure is the spatial autocorrelation in the intensity of discrimination across Indian districts. We confirm this formally by computing the Moran's I spatial

[22] Even if historical data may suffer of measurement errors, Jaques (2007) can be regarded as a reliable and comprehensive source of information for conflict data. In particular, the author's stated objectives are twofold: document as many conflict events as possible and ensure maximum accuracy in reporting. To achieve this, Jacques made sure that each event recorded was supported by at least two independent sources. As suggested by Dincecco et al. (2022), this approach may result in the exclusion of events whose memory is only within the oral tradition, thus being more severe in settings such as pre-colonial Africa rather than in the Indian context.

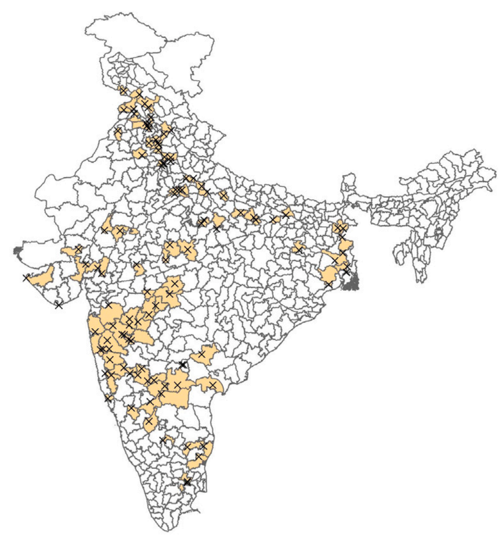
[23] We concentrate on land battles because they clearly occurred within specific district borders and were by far the most common type of precolonial conflicts — 235 out of 245 events (which corresponds to 95.92%) are land battles.



(a) Anti-Muslim Fake News (residuals), Week After March 30



(b) Average Residuals by Quintiles of Distance to New Delhi, Week After March 30



(c) Location of Precolonial Conflicts with Muslim Offenders

Fig. 3. Spatial distribution of Anti-Muslim fake news and precolonial muslim attacks. *Notes*: Fig. 3(a) depicts the residuals of an OLS regression of the number of tweets with anti-Muslim fake news posted in a district during the week after the shock (March 31 to April 6) on the total number of tweets posted in a district during the week before the shock (March 24 to March 30); the color scale represents different deciles of the distribution of residuals, where darker colors correspond to higher residuals. In Fig. 3(b), the same residuals are averaged by quintiles of distance to New Delhi. Distance is calculated from the district centroid and the district polygons have been merged within each quintile. In Fig. 3(c), the black crosses coincide with the exact locations of land-based precolonial (1000–1757) conflicts where the offender was a Muslim group or entity; the districts in which at least one such conflict is recorded are highlighted in yellow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

autocorrelation index on the same residuals depicted in the figure: we obtain a value of 0.28 (significant at the 1% level).

Fake news spread spatially. First, in Fig. 3(b), we averaged the residuals reported in Fig. 3(a) for five rings corresponding to the quintiles of the distribution of distance to New Delhi.[24] We consider distance to New Delhi because the shock is narrowly localized there, where the *Tablighi Jamaat* convention took place. As the map shows, there is a clear negative gradient from New Delhi to the outer regions of India, suggesting that distance mattered for the diffusion of anti-Muslim fake news, and starting to corroborate our first hypothesis.

Historical conflicts affect the reception of fake news. Second, Fig. 3(c) shows the geography of precolonial Muslim attacks. The black crosses coincide with the exact conflict locations, while individual districts are highlighted if at least one conflict occurred in their territory. There is a striking resemblance to the map of anti-Muslim fake news shown in Fig. 3(a). This suggests that historical Muslim attacks may have

[24] In the figure, the district polygons for districts within the same distance quintile have been merged into a single polygon.

left some anti-Muslim sentiments and perceptions that could shape the reaction of different Indian districts to the *Tablighi* shock. This provides preliminary evidence on our second hypothesis.

4. Empirical specification

To investigate the determinants of the diffusion of anti-Muslim fake news on Twitter, we use a difference-in-difference strategy. Specifically, we leverage the evening of March 30 as the shock cutoff date during two periods: (i) from March 28 to April 2 (capturing three-day windows before and after the shock), and (ii) from March 24 to April 6 (capturing seven-day windows before and after the shock).

The estimated equation is:

$$Y_{it} = \gamma Post_t + \beta Z_i + \delta Z_i \times Post_t + \phi X_{it} + \theta_i + \eta_t + \varepsilon_{it}$$

where Y_{it} is the number of tweets with anti-Muslim fake news posted in district i on day t, $Post_t$ is a dummy taking the value one for all dates after March 30, and Z_i captures our potential determinants for the rise of anti-Muslim fake news, i.e., a measure of spatial diffusion or of historical Muslim attacks. Our main coefficient of interest is δ, which is the difference-in-difference coefficient of the interaction term between Z_i and $Post_t$. It shows how much stronger the increase in the diffusion of fake news after the shock is further away from New Delhi or in an area where a conflict was initiated by Muslim entities, compared to the rest of India.

Moreover, we account for a set of district-level time-variant controls X_{it}, including: (i) the daily number of COVID-19 deaths, as the spread of fake news could simply be related to the spread of the coronavirus, rather than to distance to the *Tablighi* hotspot news or to the legacy of precolonial Muslim attacks; (ii) the average number of tweets in the week prior to the shock interacted with the $Post_t$ indicator to control for differential effects over time of Twitter penetration and—when we focus on the relationship between Muslim attacks and fake news diffusion–; (iii) the *Delhi Dummy*, which is a dummy variable taking the value one for the Delhi state, interacted with the $Post_t$ indicator; and (iv) a dummy for districts experiencing other precolonial conflicts (not originated by Muslim attacks) interacted with the $Post_t$ indicator to control for differential effects of the legacy of historical violence beyond Muslim attacks.

Finally, θ_i and η_t denote, respectively, districts fixed effects and date fixed effects to account for unobserved district-specific and time-specific factors.[25] The error term ε_{it} is clustered to account for both spatial autocorrelation up to 250 km based on Conley (1999) and serial autocorrelation.

5. Empirical analysis

In this Section we empirically test our two hypotheses on the potential drivers of the propagation of discriminatory false stories at the onset of the coronavirus pandemic. In Section 5.1 we assess whether fake news spread spatially from New Delhi and discuss potential mechanisms underlying the results. Section 5.2 turns to the role of historical Muslim attacks in affecting the diffusion of anti-Muslim fake news.

5.1. Fake news spread spatially from Delhi

In our first hypothesis we postulated that fake news diffused spatially from New Delhi after March 30, when the *Tablighi Jamaat* convention became publicly recognized as a COVID hotspot. Conceptually, we may observe a spatial diffusion pattern through two channels: (i) the *communication channel*, based on which the probability to tweet fake news depends on the probability of receiving the news of the Tablighi shock or seeing initial fake-news messages, and (ii) the *salience channel*, according to which the probability to tweet fake news depends on the importance of the Tablighi shock for a given individual (Twitter user). For instance, individuals with stronger personal ties to Delhi may perceive a greater risk of adverse consequences for people within their network due to the virus spreading closer to the Tablighi convention.

We study the diffusion of fake news from New Delhi using the continuous variable of physical distance interacted with the *Post-March 30* dummy. Table 1 shows the results. Column 1 controls for state and date fixed effects as well as for district-level time-invariant and time-variant controls.[26] Column 2 adds district fixed effects. In both specifications, the interaction term is negative and significant at the 1%

Table 1
Number of anti-muslim fake news tweets in space.

Period:	Three days		Seven days	
Dependent variable:	N. Tweets FN		N. Tweets FN	
	(1)	(2)	(3)	(4)
Dist. New Delhi (×1000)	−1.09 (1.02) [0.29]		−0.64 (0.80) [0.42]	
Dist. New Delhi (×1000) × Post	−1.95*** (0.40) [0.00]	−1.95*** (0.40) [0.00]	−1.38*** (0.33) [0.00]	−1.38*** (0.32) [0.00]
Time-variant controls	Yes	Yes	Yes	Yes
Time-invariant controls	Yes	No	Yes	No
State FE	Yes	No	Yes	No
District FE	No	Yes	No	Yes
Day FE	Yes	Yes	Yes	Yes
R-squared	0.88	0.92	0.76	0.79
Observations	3750	3750	8750	8750

Notes: OLS estimates. Observations are districts in each day in two periods: March 28–April 2 in columns 1–2, and March 24–April 6 in columns 3–4. We exclude Delhi from the sample. In all specifications, the dependent variable is the number of tweets with anti-Muslim fake news. Time-invariant controls are described in footnote 26, whereas time-variant controls include the daily number of COVID deaths and the average number of tweets in the week before March 31, interacted with the *Post-March 30* dummy. In columns 1 and 3, we control for *Dist. New Delhi (x 1000)*, which is the distance from each district's centroid to the coordinates of New Delhi (per 1000 km). In all columns this variable is also interacted with the *Post-March 30* dummy. Standard errors in parentheses are clustered to account for both spatial correlation (up to 250 km) and serial correlation. P-values are reported in brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

[25] Appendix Table D1 and D2 show that our main results are robust to using subdistricts and towns as units of observations, respectively including subdistrict and town fixed effects.

[26] State fixed effects capture unobserved historical, linguistic, political, and economic characteristics at the state level but do not account for differences at a finer spatial level. To address this issue, in all specifications with state (rather than district) fixed effects, we control directly for prominent time-invariant socio-economic and geographic characteristics at the district level. In particular, following insights from Dincecco et al. (2022) and Michalopoulos and Papaioannou (2018), we include the log of luminosity (+0.01) averaged in the 1992–2010 period to account for local economic development – with luminosity being a better proxy than official GDP data, especially in poorer areas – and the log of population density in 1990 – the most recent year prior to the years in which luminosity is measured –. Moreover, we control for the log share of Muslim and Hindu population, which may account for district-specific Muslim *versus* Hindu contact, and we include the log shares of literate and of urban population to account for broader measures of well-being and development. To capture indirect or direct correlates of district-level differences in Twitter use, we control for the average number of tweets during the week before the shock. All these data come from the 2011 census, except for the average number of tweets, computed using our database. Moving to the geographical controls, following Dincecco et al. (2022), we include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Finally, we add distance to the border as well as distance to the closest border of Pakistan or Bangladesh—the two countries bordering India with the highest shares of Muslim population.

level suggesting that fake news diffused out of New Delhi. In particular, focusing on column 2, the coefficient estimate suggests that moving one standard deviation farther from New Delhi (roughly 558 km) reduced the daily number of tweets with anti-Muslim fake news by 1.1 tweets. This compares with an average number of Anti-Muslim fake news tweets increase of 4.47 in the three days after the shock. Columns 3–4 replicate the analysis of columns 1–2, focusing on the longer, seven-day time horizon. The results are qualitatively similar, although smaller in magnitude. Columns 2 and 4 of Table 1 represent our preferred specifications and will be used as baseline in our robustness exercises and in the analysis of the mechanisms.

Robustness. To corroborate our finding that fake news spread spatially from the location of the shock, in Appendix B, we perform a series of robustness checks.

First, we use an alternative measure of our explanatory variable: the *New Delhi Region Dummy*, which is a dummy variable taking the value one for the districts located around the Delhi state.[27] Consistently with our baseline result, we find a significantly higher rise in the number of tweets with anti-Muslim fake news after the *Tablighi* shock in the New Delhi region, compared to districts outside this region (see columns 1–2 and 6–7 of Appendix Table B2).[28]

Second, our conjecture of the spatial diffusion of anti-Muslim fake news is confirmed when using a longer time horizon, i.e., the sample of tweets posted one month before and one month after the shock. Specifically, we perform an event-study analysis in which, rather than interacting distance to New Delhi with the Post indicator, we interact the distance measure with each of the time-period dummies, using March 1 as the baseline period. The results are depicted in Fig. 4, plotting the coefficients of the interaction terms. Interestingly, a clear negative distance gradient emerges in the week following the shock, gradually fading out over time. Reassuringly, there were no differential trends in anti-Muslim fake news before March 30, i.e., before the public began connecting the Tablighi convention and the COVID outbreak.

Third, our results are robust to accounting for spatial autocorrelation over distance thresholds ranging from 100 to 1500 km (Figure B1) and when considering alternative definitions of our dependent variable, such as the daily share of tweets with anti-Muslim fake news (Table B3).

Fourth, note that, being New Delhi the capital city, our findings may in principle be biased by a set of time-varying omitted variables, correlated with distance to the capital, such as the strength of government intervention during the pandemic period. In particular, a first concern could be that the accessibility to health-care services may have played a role in containing the epidemics in the aftermath of the Tablighi shock, thus mitigating its effect on the diffusion of anti-Muslim fake news.

To partially address this concern, we use data on the presence of medical professionals (doctors, hospital doctors, and paramedics per thousand inhabitants) and of health care facilities (clinics and hospitals per thousand inhabitants) at the district level from the 2011 census. Along the same lines, we exploit information on the number of COVID tests per thousand inhabitants (source: https://data.covid19india.org/). A second concern is that a stronger enforcement of lockdown measures by the local police may have favored scapegoating behaviors against Muslims, thus enhancing the diffusion of fake news. To account for this, we gathered data on: (i) the number of all entities labeled as police stations from Open Street Map; and (ii) the number of police stations authorized to issue passports (source: https://passportindia.gov.in/AppOnlineProject/online/LocatePSAction).[29] As displayed in Fig. 5, all results are robust to the inclusion of these variables (interacted with the *Post* indicator) separately and when simultaneously accounting for all measures of healthcare capacity and the reported number of police stations or the number of police stations issuing passports.[30]

Finally, Appendix Figure B2 shows the sensitivity of our main coefficient of interest—*Distance to New Delhi* × *Post*—to including in the regression other potentially confounding factors interacted with the *Post March 30* dummy. We start with district-level socio-economic and geographic controls (as described in footnote 26) and then perform the same exercise using (i) *additional geographic controls*, (ii) proxies of *initial conditions*, (iii) *colonization controls* to account for the role of district-specific differential exposure to the subsequent British colonization, (iv) *fractionalization controls* to specifically consider the role of ethnic, linguistic, and religious diversity in generating anti-Muslim fake news, and (v) *current trade routes controls* to account for the role of present-day communication ways in the diffusion of both the COVID epidemics and anti-Muslim discrimination.[31] None of these perturbations of our main specification sensibly affects our results.

All these exercises, taken together, support the spatial diffusion hypothesis and alleviate the concern that other time-varying variables, correlated with proximity to the capital, are confounding our results.

Mechanisms. To investigate the mechanisms behind the relevance of geographical distance in the diffusion of fake news, we explore the role of two possible factors: Twitter connectedness and district-level personal ties to Delhi.

First, we study whether the network of social interactions on Twitter played a role. To do this, we exploit the fact that when tweets are posted in response to a tweet (known as a reply), or by quoting another tweet, the Twitter API also provides the ID of the original tweet. To capture *Twitter connectedness to Delhi*, we focus on tweets posted in December 2019 and January 2020, and for each district i we compute the (standardized) share of quotes and replies to tweets

[27] The New Delhi region area, defined in 1985 when the government approved a regional development plan, encompasses the Delhi state and 21 surrounding districts across the states of Haryana, Uttar Pradesh, and Rajasthan. Its population is over 46 million inhabitants, and its urbanization rate is 62.6% (Census of India, 2011).

[28] Results hold if we use as explanatory variable the *Delhi Dummy*, that is a dummy variable taking the value one for the Delhi state. See Appendix Table B1.

[29] We geocoded this information using the Google API.

[30] Appendix Tables B4 and B5 present both the coefficients associated with *Distance to New Delhi* and those related to the different indicators of government health and enforcement capacity (interacted by *Post*).

[31] *Additional geographic controls* include distance to the coast, presence of a river, irrigation potential, the coefficient of variation in rainfall, the percentage of forested area, and distance to petroleum, diamonds, gems, and gold deposits. These variables are from several sources, including the Natural Earth Data website (https://www.naturalearthdata.com/), Matsuura and Willmott (2009) and Tollefsen et al. (2012), the India Institute of Forest Management (2015), and Bentzen et al. (2017). We proxy *initial conditions* by including initial state-capacity measures such as the number of Indian settlements during the Neolithic or Chalcolithic Ages from Nag (2007), the number of important Indian cultural sites between 300 and 700 and between 800 and 1200 from Schwartzberg (1978), and the natural logarithm of (one plus) the total urban population in the year 1000 according to Chandler (1987). *Colonization controls* include a dummy variable for direct British rule from Iyer (2010), the number of years a district was ruled by the British from Verghese (2016), and a variable tracking the year in which each district was connected to the first colonial railroad from Fenske et al. (2021). *Fractionalization* measures come from various sources and include medieval ports and a dummy for districts intersected by the Ganges river, religious fractionalization and polarization, ethnic polarization, and the scheduled caste and tribe shares from the 2011 census as well as the number of years a district was ruled by Muslims in medieval times. All these variables (except for the number of years a district was ruled by the British) are provided in Dincecco et al. (2022), which also discusses details on their sources and their construction. Finally, we proxy *current trade routes* by computing the total number of kilometers of roads and railroads within each district from the shapefiles for the road and railroad networks in the year 2016 from the StanfordDigitalrepository.

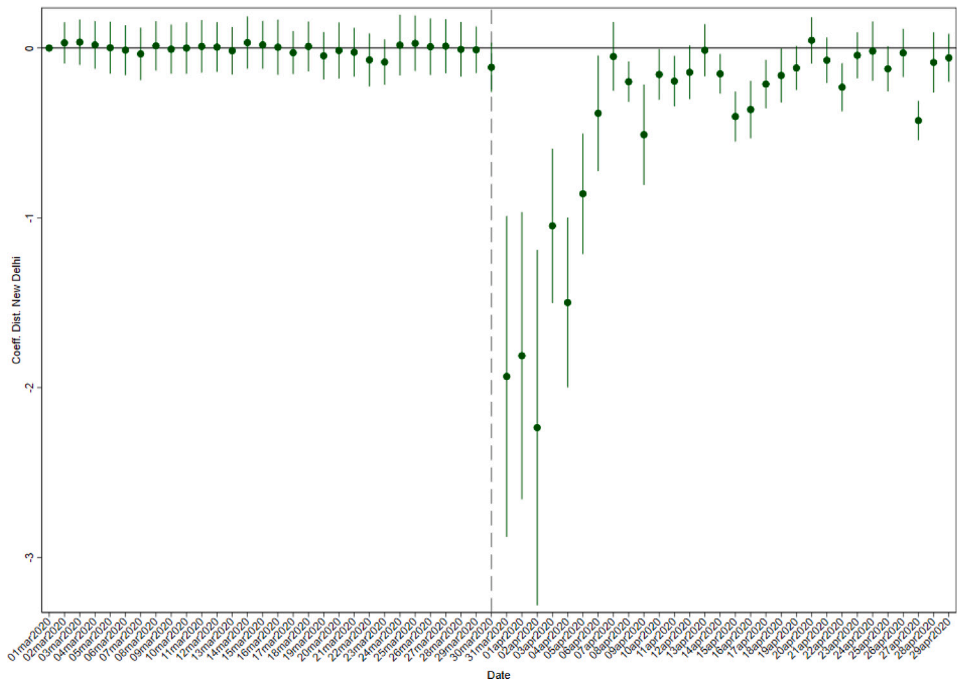
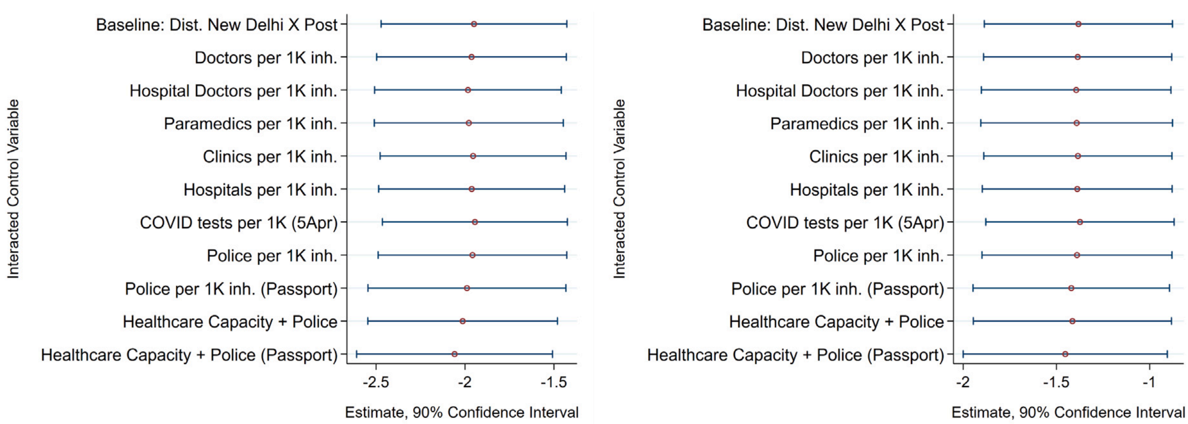


Fig. 4. N. of fake news tweets and distance to New Delhi — Event analysis. *Notes*: Daily panel analysis in the March 1–April 29 period. Each dot is a coefficient from a version of specification 2 in Table 1 that replaces *Distance to New Delhi* interacted with the *Post March 30* indicator, with *Distance to New Delhi* interacted with each date dummy. The reference category is March 1, and the specification also controls for the average number of tweets in the week before the shock interacted with date dummies. Vertical bars indicate 90% confidence intervals.



(a) 3 days window (b) 7 days window

Fig. 5. Sensitivity of distance to New Delhi to accounting for the government's ability to intervene. *Notes*: The first dot displays the coefficient associated with the variable *Dist. New Delhi* × *Post* from estimating specifications 2 (panel a) and 4 (panel b) in Table 1 as baseline estimate. In each panel the other dots display the sensitivity of this estimate to further controlling for the variables indicated on the vertical axis (interacted with the *Post-March 30* indicator). Horizontal bars indicate 90% confidence intervals. See Section 5.1 for details on all variables.

initially posted in the Delhi state.[32] The pairwise correlation between distance to New Delhi and Twitter connectedness to Delhi is −0.17, consistently pointing to a negative relationship between distance and Twitter connectedness.[33] Thus, in columns 1 and 4 of Table 2, we also account for Twitter connectedness to Delhi interacted with the *Post-March 30* dummy. Interestingly, both physical proximity and Twitter connectedness are significantly associated to the diffusion of fake news. Moreover, focusing on column 1, the coefficient on physical distance is only slightly smaller after controlling for Twitter connectedness (−1.86 with respect to −1.95 from column 2 of Table 1), suggesting that there is a physical proximity effect over and above the network of Twitter

[32] We chose the period from December 1, 2019 through January 31, 2020 to rule out that the diffusion of COVID affected social interactions between districts. Moreover, we exclude from the denominator the quotes and replies to tweets posted within the district because with this measure we want to specifically focus on Twitter connections outside the district.

[33] More generally, we also construct the full matrix of Twitter connections across Indian districts. In line with previous literature (e.g., Bailey et al., 2020), the pairwise correlation between physical distance and Twitter connectedness between district-pairs (both in logarithm) is negative and significantly different from zero, consistently suggesting that districts farther from each other display a lower level of Twitter connectedness as well. Appendix Table A5 shows that this relationship is robust to accounting for fixed effects for the districts i and j included in the pair, and for several measures of dissimilarity between districts.

Table 2
Number of anti-muslim fake news tweets in space: Mechanisms.

Period:	Three days			Seven days		
Dependent variable:	N. Tweets FN			N. Tweets FN		
	(1)	(2)	(3)	(4)	(5)	(6)
Dist. New Delhi (×1000) × Post	−1.86***	−1.15***	−1.09***	−1.32***	−0.74***	−0.70***
	(0.39)	(0.36)	(0.36)	(0.31)	(0.27)	(0.27)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.01]	[0.01]
Std.Twitter Conn. to Delhi × Post	0.31***		0.21**	0.22***		0.14**
	(0.10)		(0.09)	(0.08)		(0.06)
	[0.00]		[0.01]	[0.01]		[0.03]
Std.Born in Delhi × Post		1.71***	1.70***		1.38**	1.37**
		(0.53)	(0.53)		(0.69)	(0.69)
		[0.00]	[0.00]		[0.04]	[0.05]
Time-variant controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.92	0.92	0.92	0.79	0.79	0.79
Observations	3750	3750	3750	8750	8750	8750

Notes: OLS estimates. Observations are districts in each day in two periods: March 28–April 2 in columns 1–3, and March 24–April 6 in columns 4–6. We exclude Delhi from the sample. In all specifications, the dependent variable is the number of tweets with anti-Muslim fake news. Time-variant controls are as in Table 1. *Dist. New Delhi (x 1000)* is the distance from each district's centroid to the coordinates of New Delhi (per 1000 km); *Std. Twitter Connectedness to Delhi* computes, for each district i, the (standardized) share of quotes and replies to tweets posted in the Delhi state by users located in district i (excluding from the denominator the quotes and replies to tweets posted within the district); *Std. Born in Delhi* computes, for each district i, the (standardized) share of district inhabitants who were born in the Delhi state. All these variables are interacted with the *Post-March 30* dummy. Standard errors in parentheses are clustered to account for both spatial correlation (up to 250 km) and serial correlation. P-values are reported in brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

social interactions. This may be partly due to the fact that our measure of Twitter connectedness does not account for other social media, such as Facebook, Instagram, Snapchat, WhatsApp or Telegram.

Second, we explore whether personal ties partly explain the relationship between physical distance and the diffusion of anti-Muslim fake news on Twitter. Using data from the 2011 census, we compute, for each district, the (standardized) share of inhabitants born in Delhi and the share of inhabitants whose prior residence was in Delhi. Appendix Figure A1 presents a visual illustration of the intensity of personal ties with Delhi. Two observations stand out. First, the two measures are highly correlated (the pairwise correlation is 0.99). Second, stronger ties to Delhi tend to be concentrated in space around Delhi, as well as in coastal districts on the western part of India and in big cities such as Mumbai, Hyderabad, and Chennai. In particular, the pairwise correlation between distance to New Delhi and personal ties is negative, as in the case of Twitter connectedness, but larger in magnitude (−0.27). More formally, in columns 2 and 5 of Table 2 we augment our preferred specifications with district-specific personal ties to Delhi.[34] In both columns, the interaction between the share of population born in Delhi and the *Post-March 30* dummy is positive and significant at the 1% level, suggesting that in districts with stronger personal ties to Delhi we will observe a higher increase in tweets with anti-Muslim fake news. Note also that controlling for personal ties sensibly reduces the magnitude of the coefficient on the interaction between distance to New Delhi and the *Post-March 30* indicator (from −1.95 in column 2 of Table 1 to −1.15 in column 2 of Table 2). This points out that personal ties is likely to be one of the channels through which physical distance affects the diffusion of anti-Muslim fake news. At the same time, physical distance remains significant, probably due to the role played by other types of personal ties not captured by our measure (i.e., number of relatives and friends living in Delhi), by unobserved connections on other social media platforms, or by unobserved face-to-face interactions.[35], [36]

Finally, the picture does not change if we control for both Twitter connectedness and personal ties to Delhi (see columns 3 and 6 of Table 2).[37]

5.2. Historical determinants of Anti-Muslim fake news and the Tablighi shock

After showing that fake news spread spatially out of New Delhi, we now investigate our second hypothesis, i.e., whether deeper determinants, rooted in history, are affecting the spatial diffusion of anti-Muslim discrimination. Following the rich historical record of Muslim invasions and conquests in present-day India (see Section 2), we focus on Muslim attacks. Table 3 uses as main explanatory variables a dummy equal to one for districts that experienced Muslim attacks and its interaction with the *Post-March 30* indicator.

Columns 1–2 and 3–4 report the results of the difference-in-difference specification using the three-day and seven-day windows, respectively. The coefficients of the interaction term (*Muslim Attack* × *Post*) imply that exposure to Muslim attacks in precolonial times increased the daily number of anti-Muslim fake news tweets by one

[34] Given the high correlation between the two proxies of personal ties, we only report results using the share of population born in Delhi. Results using the share of inhabitants previously residing in Delhi are essentially the same.

[35] Some face-to-face interactions were likely allowed during the lockdown, for instance because people working in essential sectors continued their activities at their workplaces.

[36] Some studies also point to the role of algorithmic targeting of content to users based on their geolocation (see, among others, Pariser, 2011; Hannak et al., 2013). However, recent evidence shows that this may not be crucially driving the complementarity between electronic and face-to-face interactions (Guess et al., 2023; Bozarth et al., 2023).

[37] Columns 3–5 and 8–10 of Appendix Table B2 replicate specifications 1–3 and 4–6 of Table 2, focusing on the New Delhi region dummy (instead of distance) interacted with the *Post* indicator. Remarkably, once we account for personal ties to Delhi, the New Delhi region dummy interacted with *Post* becomes insignificant (compare columns 2 and 7 with columns 4–5 and 9–10), further supporting a solid association between physical distance and personal interactions.

Table 3
Number of tweets with anti-muslim fake news and historical muslim attacks.

Period:	Three days		Seven days	
Dependent variable:	N. Tweets FN		N. Tweets FN	
	(1)	(2)	(3)	(4)
Muslim attack	−0.16		−0.16	
	(0.16)		(0.11)	
	[0.30]		[0.12]	
Muslim attack × Post	1.64**	1.63***	1.06**	1.07**
	(0.64)	(0.56)	(0.50)	(0.47)
	[0.01]	[0.00]	[0.03]	[0.02]
Time-variant controls	Yes	Yes	Yes	Yes
Time-invariant controls	Yes	No	Yes	No
State FE	Yes	No	Yes	No
District FE	No	Yes	No	Yes
Day FE	Yes	Yes	Yes	Yes
R-squared	0.95	0.96	0.80	0.81
Observations	3756	3756	8764	8764

Notes: OLS estimates. The table replicates the specifications of Table 1 including the Delhi state in the sample. Columns 1 and 3 also include the dummy variable *Muslim Attack*, tracking districts that experienced attacks from Muslim groups in the 1000–1757 period, and control for a dummy taking the value one for districts that experienced precolonial conflict in which Muslims groups were not the aggressors. All columns further control for (i) *Muslim Attack × Post*, which is the interaction term between *Muslim Attack* and the *Post-March 30* dummy, (ii) the dummy for conflicts with non-Muslim aggressors interacted with the *Post-March 30* indicator, and (iii) the *Delhi Dummy × Post*, which is a dummy variable taking the value one for the Delhi state, interacted with the *Post-March 30* indicator. See Section 3 and footnote 26 for details on all variables. Standard errors in parentheses are clustered to account for spatial correlation up to 250 km in the cross-section, and for both spatial and serial correlation in the panel specifications. P-values are reported in brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

tweet compared to districts that did not experience any attack. This suggests that exposure to precolonial Muslim attacks partly explains the district variation in the diffusion of anti-Muslim fake news.

Robustness. In this section, we perform a series of robustness checks to corroborate our findings on the role of precolonial Muslim attacks in the diffusion of fake news. We use specifications 2 and 4 of Table 3 as benchmarks.

In particular, we show that our results are robust when accounting for spatial correlation of distance thresholds ranging from 100 to 1500 km (in Appendix Figure C1) and when using alternative definitions of the dependent variable, such as the share of anti-Muslim fake news (Table C1).

Next, to further shed light on the dynamic patterns of fake news diffusion, we perform an event-study analysis where, rather than interacting the *Muslim Attack* dummy with the *Post-March30* indicator, we interact the attack measure with each of the time-period dummies. The data used is from March 1 to April 29. Figure C2 displays the results. Interestingly, the difference between the number of tweets with anti-Muslim fake news in areas that experienced Muslim attacks and areas without conflicts started to increase on March 31, reached its peak two days later, and declined in the subsequent days. No differential trends in anti-Muslim fake news are displayed in the preshock period.

Moreover, we perform different exercises to explore alternative measures of exposure to Muslim attacks. As approximately 30% of the districts with historical Muslim attacks experienced this type of event more than once, in the first exercise we focus on the number of conflicts. Specifically, in columns 1 and 5 of Table 4, we use the variable *Number of Muslim Attacks* – tracking the number of conflict events in which Muslim groups were the aggressors during the 1000–1757 period – interacted with the *Post-March 30* dummy. The coefficient of the interaction term is consistently positive and significant. It suggests that after March 30 in the district with the highest number of Muslim attacks (eight events, which occurred in the district of Rupnagar), the per-day number of tweets with anti-Muslim fake news increased, on average, by up to 5 tweets more compared to districts without any conflict (column 5). In our second exercise, we consider the fact that districts not directly experiencing a Muslim attack may have nevertheless been affected, either because the conflict occurred in the proximity of the border or because the movement of armies in the territory left some scars. Building on the analysis in Dincecco et al. (2022), in columns 2 and 6 of Table 4, we compute Muslim conflict exposure as

$$\sum_{c \in C} (1 + distance_{i,c})^{-1}$$

where $distance_{i,c}$ is the distance between the centroid of district i and the location of a Muslim attack c. This measure implies that the nearer a district is to a particular Muslim attack, the more exposed it is. Muslim attacks occurring at the district centroid receive a weight of one, or full weight; as the distance of Muslim attacks from the centroid increases, they receive lower weights. In this way, we impose no cutoff at the district's borders. For each district, we consider all conflicts within a radius of 250 km. Our results hold when using this continuous measure. Note that these results are also robust to the use of an alternative radiuses (see columns 1–2 of Appendix Tables C2 and C3). In the third exercise, we construct our measure of Muslim-related conflicts over a shorter and longer time horizon. In particular, column 3 of Appendix Tables C2 and C3 considers Muslim attacks occurring between the birth of the Delhi Sultanate (around 1200) and the establishment of the British East India Company in India (in 1757); column 4 of Appendix Tables C2 and C3, instead, considers conflicts occurring until 1840, when the British established their full military and political power over the Indian subcontinent. Results are robust. In the fourth exercise, we show that, within our time period more recent Muslim attacks matter more than less recent ones (column 5 of Appendix Tables C2 and C3). At the same time, our results are stable if we consider Hindu–Muslim clashes occurring in the 1950–2000 period, rather than historical Muslim attacks (column 6 of Appendix Tables C2 and C3).[38]

Then, we show that the result on the legacy of precolonial Muslim attacks is still positive and significant if we remove Delhi from the sample (columns 3–4 and 7–8 of Table 4) and when we control for physical proximity, Twitter connectedness, and personal ties to Delhi (columns 4 and 8 of Table 4).

Finally, Fig. 6 displays the sensitivity of the coefficient of interest to the inclusion of further control variables (interacted with the *Post-March30* dummy), over the three-day and seven-day windows. In particular, we include in the regression district-level socio-economic and geographic controls (as described in footnote 26) and the following sets of additional controls: (i) additional geographic controls, (ii) variables capturing initial conditions, (iii) colonization controls, (iv) fractionalization controls, and (v) measures of availability of current trade routes (see footnote 31 for details on the variables included in each set). All results are qualitatively similar.

Altogether, the results support the existence of deeply-rooted determinants of the increase in anti-Muslim fake news after the *Tablighi* shock. This is in line with the historical narratives on the relevance of Muslim invasions and governance for local Indian history, and with present-day anecdotal evidence recalling the legacy of these experiences (see Section 2).

[38] Data on Hindu–Muslim conflicts during the 1950–2000 period (as reported in the newspaper *The Times of India*) were assembled by Ashutosh Varshney and Steve Wilkinson (2006) for the 1950–1995 period and extended to the year 2000 by Mitra and Ray (2014). We then geolocate all events and match them to present-day districts.

Table 4
Robustness.

Period:	Three days				Seven days			
Dependent variable:	N. Tweets FN				N. Tweets FN			
Specification:	N.Conflicts	Exposure	No Delhi	No Delhi Diffusion	N.Conflicts	Exposure	No Delhi	No Delhi Diffusion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Muslim attack × Post			1.64*** (0.56) [0.00]	1.52*** (0.53) [0.00]			1.07** (0.49) [0.03]	1.01** (0.47) [0.03]
N. Muslim attacks × Post	0.75** (0.35) [0.03]				0.58*** (0.21) [0.01]			
Exp. Muslim attack × Post		11.85*** (3.12) [0.00]				8.79*** (2.58) [0.00]		
Dist. New Delhi (×1000) × Post				−0.89*** (0.27) [0.00]				−0.58** (0.25) [0.02]
Std.Twitter Conn. to Delhi × Post				0.22*** (0.07) [0.00]				0.14** (0.07) [0.05]
Std.Born in Delhi × Post				1.72*** (0.38) [0.00]				1.39* (0.75) [0.06]
Time-variant controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.96	0.96	0.92	0.92	0.81	0.81	0.79	0.79
Observations	3756	3756	3750	3750	8764	8764	8750	8750

Notes: OLS estimates. Observations are districts in each day in two periods: March 28–April 2 in columns 1–4, and March 24–April 6 in columns 5–8. In all specifications the dependent variable is the number of tweets with anti-Muslim fake news. Time-variant controls are as in Table 3. Columns 1 and 5 replicate specifications 2 and 4 of Table 3, respectively, replacing the dummy for precolonial Muslim attacks with the number of Muslim attacks in the district. Columns 2 and 6 replicate specifications 2 and 4 of Table 3, respectively, replacing the dummy for precolonial Muslim attacks with a distance-based measure of exposure to Muslim attacks as in Dincecco et al. (2022). Columns 3–4 and 7–8 include our baseline dummy for precolonial Muslim Attacks and exclude Delhi from the set of observations. Columns 4 and 8 further account for Twitter connectedness and personal ties to Delhi. In all specifications, the newly included variables are interacted with the *Post-March30* dummy. See Section 3 for details on all variables. Standard errors in parentheses are clustered to account for spatial (up to 250 km) and serial correlation. P-values are reported in brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

6. Conclusion

Fake news spread rapidly on social media and the Internet. Given the large diffusion of these technologies in recent years, concern is growing over the spread of false stories and their social, economic, and political consequences.

In this paper, we study the diffusion of fake news under a novel perspective, namely as a vehicle to propagate hate and discriminatory attitudes towards minorities. In particular, we analyze the diffusion of false stories against Indian Muslims at the onset of the coronavirus outbreak in India, exploiting a tight sequence of events on March 30, 2020 that led many to identify a Muslim religious congregation (the *Tablighi Jamaat* convention) held in New Delhi as a COVID-19 hotspot. This coincided with an outburst of false stories reporting that Muslims were deliberately infecting other people and associating the spread of the virus with a form of jihad conducted by Muslim communities (see the trending hashtag "#coronajihad").

Using a comprehensive novel dataset of georeferenced text data from Twitter, we document a large spike (from nearly zero to above 3%) in the share of tweets reporting fake news against Muslims after March 30 and we investigate the spatial patterns of their diffusion. More specifically, we empirically test two main hypotheses: (i) whether fake news spread spatially, and (ii) whether local events that happened far back in history influence the dissemination of discriminatory false stories.

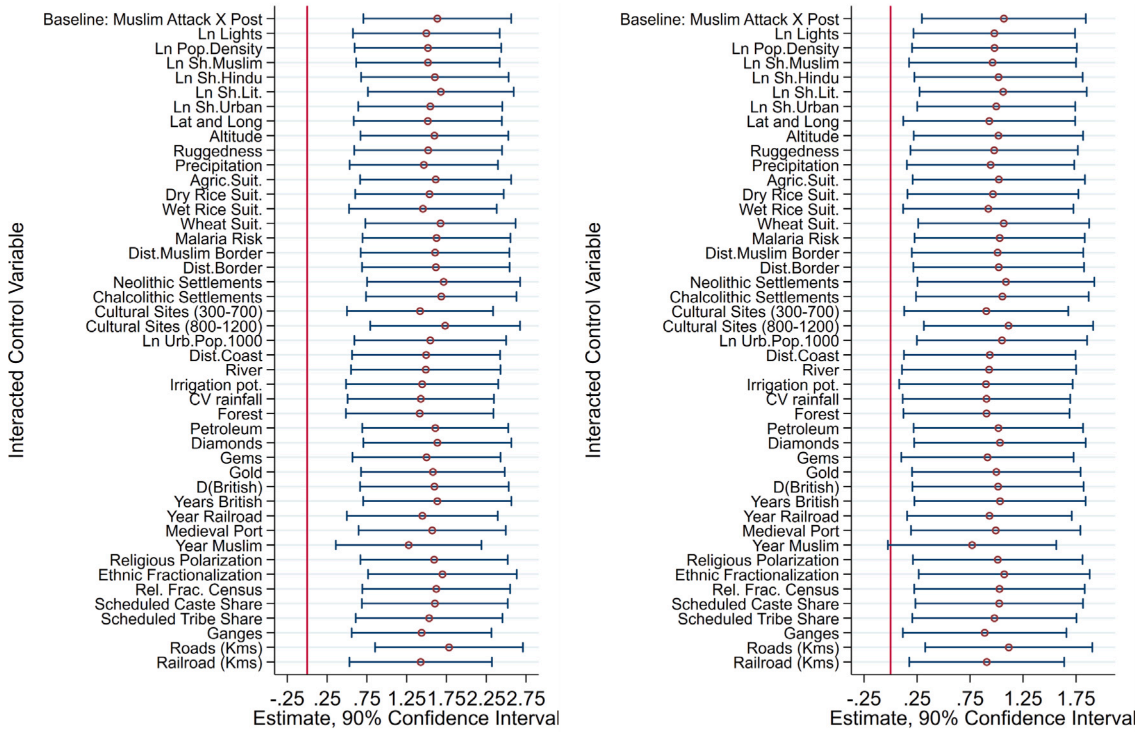
Our econometric analysis delivers two main results. First, in line with our first hypothesis, we find that, following the shock, the intensity of anti-Muslim false news was more pronounced in districts that are spatially closer to Delhi. Moreover, we explore the mechanisms behind the spatial diffusion of fake news and find that personal ties to Delhi may play a role. Second, in line with our second hypothesis, we show that the diffusion of fake news can also be linked to the legacy of precolonial Muslim attacks. In particular, districts that experienced historical Muslim attacks saw a larger diffusion of fake news compared to districts that did not experience any attack.

These findings on present-day India suggest that (epidemic) shocks may affect a country's overall environment of discrimination through the spread of anti-minority false news on social media. This is especially relevant in the case of (location-specific) persistent beliefs regarding the role of minorities as possible threats to national security and well-being.

It is still an open question how minorities react to the spread of fake news discriminating against them, whether they isolate themselves to protect their identity or whether they assimilate more into the local culture. More broadly, future research should also investigate more closely the dynamic interactions among minorities, members of the majority group, and political actors, who often contribute to the outbreak and diffusion of discriminatory false stories.

Appendix. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jue.2023.103613.



(a) 3 days window (b) 7 days window

Fig. 6. Sensitivity of Muslim attack to accounting for further control variables (Interacted by *Post*). *Notes*: The first dot displays the coefficient associated with the variable *Muslim Attack* × *Post* from estimating specifications 2 (panel a) and 4 (panel b) in Table 3 as baseline estimate. In each panel the other dots display the sensitivity of this estimate to further controlling for the variables indicated on the vertical axis (interacted with the *Post March 30* indicator). Horizontal bars indicate 90% confidence intervals. See footnotes 26 and 31 for details on all variables.

References

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., Zhuravskaya, E., 2015. Radio and the rise of the nazis in prewar Germany. Q. J. Econ. 130 (4), 1885–1939.

Anderson, D.M., Crost, B., Rees, D.I., 2020. Do economic downturns fuel racial animus? J. Econ. Behav. Organ. 175, 9–18.

Anderson, R.W., Johnson, N.D., Koyama, M., 2017. Jewish persecutions and weather shocks: 1100–1800. Econ. J. 127 (602), 924–958.

Aneez, Z., Neyazi, T.A., Kalogeropoulos, A., Nielsen, R.K., 2019. Reuters institute India digital news report. Accessed October 10, 2023 [Online], available at https://reutersinstitute.politics.ox.ac.uk/our-research/india-digital-news-report.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., Wong, A., 2018. Social connectedness: Measurement, determinants, and effects. J. Econ. Perspect. 32 (3), 259–280.

Bailey, M., Farrell, P., Kuchler, T., Stroebel, J., 2020. Social connectedness in urban areas. J. Urban Econ. 118, 103264.

Banerjee, A., Iyer, L., 2005. History, institutions, and economic performance: The legacy of colonial land tenure systems in India. Am. Econ. Rev. 95 (4), 1190–1213.

Bartoš, V., Bauer, M., Cahlíková, J., Chytilová, J., 2021. Covid-19 crisis and hostility against foreigners. Eur. Econ. Rev. 137, 103818.

Bentzen, J., Kaarsen, N., Wingender, A., 2017. Irrigation and Autocracy. J. Eur. Econom. Assoc. 15 (1), 1–53.

Bharadwaj, P., Khwaja, A.I., Mian, A., 2015. Population exchange and its impact on literacy, occupation and gender–evidence from the partition of India. Int. Migr. 53 (4), 90–106.

Bharadwaj, P., Mirza, R.A., 2019. Displacement and development: Long term impacts of population transfer in India. Explor. Econ. Hist. 73, 101273.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent Dirichlet Allocation. J. Mach. Learn. Res. 3 (Jan), 993–1022.

Bozarth, L., Quercia, D., Capra, L., Šćepanović, S., 2023. The role of the big geographic sort in online news circulation among US Reddit users. Sci. Rep. 13 (1), 6711.

Britannica, 2022. Mahmūd, king of Ghazna. Accessed April 14, 2022 [Online], available at https://www.britannica.com/biography/Mahmud-king-of-Ghazna ..

Buchel, K., Ehrlich, M.V., 2020. Cities and the structure of social interactions: Evidence from mobile phone data. J. Urban Econ. 119, 103276.

Castelló-Climent, A., Chaudhary, L., Mukhopadhyay, A., 2018. Higher education and prosperity: From Catholic missionaries to luminosity in India. Econ. J. 128 (616), 3039–3075.

Chandler, T., 1987. Four Thousand Years of Urban Growth: A Historical Census. St. David's University Press.

Chaudhary, L., Rubin, J., Iyer, S., Shrivastava, A., 2020. Culture and colonial legacy: Evidence from public goods games. J. Econ. Behav. Organ. 173, 107–129.

Comin, D.A., Dmitriev, M., Rossi-Hansberg, E., 2012. The Spatial Diffusion of Technology. NBER Working Paper No. w18534.

Conley, T.G., 1999. GMM estimation with cross sectional dependence. J. Econometrics 92 (1), 1–45.

Couttenier, M., Hatte, S., Thoenig, M., Vlachos, S., 2021. Anti-muslim voting and media coverage of immigrant crimes. Rev. Econ. Stat. 1–33.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., Zhuravskaya, E., 2014. Cross-border media and nationalism: Evidence from Serbian radio in Croatia. Am. Econ. J.: Appl. Econ. 6 (3), 103–132.

Dincecco, M., Fenske, J., Menon, A., Mukherjee, S., 2022. Pre-colonial warfare and long-run development in India. Econ. J. 132 (643), 981–1010.

Dipoppa, G., Grossman, G., Zonszein, S., 2021. Locked down, lashing out: Situational triggers and hateful behavior towards minority ethnic immigrants. Available at SSRN: https://ssrn.com/abstract=3789339.

Doerr, S., Gissler, S., Peydro, J.-L., Voth, H.-J., et al., 2021. Financial Crises and Political Radicalization: How Failing Banks Paved Hitler's Path to Power. Technical Report, Bank for International Settlements.

Eaton, R.M., 2000. Temple desecration and indo-muslim states. J. Islamic Stud. 11 (3), 283–319.

Fenske, J., Kala, N., Wei, J., 2021. Railways and Cities in India. Centre for Competitive Advantage in the Global Economy Working Paper 559.

Figueiredo, O., Guimarães, P., Woodward, D., 2015. Industry localization, distance decay, and knowledge spillovers: Following the patent paper trail. J. Urban Econ. 89, 21–31.

Gaspar, J., Glaeser, E., 1998. Information technology and the future of cities. J. Urban Econ. 43, 136–156.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., Lazer, D., 2019. Fake news on Twitter during the 2016 U.S. presidential election. Science 363 (6425), 374–378.

Guess, A.M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., et al., 2023. How do social media feed algorithms affect attitudes and behavior in an election campaign? Science 381 (6656), 398–404.

Guess, A., Nagler, J., Tucker, J., 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Sci. Adv. 5 (1), eaau4586.

Guess, A., Nyhan, B., Reifler, J., 2018. Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. Eur. Res. Counc. 9 (3), 4.

Halarnkar, S., 2020. Coronavirus is proving to be another excuse to marginalise India's Muslims. Accessed June 10, 2021 [Online], available at https://qz.com/india/1836768/coronavirus-is-another-excuse-to-marginalise-indias-muslims/..
Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., Wilson, C., 2013. Measuring personalization of web search. In: Proceedings of the 22nd International Conference on World Wide Web. pp. 527–538.
Henderson, R., Jaffe, A., Trajtenberg, M., 2005. Patent citations and the geography of knowledge spillovers: A reassessment: Comment. Amer. Econ. Rev. 95 (1), 461–464.
India Institute of Forest Management, 2015. India state of forest report.
Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., Puranen, B., et al., 2018. World Values Survey: Round Two - India. JD Systems Institute and WVSA Secretariat.
Ioannides, Y.M., Overman, H.G., Rossi-Hansberg, E., Schmidheiny, K., 2008. The effect of information and communication technologies on urban structure. Econ. Policy 23 (54), 202–242.
Iyer, L., 2010. Direct versus indirect colonial rule in India: Long-term consequences. Rev. Econ. Stat. 92 (4), 693–713.
Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Q. J. Econ. 108 (3), 577–598.
Jaques, T., 2007. Dictionary of Battles and Sieges: A Guide To 8,500 Battles from Antiquity Through the Twenty-First Century. Greenwood Press.
Jedwab, R., Johnson, N.D., Koyama, M., 2019. Negative shocks and mass persecutions: evidence from the Black Death. J. Econ. Growth 24 (4), 345–395.
Jha, S., 2013. Trade, institutions, and ethnic tolerance: Evidence from South Asia. Am. Political Sci. Rev. 107 (4), 806–832.
Lanzara, G., Lazzaroni, S., Masella, P., Squicciarini, M., 2023. The Dynamics of Discrimination and Assimilation: Theory and Evidence. Working Paper.
Levy, M., Goldenberg, J., 2014. The gravitational law of social interaction. Physica A 393, 418–426.
Lu, R., Sheng, Y., 2022. How racial animus forms and spreads: Evidence from the coronavirus pandemic. arXiv preprint arXiv:2007.01448.
Matsuura, K., Willmott, C., 2009. Terrestrial Air Temperature: 1900–2008 Gridded Monthly Time Series. Center for Climatic Research, University of Delaware.
Michalopoulos, S., Papaioannou, E., 2018. Spatial patterns of development: A meso approach. Annu. Rev. Econ. 10, 383–410.
Mitra, A., Ray, D., 2014. Implications of an economic theory of conflict: Hindu-Muslim violence in India. J. Polit. Econ. 122 (4), 719–765.
Mok, D., Wellman, B., Carrasco, J., 2010. Does distance matter in the age of the internet? Urban Stud. 47 (13), 2747–2783.
Mukhia, H., Chandra, B., Thapar, R., 2017. Communalism and the writing of Indian history.
Müller, K., Schwarz, C., 2020. From hashtag to hate crime: Twitter and anti-minority sentiment. Available at SSRN 3149103.
Müller, K., Schwarz, C., 2021. Fanning the flames of hate: Social media and hate crime. J. Eur. Econom. Assoc. 19 (4), 2131–2167.
Nag, P., 2007. Historical Atlas of India. NATMO.
Nielsen, 2022. Nielsen's bharat 2.0 study reveals a 45 users in rural India since 2019. Accessed May 12, 2022 [Online], available at https://www.nielsen.com/in/en/press-releases/2022/nielsens-bharat-2-0-study-reveals-a-45-growth-in-active-internet-users-in-rural-india-since-2019/.
Ochsner, C., Roesel, F., 2019. Mobilizing History. Technical Report, CERGE University Mimeo.
Outlook, 2020. Communal corona? Is it justified to blame Tablighi Jamaat for Nizamuddin outbreak?. Accessed April 14, 2022 [Online], available at https://www.outlookindia.com/website/story/india-news-corona-outbreak-afflicted-by-communal-virus-blaming-tablighi-jamaat-could-be-misdirected/349784.
Pariser, E., 2011. The Filter Bubble: How the New Personalized Web Is Changing What We Read and how We Think. Penguin.
Perrigo, B., 2020. It was already dangerous to be Muslim in India. Then came the coronavirus. Times.
Pew Research Center, 2021. Religion in India: Tolerance and segregation. Available at https://www.pewresearch.org/religion/2021/06/29/religion-in-india-tolerance-and-segregation..
Pooja Chaudari, Alt News, 2020. The year that was: Misinformation trends of 2020. Accessed April 14, 2022 [Online], available at https://www.altnews.in/the-year-that-was-misinformation-trends-of-2020/.
Press Trust of India, 2020. Nizamuddin congregation: Arvind Kejriwal orders FIR against maulana. Accessed April 14, 2022 [Online], available at https://www.indiatoday.in/india/story/nizamuddin-congregation-arvind-kejriwal-orders-fir-against-maulana-1661514-2020-03-30.
Ritika Jain, Article14, 2020. Covid-19: How fake news and Modi government messaging fuelled India's latest spiral of Islamophobia. Accessed April 14, 2022 [Online], available at https://scroll.in/article/959806/covid-19-how-fake-news-and-modi-government-messaging-fuelled-indias-latest-spiral-of-islamophobia.
Sachau, E.C., 2013. Alberuni's India: An Account of the Religion, Philosophy, Literature, Geography, Chronology, Astronomy, Customs, Laws and Astrology of India: Volume I. Routledge.
Sarkar, S.J., 1930. A-short history of Aurangzib.
Scellato, S., Mascolo, C., Musolesi, M., Latora, V., 2010. Distance matters: Geo-social Metrics for Online Social Networks. In: WOSN.
Schwartzberg, J., 1978. A Historical Atlas of South Asia. University of Chicago Press.
Soundararajan, T., Kumar, A., Nair, P., Greely, J., 2019. Facebook India towards the tipping point of violence: Caste and religious hate speech. Accessed October 10, 2023 [Online], available at https://www.equalitylabs.org/research/publications-resources/.
Sunstein, C.R., 2018. # Republic: Divided Democracy in the Age of Social Media. Princeton University Press.
Takhteyev, Y., Gruzd, A., Wellman, B., 2012. Geography of Twitter networks. Social Networks 34 (1), 73–81.
The Quint, 2021. 214% Rise in 'fake news' circulation cases in 2020: NCRB report. Accessed October 9, 2023 [Online], available at https://www.thequint.com/news/webqoof/214-percent-rise-in-fake-news-cases-in-2020-ncrb-report.
The Week, 2020. COVID-19: Centre blames Tablighi Jamaat for sudden spike in cases. Accessed May 12, 2022 [Online], available at https://www.theweek.in/news/india/2020/04/01/covid-19-centre-blames-tablighi-jamath-for-sudden-spike-in-cases.html.
Thompson, P., Fox-Kean, M., 2005. Patent citations and the geography of knowledge spillovers: A reassessment. Amer. Econ. Rev. 95 (1), 450–460.
Times of India, 2020. Coronavirus: About 9,000 Tablighi Jamaat Members, Primary Ontacts Quarantined in Country, Mha Says. Times of India.
Tollefsen, A., Strand, H., Buhaug, H., 2012. PRIO-GRID: A unified spatial data structure. J. Peace Res. 49 (2), 363–374.
Tranos, E., Nijkamp, P., 2013. The death of distance revisited: cyber-place, physical and relational proximities. J. Reg. Sci. 53 (5), 855–873, Special Issue: Mini-Special Issue: Regional Innovation Hotspots and Spatial Development.
Varshney, A., Wilkinson, S., 2006. Varshney-Wilkinson Dataset on Hindu-Muslim Violence in India, 1950–1995, Version 2. Inter-university Consortium for Political and Social Research [distributor], http://dx.doi.org/10.3886/ICPSR04342.v1, 2006-02-17.
Verghese, A., 2016. The Colonial Origins of Ethnic Violence in India. Stanford University Press.
Verghese, A., 2021. Ethnic violence in India 1. In: Routledge Handbook of Race and Ethnicity in Asia. Routledge, pp. 31–45.
Voigtländer, N., Voth, H.-J., 2012. Persecution perpetuated: the medieval origins of anti-Semitic violence in Nazi Germany. Q. J. Econ. 127 (3), 1339–1392.
Vosoughi, S., Roy, D., Aral, S., 2018. The spread of true and false news online. Science 359 (6380), 1146–1151.
Voth, H.-J., 2021. Persistence–myth and mystery. In: The Handbook of Historical Economics. Elsevier, pp. 243–267.
Wolpert, S.A., 2004. A New History of India. Oxford University Press, USA.
Yanagizawa-Drott, D., 2014. Propaganda and conflict: Evidence from the Rwandan genocide. Q. J. Econ. 129 (4), 1947–1994.
Zhuravskaya, E., Petrova, M., Enikolopov, R., 2020. Political effects of the internet and social media. Annu. Rev. Econ. 12, 415–438.