

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active. Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

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

How racial animus forms and spreads: Evidence from the coronavirus pandemic^{\star}

Runjing Lu^{a,*}, Sophie Yanying Sheng^b

^a Alberta School of Business, University of Alberta, 11211 Saskatchewan Dr NW, Edmonton, AB T6G 2R6, Canada ^b Economics Department, University of California San Diego, 9500 Gilman Dr, La Jolla, CA 92093, United States

ARTICLE INFO

Article history: Received 6 March 2021 Revised 26 April 2022 Accepted 20 May 2022 Available online 27 May 2022

- JEL classification: J15 D74 I18
- Keywords: COVID-19 Racial animus Xenophobia Social media Hate crime

ABSTRACT

This paper studies the formation and the spread of crisis-driven racial animus during the coronavirus pandemic. Exploiting plausibly exogenous variation in the timing of the first COVID-19 diagnosis across US areas, we find that the first local case leads to an immediate increase in local anti-Asian animus, as measured by Google searches and Twitter posts that include a commonly used derogatory racial epithet. This rise in animus specifically targets Asians and mainly comes from users who use the epithet for the first time. These first-time ch-word users are more likely to have expressed animosity against non-Asian minorities in the past, and their interaction with other anti-Asian individuals predicts the timing of their first ch-word tweets. Moreover, online animosity and offline hate incidents against Asians both increase with the salience of the connection between China and COVID-19; while the increase in racial animus is not associated with the local economic impact of the pandemic. Finally, the pandemic-driven racial animus we documented may persist beyond the duration of the pandemic, as most racist tweets do not explicitly mention the virus.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Racial animus can affect welfare in measurable ways, as economists have noted since the seminal work of Becker (1957). Recent papers have shown that racial animus can hinder economic development, affect political institutions, and induce social unrest (e.g., Charles and Guryan, 2008; Card et al., 2008; Stephens-Davidowitz, 2014; Healy and Searcey, 2020). To curb racial animus at the outset and to mitigate its consequences, a crucial first step is to understand how it forms and grows.

In this paper, we shed light on what factors motivate racial animus, which individuals are more susceptible to such factors, and how racial animus spreads, using the coronavirus (COVID-19) pandemic as a natural experiment. The Centers for Disease Control and Prevention (CDC) has emphasized that people of Asian descent are at no greater risk of spreading the virus than other Americans. Nonetheless, since the outbreak of the virus, news reports of hate crimes against Asian



JOURNAL O Economi

Behavior & Organization



^{*} An earlier version of this paper circulated under the title "From Fear to Hate: How A Pandemic Sparks Racial Animus in the United States." We are grateful for the support and guidance from Eli Berman, Julie Cullen, and Gordon Dahl. We would also like to thank David Arnold, Julian Betts, Georgy Egorov, Itzik Fadlon, Ruixue Jia, William Mullins, Jacob Shapiro, Jesse Shapiro, and the participants at the UCSD applied lunch seminar, AEA, and SOLE for their constructive feedback.

^{*} Corresponding author.

E-mail addresses: runjing1@ualberta.ca (R. Lu), yas107@ucsd.edu (S.Y. Sheng).

Americans have increased (Mullis and Glenn, 2020).¹ The unexpected nature and regional variation of the pandemic provide a valuable opportunity to study the rise and spread of racial animus – in this case, against Asians.

To proxy for an area's racial animus against Asians, we use the percentage of Google searches and Twitter posts (tweets) that include the words "chink" or "chinks" (hereafter, the ch-word).² Google searches can capture *private* racial animus given others cannot view one's searches. Past papers have documented a clear relationship between Google searches of racial slurs and racial animus against minorities (Anderson et al., 2020; Depetris-Chauvin, 2015; Stephens-Davidowitz, 2014). Furthermore, as we will show below, an area's monthly Google searches for the epithet are positively correlated with monthly anti-Asian hate crimes and are negatively correlated with monthly visits to Chinese restaurants. Our second proxy is based on tweets, which has been used to measure *public* displays of racial animus (Nguyen et al., 2018). These two proxies are valuable alternatives to more traditional measures, such as offline hate crimes which may only capture the most extreme hatred and may not fully reflect the levels of racial animus due to blanket stay-at-home orders during the pandemic. In addition, use of racial slurs online is an important outcome in and of itself, as researchers have shown a strong relationship between exposure to racial discrimination online and depression and anxiety measured offline (Tynes et al., 2008).

To motivate, we exploit the timeline of COVID-19 developments in the United States to understand the general evolution of anti-Asian animus during the pandemic. We find little increase in the national racial animus upon the first US COVID-19 case and only a small uptick in the week when the World Health Organization (WHO) declared COVID-19 a pandemic. In contrast, we observe a clear jump in the week when President Trump tweeted "Chinese virus."

In order to causally identify the effects of COVID-19 on racial animus against Asians, we use a difference-in-differences (DID) event study design exploiting the variation in the timing of the *first local* COVID-19 diagnosis across areas. Specifically, we compare the change in racial animus following the first diagnosis in an area to the change in other areas during the same period. First local diagnoses are likely to increase the salience of the virus, and the salience of diseases has been shown to induce xenophobia in lab experiments (Faulkner et al., 2004). The identifying assumption is that the *precise* timing of the first diagnosis in an area is plausibly exogenous; whether an area has its first diagnosis this week (day) or the next is largely unpredictable and unlikely to correlate with other factors that simultaneously change local racial animus.³

Our DID event study reveals that, in the week after the first local COVID-19 diagnosis, an area's Google search rate of the ch-word increases by 22.6% of the area's maximum search rate during the sample period, and an area's Twitter post rate of the epithet increases by 118.6% of the average post rate across all areas during the sample period. These effects persist for at least six weeks after the first local case. Given the correlation based on historical data, the increase in Google search rate of the ch-word would be associated with a 6.5% increase in anti-Asian hate crimes holding everything else constant. The results, where applicable, are quantitatively unchanged under a dynamic event study design which allows for varying treatment effects across event periods (see Sun and Abraham, 2020). Our results are also robust to using alternative racial animus measures based on tweets which include other anti-Asian slurs and are not counter-hate; to excluding early- and hard-hit states; and to controlling for severity of local infection, existence of stay-at-home orders, general local attention to Asians, and area and year-month fixed effects.

When we examine the content of ch-word tweets, we find that the share showing emotions of anger and disgust increases from 23.3 to 40.8% after the first local diagnosis. This shift in sentiment suggests that the increase in racially charged tweets represents a real change in attitude towards Asians. Moreover, the increase in racial animus is directed *only* at Asians and not at other minorities. The singling out of Asians implies that the increase is likely not due to an overall rise of ethnic distrust or tensions from general uncertainty about cross-group differences in health status or risk-taking behavior. Rather, it is targeted at a specific group associated with the geographical origin of the virus. In addition, 75% of ch-word tweets posted following the first local case do not explicitly mention COVID-19, implying that the pandemic-induced racial animus towards Asians extends to broader topics and may persist beyond the duration of the pandemic.

We also leverage the rich information in historical tweets and Twitter user network to study *which individuals* are more likely to start expressing hate because of the pandemic. We find that the surge in ch-word tweets is driven primarily by the extensive margin (i.e., existing Twitter users who post the term for the first time) rather than the intensive margin (i.e., increase in tweets from users who have previously used the term). These first-time ch-word users are 40% of the mean more likely than never users to have tweeted racial slurs against non-Asian minorities in the past, implying that the pandemic may have redirected their anti-minority sentiment towards Asians. They are also 58 and 28% of the mean more likely to list "Trump" and "politics" in their user profiles.

Finally, we turn our attention to the factors fueling the spread of racial animus among individuals. Exposure to anti-Asian users is one such factor. We find that interacting with anti-Asian users in a day predicts a 22% higher likelihood (relative to the mean) of tweeting the ch-word the next day. The salience of the connection between COVID-19 and the Asian population is another factor. We proxy for this salience by using the number of President Trump's tweets that mention China and COVID-19 simultaneously. We find that one additional such China-and-COVID tweet in a day corresponds to an

¹ For example, see NBC News, New York Times, and USA Today.

² We focus on the use of the ch-word because it is the most salient and unambiguously pejorative racial slur against Asians. According to the Philadelphia Bar Association, the epithet "is now widely used throughout the United States as a racial slur against people of Asian descent" (Association). Importantly, it has not been reclaimed by the Asian American community (Anderson and Lepore, 2013).

³ Papers like Egorov et al., 2020 have noted that areas with larger population sizes or better medical systems tend to have first diagnoses earlier. We include area fixed effects to control for these time-invariant characteristics.

eight percent increase in anti-Asian hate incidents and an increase in national ch-word tweets on the same day which is equivalent to 14% of the daily average. An event study using hourly tweet data also reveals an immediate increase in ch-word tweets following the president's China-and-COVID tweets but not before. In contrast, we find little evidence that negative economic impacts from the pandemic motivates the initial rise of racial animus. Areas with a more severe economic damage from the pandemic do not exhibit a higher increase in racial animus than areas with a less severe impact.

This paper contributes to the literature studying the causes of animus toward minorities. This body of work has shown that negative shocks such as terrorist attacks and deterioration of economic conditions induce animus against racial or religious minorities. For example, Kaushal et al. (2007), Hanes and Machin (2014), and Ivandic et al. (2019) document that 9/11 and jihadi terror attacks lead to increases in anti-Muslim hate crimes. Anderson et al. (2017) and Anderson et al. (2020) find that the Great Recession and negative shocks to agricultural income in historical Europe contribute to animus against minorities. In addition, desire to avoid health threats has also been postulated to motivate racial bias (Schaller and Neuberg, 2012). Lab experiments have shown that exposing subjects to disease-related primes leads to increased xenophobia (Faulkner et al., 2004; O'Shea et al., 2020; Bartos et al., 2020). However, the causal evidence on whether infectious diseases lead to racial animus in the field is still lacking.⁴ An exception is Jedwab et al. (2019), which documents that the black death caused an increase in anti-Jewish pogroms in medieval Europe.

Our contribution is to provide causal evidence on how negative shocks, such as pandemics, trigger racial animus, and shed light on who are more susceptible to such shocks and how racial animus spreads. Our findings have implications for mitigating animus amid future crises. We find that the rise in racial animus is specific to Asians who are associated with the geographical origin of the virus and that the salience of this association amplifies animus against the group. Therefore, careful naming of a disease (e.g., COVID-19 and Delta variant as opposed to Chinese virus and Indian variant) and debunking claims of any purported connection between a disease and a group could be helpful in curbing animus. Additionally, our findings reveal that the extensive margin and social media play an important role in spreading racial animus, suggesting that moderating racist individuals and their interaction with others on social media could help constrain racial animus in the future.

Finally, our paper speaks to the literature on political rhetoric. Political rhetoric has been shown to influence public opinion and behavior, such as presidential approval (Druckman and Holmes, 2004), public perception of a foreign country (Silver, 2016), and anti-minority hate crimes (Müller and Schwarz, 2019). We add to this literature by providing another example of how the rhetoric of political figures regarding a public crisis influences racial animus at the national level. Harnessing the opinion-shaping power of these public figures could be useful in curbing animus.

2. Measures of racial animus

2.1. Google and twitter proxies

We use two measures to proxy for an area's racial animus against Asians: the percentage of Google searches and the percentage of tweets that include the words "chink" or "chinks." The ch-word is not uncommon in Google searches or tweets. Between June 2019 and June 2020, this racial epithet was included in more than a quarter million searches and 60,000 tweets.⁵ Google searches and tweets that include the epithet are mostly negative. For instance, "chink eye" and "chink virus" are common terms in such Google queries and Twitter posts. People may search the epithet to look for jokes or memes about Asians or to look for like-minded others with whom they can share anti-Asian sentiments.

We use Google Trends to obtain weekly Google search data for the ch-word at the media market level between July 2019 and April 2020. The data are not the raw number of searches but the weekly percentage of searches that include the term (*search rate*), taken from a random sample of total searches representative at the media market and time levels and scaled by the highest weekly search rate in the same market during the entire extraction period – in our case, between July 2019 and April 2020. In particular, the racially charged Google search index for media market *m* at time *t* extracted over period *T* is

Google search index_{*mt,T*} = 100 ×
$$\frac{\frac{\text{Searches including "chink(s)''_{mt}}}{\text{Total searches_{mt}}}}{\max_{t \in T} \{\frac{\text{Searches including "chink(s)''_{mt}}}{\text{Total searches_{mt}}}\}}$$
(1)

Note that Google returns a zero value when the racially charged search index for a given area and time falls below an unreported threshold. Because of this, we only include media markets that have a valid racially charged Google search index in the baseline period (2014–2018) in our sample. This leaves us with 60 of 210 media markets, covering approximately 74% of the 2019 US population and 78% of the 2019 US GDP across 33 states. Compared to other media markets, the ones in our sample tend to have a larger population, higher percentage of Asians, slightly lower baseline anti-Asian hate crime rate,

⁴ More recent papers on the prevalence of hate during the COVID-19 pandemic are mostly descriptive (e.g., Croucher et al., 2020; Lyu et al., 2020, Schild, Ling, Blackburn, Stringhini, Zhang, Zannettou, Ziems, He, Soni, Kumar) or take a structural approach (e.g., Deng and Hwang, 2021).

⁵ The number of Google searches is an approximation from https://searchvolume.io/. The data are only available for the 12-month period before our query on June 8, 2020.

and more enplanements of international airports (Table A1 column (1)). Shaded areas in Fig. A1 panel A indicate the media markets in our sample. Analyses using Google data are conducted at the media market level.

The above metric can capture the timing but not the level of a change in an area's search index. As an alternative, we rescale the Google search index so that the search rate in different media markets is normalized by *one* base search rate. We try three different bases: Huntsville-Decatur (Florence)'s search rate on March 15, 2020; Wilkes Barre-Scranton on March 29, 2020; and Buffalo on April 5, 2020. We choose these bases to obtain rescaled indexes for as many media markets as possible, i.e., 35, 29, and 29, respectively. As detailed in Appendix 1, rescaling drops many media markets whose search rate is zero on the date when the base search rate occurs (*benchmark date*). For this reason, we only use the rescaled version as a robustness check.

We obtain Twitter data from Crimson Hexagon, which houses all public tweets through a direct partnership with Twitter. We download all geo-located tweets that include the ch-word between November 1, 2019, and May 2, 2020. Crimson Hexagon does not provide the total number of tweets posted in a given area and time. We thus extract the number of all public tweets that include the word "the," the most common word on Twitter, in a given area and time as a substitute. Assuming that the proportion of tweets that include "the" is stable across areas, the number of tweets that include "the" can approximate overall Twitter activity. We define the racially charged Twitter post index for a given area and time as the number of tweets including the ch-word per 100,000 tweets including the word "the."

We calculate the Twitter post index for 658 counties across 50 states and Washington D.C., encompassing 60% of the US population and nearly 70% of the US GDP in 2019. Counties are included if their residents ever posted "the" tweets between 2014 and 2018.⁶ Counties with Twitter data tend to have a larger population, higher support for the Democratic Party, and higher enplanements of international airports, but show no difference in baseline anti-Asian hate crime rate compared to other counties (Table A1 column (2)). Shaded areas in Fig. A1 panel B are counties with Twitter data. Analyses using Twitter data are conducted at the county level unless noted otherwise.

The fact that Google and Twitter data do not cover the full of the US should not affect internal validity of our study, but it could pose a threat to external validity. This is why we use *both* data sources, which could alleviate concerns about the external validity of our findings.

2.2. Relationship between racial animus, hate crimes, and restaurant visits

For the racially charged Google search index and Twitter post index to be meaningful proxies for racial animus, the only assumption we need is that an increase in racial animus makes a person more likely to use the ch-word. Under this assumption, higher racial animus results in a higher percentage of searches and tweets that include the racial epithet. Existing papers that use similar proxies for racial animus suggest that the assumption is likely to hold (Anderson et al., 2020; Depetris-Chauvin, 2015; Stephens-Davidowitz, 2014).

To better understand the above proxies, we check how they predict anti-Asian hate crimes and visits to Chinese restaurants. Hate crime data come from the FBI Uniform Crime Reports (UCR) and are available up to 2018. A majority of these hate crimes are simple or aggravated assault (30%) and in-person intimidation (34%). Table 1, panel A, columns (1) through (4) report the media market-level correlation between the monthly Google search index and the monthly number of anti-Asian hate crimes between January 2014 and December 2018, controlling for local population size, unemployment rate, year-month fixed effects, and media market fixed effects. On average, a one-standard-deviation increase in the Google search index corresponds to an increase in the anti-Asian hate crimes in the same month, amounting to 8.9% of the monthly average.⁷ The correlation is robust to controlling for the Google search index for "Asian(s)," which is related to the ch-word but neutral in connotation, as shown in column (2). In columns (3) and (4), we include both the index in the current month and the index in the prior month. The relationship between the Google search index and hate crimes is mainly contemporaneous.

Next, we change the dependent variable to monthly visits to Chinese restaurants in a media market between January 2018 and December 2019, additionally controlling for the monthly visits to all local restaurants. The visit data are from Safegraph and are available starting in 2018.⁸ Table 1, panel A, columns (5) and (6) show that a one-standard-deviation increase in the Google search index is linked to 484 fewer monthly visits to Chinese restaurants, equaling 0.5% of the monthly average. The relationship between the Google search index and visit rate is also contemporaneous.

Finally, we replicate the above correlations using Twitter data in Table 1, panel B. We aggregate hate crimes to the media market level due to their low occurrences at the county level. To maintain consistency, we also aggregate restaurant visits to the media market level. Overall, the Twitter post index does not correlate with anti-Asian hate crimes or visits to Chinese restaurants. One potential explanation is that Twitter data represent public displays of racial animus and undergo more social censoring. We may only see a change on Twitter when the shift in racial animus is substantially large.

⁶ About half of the tweets in the sample lack geo-identifiers and hence cannot be associated with a certain county.

⁷ The percent increase is calculated by multiplying the standard deviation of the index (23.07) with the coefficient and dividing the product with the outcome mean.

⁸ Safegraph provides data on foot traffic to roughly 4.1 million points of interest in the United States.

Table 1

Relationship between Racial Animus, Hate Crimes, and Chinese Restaurant Visits.

VARIABLES	(1) Incidents	(2) Incidents	(3) Incidents	(4) Incidents	(5) Visits	(6) Visits	(7) Visits	(8) Visits
Panel A: Google search index Google ch-word(t)	0.00057* (0.00031)	0.00057* (0.00031)	0.00057* (0.00031)	0.00057* (0.00031)	-21.069^{*}	-22.017^{*}	-23.489** (11.632)	-24.381** (11.697)
Google ch-word $(t-1)$	()	()	-0.00019 (0.00035)	-0.00018 (0.00035)	()	()	-13.934 (11.467)	-14.900 (11.594)
Total visits					0.052*** (0.005)	0.052*** (0.005)	0.048*** (0.006)	0.048*** (0.006)
Population(m)	0.51926*** (0.19939)	0.51907*** (0.19942)	0.52042*** (0.19936)	0.52287*** (0.19991)				
Google Asian(s)(t)		-0.00005 (0.00109)		-0.00018 (0.00111)		-138.636** (60.169)		-120.800** (54.317)
Google Asian(s)($t - 1$)				0.00075 (0.00163)				10.816 (59.669)
Observations	3600	3600	3600	3600	1440	1440	1380	1380
R-squared	0.309	0.309	0.309	0.309	0.996	0.996	0.997	0.997
Outcome mean	0.147	0.147	0.147	0.147	104962.736	104962.736	104962.736	104962.736
Panel B: Twitter post index	0.000.11	0.00007	0.000.40	0.000.47	60.040	64.400	22.002	20.470
lwitter ch-word	-0.00041 (0.00099)	-0.00037 (0.00096)	-0.00048 (0.00099)	-0.00047 (0.00100)	-60.249 (49.798)	-61.190 (50.150)	-30.883 (42.641)	-30.178 (42.737)
Twitter ch-word $(t - 1)$			-0.00063 (0.00065)	-0.00064 (0.00065)			-10.216 (36.060)	-10.089 (36.286)
Total visits			. ,	. ,	0.047*** (0.005)	0.047*** (0.005)	0.044*** (0.006)	0.044*** (0.006)
Population(m)	0.56488*** (0.17387)	0.56517*** (0.17391)	0.58097*** (0.17623)	0.58136*** (0.17630)				
Twitter Asian(s)(t)	(,	-0.00003		-0.00002		0.294 (0.682)		-0.279
Twitter Asian(s)($t - 1$)				-0.00001 (0.00002)				-0.266 (0.693)
Observations	11,116	11,116	10,921	10,921	4493	4493	4300	4300
R-squared	0.220	0.220	0.220	0.230	0.996	0.996	0.997	0.997
Outcome mean	0.057	0.057	0.057	0.057	41065.784	41065.784	41065.784	41065.784

Notes: The table presents the relationship between the racially charged Google search index and the Twitter post index, anti-Asian hate crimes, and visits to Chinese restaurants. Hate crime data are from the FBI UCR, visit data are from Safegraph, and all data are at the media market×year-month level. Outcome variables are the monthly number of anti-Asian hate crimes between January 2014 and December 2018 (columns (1)-(4)) and the monthly number of visits to Chinese restaurants between January 2018 and December 2019 (columns (5)-(8)). *Google Asian(s)* is the Google search index for the word "Asian(s)." *Twitter Asian(s)* is the number of tweets including "Asian(s)" per 100,000 "the" tweets. All regressions control for local unemployment rate, year-month fixed effects, and media market fixed effects. *** p < 0.01, ** p < 0.1.

3. Evolution of racial animus amid the pandemic

To motivate, we study the general evolution of anti-Asian animus as the pandemic develops. An ideal experiment would be to contrast rates of racially charged Twitter posts and Google searches in the U.S. during the pandemic to counterfactual rates in the absence of the pandemic. However, a perfect counterfactual does not exist because all individuals and areas were more or less impacted by the pandemic. For this reason, we use racially charged Twitter posts and Google searches in 2019 as controls. The assumption is that the searches and tweets in 2020 would have been the same as in 2019 absent the pandemic.⁹

We begin by comparing an individual's weekly likelihood of tweeting the ch-word during the first 16 full weeks in 2020 and the same person's likelihood of doing so in the corresponding weeks in 2019. An advantage of this analysis is that it does not require geo-identifiers, so we can include all 26,065 Twitter users who ever tweeted the ch-word between 2014 and 2018.¹⁰ We use the following specification:

$$Y_{iyw} = \sum_{w=2}^{16} \beta_w \times 1\{y = 2020\} + \alpha_i + \alpha_w + \epsilon_{iyw}$$
(2)

where Y_{iyw} is a binary variable which equals one if individual *i* tweets the ch-word in week *w* of year *y*. We use w = 1, the first full week of a year, as the comparison period. Our treatment variable is $1\{y = 2020\}$, which equals one if the year is 2020, and 0 if the year is 2019. We include person fixed effects α_i and week-of-year fixed effects α_w to absorb individual baseline propensity to tweet the racial epithet and the seasonality in such tweets. We cluster standard errors by individual.

⁹ This assumption could be violated if there are other contemporaneous shocks affecting racial animus. Our strategy in the next section avoids this issue.

¹⁰ We cannot look at the universe of Twitter users because Crimson Hexagon only allows tweet extraction based on keywords.

The individual-level analysis reveals that the likelihood of tweeting the ch-word co-moves with important developments of COVID-19. In Fig. A2 panel A, we plot β_w from Eq. (2) and the timing of major COVID-19 developments. While we find little to no increase in the likelihood of tweeting the term following the first US COVID-19 case or declarations of health emergency and only a small uptick in the week when the WHO declared COVID-19 a pandemic, we observe a clear jump in the week when President Trump first tweeted "Chinese virus."

Lacking individual-level search data, we compare media market-level weekly racially charged Google search index in 2020 and the index in the corresponding markets and weeks in 2019. Specification is the same as Eq. (2) except that Y_{iyw} is now the Google search index in media market *i* in week *w* of year *y*. We plot β_w in Fig. A2 panel B. While we also see a spike in the Google search index in the week when President Trump tweeted "Chinese virus," we cannot draw definitive conclusions for other weeks.

4. Evidence from DID event study

We now turn to our main empirical strategy, a DID event study design exploiting the variation in the precise timing of the first COVID-19 diagnoses across the United States. We compare the changes in racially charged Google search index (or Twitter post index) in the weeks before and after the first local case to the changes in other media markets (or counties) during the same period. This design allows us to avoid concerns about contemporaneous shocks that influence racial animus at the same time as the pandemic develops.

4.1. Data and empirical strategy

We download the data on US COVID-19 cases and deaths between January 21, and May 2, 2020, from the Johns Hopkins University Coronavirus Resource Center. We match the date of the first case in each media market and county to those with valid Google and Twitter data. Table A2 displays the number of media markets and counties by the timing of their first local diagnoses. All media markets have their first diagnoses in the sample period and have Google data for at least six weeks after the diagnosis. These media markets make up the Google sample. Seventeen counties with Twitter data are excluded because they did not have diagnoses in the sample period.¹¹ The remaining 641 counties make up the Twitter sample and have data for at least one week after the first local diagnosis; the number of counties decreases to 636, 629, 613, 555, 416 in weeks two to six.¹² Therefore, the Google (or Twitter) sample is a panel of media markets (or counties) from six weeks before to six weeks after the first local diagnosis. Table A3 reports summary statistics for each of the samples.

To understand predictors of the diagnosis timing, we regress the week of first local diagnosis on a battery of local characteristics in Table A4. The analysis reveals that a larger population size predicts earlier diagnoses for both the Google and the Twitter samples, while enplanements of international airports predict slightly later diagnoses for the Google sample. However, the proportion of Asians does not have predictive power for the timing, consistent with the CDC's statement that Asians are at no greater risk of spreading the virus. More importantly, pre-pandemic anti-Asian hate crime rate does not predict the timing, suggesting that the treatment timing is orthogonal to baseline racial animus. We then estimate the following regression:

$$Y_{it} = \sum_{t=-6, t\neq -1}^{6} \beta_t + \gamma' X_{it} + \alpha_i + \alpha_{ym(t)} + \epsilon_{it}$$
(3)

where Y_{it} is the racially charged Google search index (Twitter post index) in media market (county) *i* in event time *t*, which is the number of time periods relative to the first local diagnosis. β_t represents event dummies for six weeks before to six weeks after the first local diagnosis, excluding our comparison period t = -1. X_{it} is a vector of area-specific time-varying characteristics such as the local number of COVID-19 diagnoses or deaths, an indicator for a state-level stay-at-home order, and the Google search index or Twitter post index for "Asian(s)". We include county or media market fixed effects α_i and year-month fixed effects $\alpha_{ym(t)}$ to control for an area's baseline racial animus and national trends in racial animus. We cluster standard errors by media market for Google data and by county for Twitter data. We also estimate Eq. (3) at the daily level, where we include event dummies from 14 days before to 21 days after the first local diagnosis while omitting the dummy for day -4 and additionally control for day-of-week fixed effects.

If the trends of racially charged Google search index (or Twitter post index) across media markets (or counties) are parallel in the absence of local COVID-19 cases, and the treatment effect of the first local case does not vary across event times, β_t identifies the weighted average treatment effect across treatment areas on local searches or posts of the ch-word in time *t*. Testing for parallel pre-trends can shed light on the first identifying assumption. As we will show, this assumption appears to hold. The second assumption is harder to test, and its violation could bias the estimates in unknown directions. For example, if earlier treated areas experience an increasing (or decreasing) treatment effect over time due to evolving local pandemic situations, using these areas as controls for later treated places could bias the average treatment effect downward (upward).

¹¹ Results are unaffected when we include these counties.

¹² Crimson Hexagon was discontinued in July, 2020, so we cannot extend the Twitter sample.

		0	8	
	(1)	(2)	(3)	(4)
VARIABLES	Ch-word index	Severity control	Asian control	Exclude states
-6w	-3.920	-2.694	-4.265	-8.979
	(6.379)	(6.620)	(6.404)	(8.341)
-5w	0.431	1.100	-0.198	-2.575
	(5.722)	(5.820)	(5.699)	(7.083)
-4w	9.764	10.088	9.419	9.205
	(6.263)	(6.316)	(6.233)	(7.649)
-3w	2.282	2.503	2.247	2.458
	(5.023)	(5.085)	(5.020)	(5.912)
-2w	4.739	4.899	4.771	2.564
	(5.469)	(5.535)	(5.467)	(6.150)
+0w	6.421	6.326	6.274	6.574
	(4.898)	(4.911)	(4.864)	(5.127)
+1w	22.628***	22.442***	22.030***	22.771***
	(5.210)	(5.246)	(5.280)	(5.721)
+2w	16.945***	15.936***	16.727***	18.104***
	(5.439)	(5.443)	(5.407)	(5.621)
+3w	8.155	5.702	7.894	8.614
	(5.359)	(5.907)	(5.403)	(5.829)
+4w	19.106***	15.972**	18.873***	19.527**
	(6.265)	(6.999)	(6.253)	(7.461)
+5w	18.263**	15.375*	18.041**	14.709*
	(7.411)	(8.113)	(7.428)	(8.679)
+6w	17.861**	15.002*	18.125**	18.017*
	(7.726)	(8.046)	(7.751)	(9.267)
Observations	780	780	780	663
R-squared	0.190	0.192	0.193	0.180
Outcome mean	30.03	30.03	30.03	30.03

 Table 2

 The effect of the first local COVID-19 diagnosis on racial animus Google Search Index.

Notes: The table presents the effect of the first local COVID-19 diagnosis on the racially charged Google search index. All columns report the estimates of coefficients on the event dummies in Eq. (3). Column (1) corresponds to Fig. A4, panel A. Column (2) controls for the number of COVID-related new cases and deaths, and whether the state has any stay-athome orders in place. Column (3) controls for the Google search index for "Asian(s)." Column (4) excludes Washington, New York, and California. All regressions control for media market fixed effects and year-month fixed effects. Standard errors are clustered by media market. *** p < 0.01, ** p < 0.05, * p < 0.1.

To alleviate concerns about time-varying treatment effects, we use a *dynamic* DID event study comparing areas with a first case before and after the case, using areas that *have not had* any cases as controls. To implement the dynamic event study, we follow Novgorodsky and Setzler (2019) and stack our data as a series of 2×2 matrices (treated/not-yet-treated \times pre/post). We define areas which have their first cases in calendar week *g* as cohort *g*, and cohort-specific event time in calendar month *m* as $e_g = m - g$. The treatment effect on cohort *g* in event time e_g is labeled as β_{e_g} . Following Sun and Abraham (2020), we define the average treatment effect for event time *e* among all cohorts *G* as:

$$\beta_e = \sum_{g \in G} \beta_{e_g} \times w_g \tag{4}$$

where the aggregation weight w_g is the population in areas belonging to cohort *g*. We calculate clustered standard errors at the area level for β_e via the delta method.

A limitation of the dynamic event study is that it requires enough not-yet-treated areas in event time e_g to estimate β_{e_g} . Since counties are smaller than media markets, there is more variation in the timing of first local diagnoses at the county level. As a result, most counties in the Twitter sample have control counties for multiple post periods while most media markets in the Google sample have none after event 0. Therefore, we only apply the dynamic event study to Twitter data and use this approach as a robustness check.

4.2. Effects of the first local case on racial animus

4.2.1. Main findings

We start by examining how an area's Google searches for the ch-word respond to the first COVID-19 case in the local area. Fig. 1, panel A plots β_w from Eq. (3) using an area's racially charged Google search index as the outcome. The Google search index jumps markedly in the week after the first local case and persists at high levels in the following weeks. The pre-trends are flat and statistically insignificant, suggesting that the parallel trend assumption is likely to hold. Regression results corresponding to this figure are found in Table 2, column (1). For example, consider the +1w coefficient: compared to the week before the first local case, in the first week, an area's racially charged search rate increases by 22.6% of the



B: Twitter post index

Fig. 1. The Effect of the First Local COVID-19 Diagnosis on Racial Animus. *Note:* The figure presents the effect of the first local COVID-19 diagnosis on the racially charged Google search index and Twitter post index. See Section 2.1 for the definitions of the indexes. Panels A and B plot the estimates and 95% confidence intervals of the coefficients on the event dummies in Eq. (3) using the Google search index and the Twitter post index as the outcome, respectively. The estimates in panels A and B correspond to column (1) of Table 2 and column (1) of Table 3. Regressions control for year-month fixed effects and media market (panel A) or county (panel B) fixed effects. Standard errors are clustered by media market (panel A) or by county (panel B).

area's maximum search rate over the sample period. The treatment effects remain mostly above 17% for the following five weeks. Given our findings of the correlation between the Google search index and hate crimes, the increase in the index in the month after the first local diagnosis translates to an increase of 0.0095 anti-Asian hate crimes in a media market or 6.5% of the monthly average.¹³

Fig. A3 shows the event time plot when we replace the original Google search index with the indexes rescaled using three different bases. The patterns are qualitatively similar to those using the original index, although the magnitude of the estimates is now roughly half the size. This is because base search rates for the rescaled indexes are higher than search rates in most media markets. The standard errors of the estimates also become much larger because rescalling forces us to drop nearly half of the media markets (see Appendix 1 for detail). Because of this, we only present results using the original Google search index in the rest of the paper.

We next turn to Twitter to understand how the first local case affects *public* use of the ch-word. In Fig. 1, panel B, we plot the effect of the first local case on the racially charged Twitter post index. Similar to the Google search index, the Twitter post index also jumps in the week after the first case. Specifically, relative to the week before the case, the Twitter post index increases by 0.7 per 100,000 "the" tweets in the week after, amounting to 118.6% of the weekly average during the sample period. The effects remain high in weeks 2 through 6. Table 3, column (1) reports the regression results.

To confirm that our results are not driven by the functional form of the Twitter post index or the specific racial epithet we choose, we use alternative functional forms and other ways of identifying anti-Asian content. Raw number of ch-word tweets and number of ch-word tweets per million population reveal similar patterns as the original Twitter post index, as shown in columns (2) and (3). Additionally, we construct a new index using COVID-related tweets posted between January

¹³ We obtain the number by multiplying the average of +1w through +4w coefficients in Table 2 with the coefficient on *Google ch-word(t)* from Table 1.

Table 3 The effect of the first local COVID-19 Diagnosis on racial animus Twitter Post Index.

VARIABLES	(1) Ch-word index	(2) Ch-word level	(3) Ch-word per capita	(4) Exclude counter-hate	(5) Dynamic DID	(6) Severity control	(7) Asian control	(8) Exclude states	(9) Exclude bots
-6w	0.075	-0.022	0.061		-0.037	0.070	0.053	0.127	0.098
	(0.159)	(0.276)	(0.307)		(0.097)	(0.159)	(0.255)	(0.165)	(0.151)
-5w	0.030	-0.069	-0.036	-0.801	-0.085	0.027	0.091	0.056	0.039
	(0.143)	(0.158)	(0.267)	(1.252)	(0.091)	(0.143)	(0.242)	(0.142)	(0.153)
-4w	0.098	-0.128	-0.117	-0.328	-0.025	0.095	0.248	0.113	0.075
	(0.140)	(0.165)	(0.240)	(1.181)	(0.107)	(0.140)	(0.239)	(0.141)	(0.144)
-3w	-0.004	0.024	-0.100	0.450	-0.082	-0.006	0.095	0.018	0.014
	(0.121)	(0.091)	(0.195)	(1.152)	(0.081)	(0.121)	(0.213)	(0.129)	(0.138)
-2w	0.150	0.065	0.412	-0.361	0.120	0.149	0.331	0.136	0.242
	(0.137)	(0.050)	(0.308)	(0.967)	(0.094)	(0.137)	(0.212)	(0.146)	(0.180)
+0w	0.158	0.012	0.390**	5.154***	0.120***	0.163	0.168	0.169	0.203
	(0.112)	(0.069)	(0.170)	(1.005)	(0.094)	(0.159)	(0.171)	(0.122)	(0.142)
+1w	0.707***	0.227**	1.037***	5.075***	0.689***	0.718***	1.077***	0.572***	0.952***
	(0.169)	(0.105)	(0.197)	(1.046)	(0.159)	(0.166)	(0.238)	(0.162)	(0.228)
+2w	0.460***	0.348***	1.140***	2.855***	0.428***	0.478***	0.763***	0.389**	0.538***
	(0.142)	(0.109)	(0.252)	(1.039)	(0.111)	(0.145)	(0.199)	(0.151)	(0.173)
+3w	0.297**	0.631***	1.331***	2.688***	0.181*	0.315**	0.526**	0.300*	0.255*
	(0.141)	(0.193)	(0.396)	(0.842)	(0.095)	(0.152)	(0.204)	(0.154)	(0.137)
+4w	0.286*	0.789**	1.947**	1.521	0.122	0.307*	0.361	0.273	0.132
	(0.173)	(0.310)	(0.771)	(1.257)	(0.103)	(0.184)	(0.269)	(0.187)	(0.157)
+5w	0.394*	0.683***	1.650***	1.158	0.240	0.421*	0.535*	0.385	0.144
	(0.221)	(0.201)	(0.466)	(1.396)	(0.154)	(0.248)	(0.323)	(0.240)	(0.178)
+6w	0.459**	0.696***	1.664***	2.264	0.340**	0.489*	0.533*	0.479**	0.373*
	(0.222)	(0.223)	(0.469)	(1.566)	(0.150)	(0.252)	(0.315)	(0.243)	(0.198)
Observations	7930	7976	7976	3141	103,694	7930	5578	7188	11,811
R-squared	0.121	0.809	0.270	0.611	0.112	0.121	0.142	0.123	0.060
Outcome mean	0.591	0.681	1.075	6.779	0.591	0.591	0.591	0.591	0.569

Notes: The table presents the effect of the first local COVID-19 diagnosis on the prevalence of ch-word tweets in an area. All columns report the estimates of coefficients on the event dummies in Eq. (3), except for column (5). Column (1) corresponds to Fig. A4, panel B. The outcome variable in column (2) is the number of ch-word tweets, and the regression controls for the number of "the" tweets. The outcome variable in column (3) is the number of ch-word tweets per one million county population. Column (4) uses an alternative Twitter post index, which removes counter-hate tweets (see Section 4.2.1). Column (5) presents the estimates from a dynamic DID event study (Sun and Abraham, 2020). Column (6) controls for the number of COVID-related new cases and deaths, and whether the state has any stay-at-home orders in place. Column (7) controls for the Twitter post index for "Asian(s)." Column (8) excludes Washington, New York, and California. Column (9) excludes tweets from users who are likely Twitter bots. All regressions control for county fixed effects and year-month fixed effects. Standard errors are clustered by county. *** p < 0.01, ** p < 0.05, * p < 0.1.

15, and April 17, 2020 that are classified as anti-Asian via machine learning.¹⁴ Column (4) shows that the effects estimated with this new index share a similar pattern to the ones in column (1) but are seven times as large. The original Twitter post index is thus likely a conservative measure of racial animus.

An evolving local pandemic situation may produce time-varying treatment effects, which could bias results of a regular DID event study. To alleviate this concern, in column (5), we re-estimate the effect using a dynamic DID event study. The estimates are quantitatively similar to those in column (1), implying that time-varying treatment effect is likely not an issue here.

Fig. A4 presents results using indexes at the *daily* frequency. Both indexes start to rise two to three days after the first local case, suggesting that residents react to the news of the first local COVID-19 case fairly quickly.

4.2.2. Discussion and robustness check

We now discuss alternative explanations for the rise in the ch-word use in an area after the first local COVID-19 case and explore the robustness of our main findings.

Increased ch-word usage may result from a general rise in online activities due to blanket stay-at-home orders rather than a change in racial animus. However, our indexes already account for an overall change in online activities because they are normalized by the total searches and tweets in a given area and time. In addition, when we include an indicator for state-level stay-at-home order in Table 2 column (2) and Table 3 column (6), results are quantitatively similar to those from our main specification reported in column (1).

¹⁴ We thank Ziems, He, Soni, Kumar for providing the data. These anti-Asian tweets include phrases like "Chinese Virus" and "Wuhan Virus" and exclude counter-hate tweets that may have racist keywords in them. Only counties that had their first diagnoses between February 16, and March 22, 2020 are included in this analysis.

Alternatively, an increase in general attention to Chinese or Asians may lead to higher ch-word usage. In Table 2 column (3) and Table 3 column (7), we control for searches or tweets of terms that capture such general attention but are neutral in connotation, i.e., "Asian(s)." Results are similar.

Our results are also robust to excluding early- and hard-hit states like New York, Washington, and California, as shown in Table 2 column (4) and Table 3 column (8). Our findings thus represent a general phenomenon across the United States rather than only in a few states particularly impacted by the pandemic.

One may also worry that Twitter bots rather than actual users are responsible for the rise in ch-word use on Twitter. However, only 10.4% of users who post anti-Asian tweets between January 2020 and April 2020 are potential bots (Ziems et al., 2020). Moreover, our results are quantitatively unchanged when we exclude users who are more likely to be bots, i.e., those who tweeted the ch-word more than five times (99 percentile in our sample) during the sample period, in Table 3 col-umn (9).

The increase in searches and tweets including the ch-word could also come from the seasonality in ch-word use and may exist absent the pandemic. To test this possibility, we generate a placebo diagnosis date for each area using the same calendar day and month of its actual diagnosis date but changing the year from 2020 to 2019. We reestimate Eq. (3) using the placebo dates and plot the effects in Fig. A5. Reassuringly, the Google search index and Twitter post index do not change around the placebo dates, suggesting that seasonality cannot explain our findings.

Finally, the increase in ch-word use on Twitter could reflect a change in the social cost of publicly expressing racial animus rather than a shift in attitudes towards Asians. However, this would not explain the increase in racist Google searches, which are done in private. Several other pieces of evidence also support a shift in attitudes. First, the proportion of ch-word tweets showing emotions of anger and disgust increases from 23.3% between November 2020 and the first local diagnosis to 40.8% in the six weeks following the first diagnosis.¹⁵ Second, data on self-reported hate incidents from Asian Pacific Policy and Planning Council (AP3CON) Stop Hate Reporting System show that the daily average of anti-Asian hate incidents nationwide was alarmingly 70 in late March 2020 and 13 between April and May 2020 (Fig. A6). Third, Pew Research Center's Global Attitudes Survey, conducted in June to August 2020, also shows that unfavorable views of China have reached historic high (Wang, 2021). Taken together, the rise in ch-word usage likely represents a real change in animus against Asians and not just a lower cost of publicly expressing it.

4.3. What motivates racial animus and who responds the most

Thus far, we have provided evidence that animus against Asians, as measured by Google searches and Twitter posts including the ch-word, surges immediately following the first diagnosis in an area. We next explore *what* motivates individuals to increase animus in response to the pandemic and *who* responds the most.

As a first step, we test whether the rise is specific to Asians. If the racial animus is motivated by an overall increase in ethnic distrust or tensions from general uncertainty about cross-group differences in health status or risk-taking behavior, we expect to see an increase in animus against non-Asian minorities too. By contrast, if the racial animus is targeted *only* at Asians, it is more likely to be motivated by the association between Asians and the geographical origin of the virus.

To proxy for racial animus against other minorities, we construct Google search and Twitter post indexes for common racial epithets against major minority groups in the United States, such as "nigger(s)" (n-word) against African Americans, "wetback(s)" (w-word) against Hispanics, and "kike(s)" (k-word) against the Jewish population.¹⁶ We estimate Eq. (3) using racially charged searches and tweets against these minorities as outcomes.¹⁷ The coefficients on the event dummies are plotted in Fig. A7. None of the examined racial epithets experience an increase in Google searches following the first local diagnosis. A similar pattern is found for tweets using the w-word and the k-word.¹⁸ The lack of response in the use of racial epithets against other minorities suggests that the pandemic-induced racial animus is mainly driven by the connection between Asians and the geographical origin of the virus.

Although the anti-Asian animus is motivated by the potential geographical origin of the virus, racially charged tweets extend to broader topics than just the virus. Fig. 2 demonstrates that the increase in ch-word tweets mostly comes from those that *do not* explicitly mention COVID-19, i.e., no mention of "virus," "COVID," "pandemic" or "epidemic." This finding implies that the pandemic-induced racial animus may persist beyond the duration of the pandemic.

¹⁵ Crimson Hexagon assigns each tweet emotion tag(s) generated via a natural language processing algorithm. Please refer to https://www.brandwatch. com/blog/understanding-sentiment-analysis for more details.

¹⁶ We do not use "spic(s)" as an epithet against Hispanics because the cleaner brand "Spic and Span" experienced growing interest during the pandemic. We do not include "redskin(s)" as an epithet against Native Americans because the corresponding queries and tweets are about an American football team formerly called "Washington Redskins."

¹⁷ When using the n-word as the outcome, we include an indicator for the week of January 26, 2020 because the word's use spiked due to an MSNBC anchor using the n-word when broadcasting Kobe Bryant's death. When using the k-word as the outcome, we include an indicator for the week of February 23, 2020 because Los Angeles Dodgers player Enrique ("Kiké") Hernandez led to a spike in the word's use.

¹⁸ We present the result for tweets using the n-word in Fig. A7 panel C. N-word tweets may not be a valid proxy for racial animus against African Americans on Twitter because of Black Lives Matter protests, Black History Month in February, and seasonality which is evident when comparing the n-word usage between 2019 and 2020 in panel D. Note that we include an indicator for the week of February 9, 2020 in panel A because a viral n-word tweet unrelated to COVID-19 contributed to 95% of the n-word tweets on that day.



Fig. 2. The Effect of the First Local COVID-19 Diagnosis on Racially Charged Tweets COVID-Related vs Non-COVID-Related Tweets. *Note:* The figure presents the effect of the first local COVID-19 diagnosis on the racially charged Twitter post index by whether or not the tweets are related to COVID-19. COVID-related racially charged Twitter post index are defined as the number of ch-word tweets including keywords: "COVID-19", "COVID", "virus", "pandemic", or "epidemic", per 100,000 "the" tweets. The solid blue (dashed red) line plots the estimates and 95% confidence intervals of the coefficients on the event dummies in Eq. (3) using the (non-)COVID-related Twitter post index as the outcome. All regressions control for year-month fixed effects and county fixed effects. Standard errors are clustered by county.



Fig. 3. The Effect of the First Local COVID-19 Diagnosis on Racially Charged Tweets First-time vs Existing Ch-word Users. *Note:* The figure presents the effect of the first local COVID-19 diagnosis on the racially charged Twitter post index by whether the posting user is a first-time or an existing ch-word user. See Section 4.3 for definitions of first-time and existing ch-word users. The solid blue (dashed red) line plots the estimates and 95% confidence intervals of the coefficients on the event dummies in Eq. (3) using the racially charged Twitter post index based on first-time (existing) ch-word users as the outcome. All regressions control for year-month fixed effects and county fixed effects. Standard errors are clustered by county.

We next study which individuals are more susceptible to the pandemic shock. We begin by examining whether the increase in ch-word usage comes from users who only start to harbor animus against Asians after the pandemic hits or from existing anti-Asian users who step up their animosity. We define *existing* ch-word users as individuals who tweeted the ch-word at least once between 2014 and the sixth week before the first local COVID-19 diagnosis. We define first-time ch-word users as individuals who never tweeted the ch-word between 2014 and the sixth week before the first local diagnosis and who posted at least 10 tweets before their first ch-word tweet. This definition avoids counting newly registered Twitter users as first-time ch-word users.

Fig. 3 plots the breakdown in effects by the first-time versus existing ch-word user status. The increase in ch-word tweets from first-time users is roughly 4.5 times of that from existing users in the first two weeks after the first local diagnosis. This breakdown suggests that the extensive margin plays an more important role than the intensive margin in driving racial animus during the pandemic. After the first local diagnosis, 4621 Twitter users started to use the racial epithet, potentially exposing their combined 13 million followers to racially charged content and creating a multiplier effect on racial animus.

To better understand the type of individuals whose anti-Asian sentiment is easily influenced by the pandemic, we analyze user profiles and historical tweets of first-time ch-word users.¹⁹ To form a comparison group, we extract information for

¹⁹ We downloaded historical tweets and user profiles for 3033 of these users in August 2021. We cannot download the rest because their accounts are private, suspended, or deactivated.



B: Twitter user profile keywords

Fig. 4. Predictors of Being First-time Ch-word Users. *Note:* The figure presents the relationship between being a first-time ch-word user and one's Twitter activity and user profile keywords. Panels A and B plot the coefficients and 95% confidence intervals from regressing an indicator for being a first-time ch-word user on user's pre- and mid-pandemic Twitter activity, and user profile keywords, respectively. Both regressions control for account age, log number of followers, and log number of followings. Regressors in panel A are defined as follows: "Anti-Asian user" is one if an user has interacted with other ch-word users before the pandemic; "Anti-minority" is one if an user has tweeted racial epithets against non-Asian minorities (the n-word, and k-word) before the pandemic; "Trump" is one if an user has ever mentioned #trump or @realDonaldTrump before the pandemic; "McCarthy", "McConnell", "Pelosi", "Schumer", "Fox", "CNN", and "CBS" are similarly defined using @kevinomccarthy, @McConnellPress (or @LeaderMcConnell), @SpeakerPelosi, @SenSchumer, @cnn, @foxnews, @cnn, and @cbsnews as keywords, respectively; "COVID consp." is one if an user has ever tweeted keywords related to COVID-19 conspiracies (i.e., plandemic, fakepandemic, scamdemic, film your hospital, 5gcoronavirus, or coronavirustruth) by the end of our sample period. Regressors in panel B are the 25 most common user profile words used by first-time ch-word users and the 25 most common user profile words, so the number of words included in the regression is less than 50. Standard errors are heteroscedasticity-consistent. Regression results are reported in Tables A6 and A6.

3000 randomly selected Twitter users who registered before July 2019 and never tweeted the ch-word by the end of our sample period (hereafter, control users).

Table A5 reports the summary statistics for first-time ch-word users and control users. Both groups of users are seasoned Twitter users: their average account age is roughly six years, and their average number of followers is well over 1000. Compared to control users, first-time ch-word users are more likely to tweet racial epithets against other minorities and have interacted with anti-Asian users before the pandemic. They also appear to pay more attention to politics and news, as evident by their much higher interaction with twitter accounts of prominent politicians and major news outlets. Interestingly, very few ch-word users and control users ever tweeted COVID-related conspiracies.

To formally characterize users that are more susceptible to the pandemic-driven racial animus, we run two user-level regressions. The first one regresses an indicator for being a first-time ch-word user on the user's pre- and mid-pandemic Twitter activity, and the second one on user profile keywords; both regressions control for account age, log number of followers, and log number of followings. Regression results are plotted in Fig. 4 and reported in appendix Tables A6 and A7.

Fig. 4 panel A presents the relationship between being a first-time ch-word user and user activity on Twitter. Users who interacted with anti-Asian users before the pandemic are over twice the mean more likely than others to tweet the ch-word for the first time upon the pandemic. As we will show in the next section, this interaction plays a key role in spreading animus against Asians. In addition, users who tweeted racial epithets against non-Asian minorities before the pandemic are 40% the mean more likely to be first-time ch-word users. This finding implies that the crisis may have redirected pre-existing anti-minority sentiment towards Asians. Interestingly, paying attention to major politicians and news outlets also predicts a slightly higher chance of being a first-time ch-word user. Finally, tweeting COVID-related conspiracies has a precisely estimated zero effect on tweeting the ch-word, suggesting that such conspiracies are likely not the main cause of racial animus among users in our sample.

Fig. 4 panel B plots the relationship between being a first-time ch-word user and user profile keywords.²⁰ Consistent with results in panel A, keywords indicating attention to politics have the largest positive predictive power. Users who list "Trump" and "politics" in their profiles have a 58 and 28% higher chance (relative to the mean) of tweeting the ch-word for the first-time after the pandemic shock, respectively. As we will show in the next section, opinions of public figures, such as those of President Trump, likely play a crucial role in inciting anti-Asian sentiment during the pandemic. In contrast, keywords related to profession and family life, such as "artist," "wife," and "husband," predict a significantly lower propensity to tweeting the ch-word upon the pandemic.

5. Factors fueling racial animus

In this section, we explore factors that may have helped propagate anti-Asian animus during the pandemic. Understanding these factors is crucial to stopping the spread of animus from the outset amid future crises.

We know from the previous section that first-time ch-word users are the main driving force behind the rise of ch-word usage on Twitter during the pandemic. In Table 4, we zoom in on these users and their Twitter activity between the date of the first local COVID-19 case and the end of sample (May 2, 2020) to understand what prompts their first ch-word tweets. We regress a user's likelihood of tweeting the ch-word in a given day on a series of indicators for whom they interacted with and what they tweeted about in the day before. We control for user characteristics as well as county, year-of-week, and day-of-week fixed effects to absorb the average propensity to tweet the ch-word in a county and the national trend of such tweets.

Exposure to anti-Asian individuals. Table 4 column (1) shows that interaction with anti-Asian users (i.e. users who have previously used the ch-word) in a given day is associated with a 0.28 percentage point increase in the likelihood of tweeting the epithet in the following day, amounting to 22% of the sample mean. This finding highlights the importance of social media in spreading racial animus and is consistent with papers which document how social media influence real outcomes like voting behaviors (e.g., Fujiwara et al., 2021). Our finding suggests that moderating racist individuals and their interaction with others on social media could constrain the spread of animus.

Opinions of public figures. The only other positive predictor in column (1) is a user's interaction with President Trump. (Re)tweeting, replying, or mentioning the president in a day is associated with a 0.33 percentage point increase in the likelihood of tweeting the ch-word the next day, or 26% of the sample mean. This finding is consistent with Müller and Schwarz (2019) which shows that President Trump's tweets affect public behavior such as hate crimes. In contrast, mentioning other prominent politicians of either parties or national news accounts has little to no predictive power, or even predicts a lower likelihood of tweeting the epithet. When we additionally control for the number of new COVID-19 cases or deaths in the local area in column (2), the results remain similar. Taken together, certain public figures play a key role in shaping public opinions of a subject matter. Harnessing their opinion-shaping power could be useful in curbing animus in the future.

Salience of Asian-COVID connection. One potential factor mediating the relationship between ch-word use and interaction with President Trump is the salience of the connection between COVID-19 and the Asian population. It is possible that President Trump's tweets that simultaneously mention COVID-19 and China may increase the salience of the connection and influence racial animus. We categorize President Trump's tweets between January 1, 2020 and May 2, 2020 that contain any of the words "china", "chinese", "huawei", "xi", "covid", "covid-19", "corona", "coronavirus", "virus", "epidemic", or "pandemic" into three categories: those mentioning only China (China-only), only COVID-19 (COVID-only), and both China and COVID-19 (China-and-COVID). Table A8 presents examples of President Trump's tweets. Fig. A8 plots the daily frequency of his tweets.

In Table 5, we regress the daily racially charged Twitter post index at the national level on the number of the president's tweets in each of the three categories while controlling for year-week and day-of-week fixed effects. Column (1) shows that one additional China-and-COVID tweet of President Trump in a day corresponds to roughly five more racially charged tweets per ten million "the" tweets nationwide on the same day. This increase is non-trivial and is equivalent to 14% of the national daily average. Importantly, the Twitter post index does not co-move with the president's China-only or COVID-only tweets, highlighting that the *connection* between China and COVID-19 is what matters. Results remain similar when we control for the daily number of new COVID-19 cases and deaths nationwide in column (2).

²⁰ For ease of presentation, we only include the 25 most common user profile words used by first-time ch-word users and those by control users. Since there is an overlap between the two sets of words, the number of words included in the regression is less than 50.

Table 4

Predictors of Tweeting Ch-word among first-time Ch-word users after the first local COVID-19 diagnosis.

	(1)	(2)
VARIABLES	P(ch-word) (t+1)	P(ch-word) (t+1)
Anti-Asian user(t)	0.281***	0.259***
	(0.072)	(0.072)
Anti-minority(t)	1.156	1.112
	(1.360)	(1.368)
COVID consp.(t)	1.314	1.177
	(1.861)	(1.845)
Trump(t)	0.325***	0.297**
	(0.122)	(0.122)
McCarthy(t)	-0.184	-0.097
	(0.456)	(0.456)
McConnell(t)	-1.969***	-2.045***
	(0.351)	(0.449)
Pelosi(t)	-0.373	-0.419
	(0.284)	(0.285)
Schumer(t)	-0.031	-0.106
	(0.379)	(0.379)
CBS(t)	-0.636	-0.692
	(0.824)	(0.825)
CNN(t)	0.164	0.173
	(0.277)	(0.277)
Fox(t)	-0.430	-0.386
	(0.348)	(0.346)
Account years	-0.002	-0.002
•	(0.005)	(0.005)
Log(followers)	-0.032**	-0.031**
	(0.013)	(0.013)
Log(followings)	-0.010	-0.009
	(0.020)	(0.020)
New diagnoses		-0.000
-		(0.000)
New deaths		0.000
		(0.000)
Observations	174,164	174,164
R-squared	0.002	0.004
Outcome mean	1.251	1.251

Notes: This table presents the relationship between first-time ch-word users' likelihood of tweeting the ch-word (in percentage point) and their Twitter activity in the day before, as well as their baseline characteristics. See note to Fig. 4 for definitions of the independent variables. The data is at the user×day level, and all regressions control for county, year-of-week, and day-of-week fixed effects. Standard errors are clustered by user. *** p < 0.01, ** p < 0.05, * p < 0.1.

The time-series correlation may be confounded by contemporaneous shocks unrelated to the president's tweets. To alleviate this concern, we conduct an event study comparing nationwide racially charged Twitter post index in the hours before and after President Trump's China-and-COVID tweets, using the index during the same hours-of-day on days without such tweets as controls. Fig. 5 shows that the index in the four hours leading up to the China-and-COVID tweets is no different from other times, but it jumps in the first hour after such tweets and continues to grow. The immediacy of the change upon the president's tweets suggests a causal interpretation of the relationship between the salience of the China-and-COVID connection and the anti-Asian sentiment at the national level.

In addition, we study whether the salience of the connection has translated into hate incidents against Asians. We obtain self-reported anti-Asian hate incidents from AP3CON Stop AAPI Hate Reporting Center, a hate incident self-reporting website that went online on March 17, 2020. This is the best hate-tracking organization specialized in anti-Asian hate incidents in the United States (CBS News, 2020). In Table 5, column (3), we regress the log of daily hate incidents at the national level on the number of the president's tweets in each of the aforementioned categories while controlling for year-week and day-of-week fixed effects. We find that one additional China-and-COVID tweet from the president in a day corresponds to a roughly eight percent increase in self-reported hate incidents against Asians nationwide on the same day.²¹ When we control for the daily number of new COVID-19 cases and deaths nationwide in column (4), results are unchanged.²²

In contrast to the clear relationship between anti-Asian sentiment and the president's tweets, we find little evidence that the sentiment co-moves with tweets from other prominent politicians or national news outlets (Table A9). The difference is likely due to the large difference in the number of Twitter followers between the president and the others. President

 $^{^{21}}$ We conduct the analysis at the daily level because the exact hour of the incidents is not available. We cannot estimate Eq. (3) with AP3CON data due to the lack of pre-periods given the late start date of the data.

²² We find little relationship between the racially charged Google search index and President Trump's tweets. Results are available upon request.

Table 5

Relationship	between	trump	tweets	and	racial	animus	nationwide
--------------	---------	-------	--------	-----	--------	--------	------------

VARIABLES	(1) Twitter ch-word	(2) Twitter ch-word	(3) Log(incidents)	(4) Log(incidents)
China-and-COVID(t)	0.0482**	0.0493**	0.0799*	0.0888**
	(0.0234)	(0.0246)	(0.0453)	(0.0398)
China only(t)	-0.0126	-0.0130	-0.0592	-0.0332
	(0.0131)	(0.0133)	(0.0815)	(0.0844)
COVID only(t)	0.0008	0.0006	-0.0014	0.0004
	(0.0038)	(0.0040)	(0.0146)	(0.0143)
New diagnoses		-0.0000		0.0000*
		(0.0000)		(0.0000)
New deaths		0.0001		0.0001
		(0.0001)		(0.0003)
Observations	123	123	45	45
R-squared	0.519	0.522	0.812	0.829
Outcome mean	0.344	0.344	3.1932	3.1932

Notes: The table presents the relationship between the number of President Trump's tweets about COVID-19 and/or China and racial animus nationwide. The outcome variable in columns (1) and (2) is the daily number of ch-word tweets per 100,000 "the" tweets nationwide between January 1, 2020 and May 2, 2020. The outcome variable in columns (3) and (4) is the natural log of the daily number of anti-Asian hate incidents nationwide from AP3CON Stop AAPI Hate Reporting system between March 19, and May 2, 2020. We categorize the president's tweets that include "china", "chinese", "huawei", "xi", "COVID", "COVID-19", "corona", "coronavirus", "virus", "epidemic", or "pandemic" into three categories: "China-and-COVID" is the daily number of the president's tweets mentioning both China and COVID-19; "China only" those mentioning only China; and "COVID only" those mentioning only COVID-19. "New diagnoses" and "New deaths" are the daily number of COVID-related new cases and deaths in the United States. All regressions control for year-week fixed effects and day-of-week fixed effects. Standard errors are clustered by date. *** p < 0.01, ** p < 0.05, * p < 0.1.



Fig. 5. Relationship between Racially Charged Tweets Nationwide and Trump Tweets. *Note:* The figure presents the relationship between the number of President Trump's tweets that mention both Covid-19 and China (China-and-COVID tweets) in an hour and the number of ch-word tweets per 100,000 "the" tweets nationwide in the four hours before and the four hours after the president's tweets. The figure plots the estimates and 90% confidence intervals of the coefficients on the interactions between hourly event dummies and the number of Trump's China-and-COVID tweets at hour zero. Event dummy for the hours outside of those being plotted are omitted. The regression controls for year-week fixed effects, day of week fixed effects, and hour fixed effects. Standard errors are clustered by date.

Trump amassed 88.7 million followers before Twitter suspended his account in January 2021, while the follower number as of October, 2021 for the prominent politicians and national news outlets are mostly below 10 millions with only Fox and CNN reaching 20.2 and 54.7 millions, respectively.

Economic downturn. The COVID-19 pandemic poses risks on both lives and livelihoods. Existing work has documented that a deterioration of economic conditions can fuel animus towards minorities (Anderson et al., 2017; 2020; Sharma, 2015). We thus study the heterogeneity in the change in racial animus by the level of the pandemic's negative impact on the local economy. We partition the main regression samples by whether the proportion of an area's annual average employment in "leisure and hospitality" and "education and health services," the two hardest-hit industries in employment according to the Bureau of Labor Statistics (BLS), is above or below the sample median (32% in Google data and 35% in Twitter data). We also partition the samples by whether the percent change in net revenue between January and March, 2020 among local small businesses is above or below the sample median (-39% in the Google sample and -37% in the Twitter sample) using data built by Chetty et al. (2020). Fig. 6 shows that the areas that experience high versus low negative economic impact respond



C: Google - Revenue shock

D: Twitter - Revenue shock

Fig. 6. The Effect of the First Local COVID-19 Diagnosis on Racial Animus by the Negative Economic Impact of COVID-19. *Note:* The figure presents the heterogeneous effects of the first local COVID-19 diagnosis on the racially charged Google search index and Twitter post index by the negative economic impact of the pandemic. Panels A and B partition the regression sample by whether the proportion of an area's annual average employment in "leisure and hospitality" and "education and health services" is above or below the sample median (i.e., 32% in the Google sample and 35% in the Twitter sample). Panels C and D partition the regression sample by whether the percent change in net revenue among local small businesses between January and March is above or below the sample median (i.e., -39% in the Google sample and -37% in the Twitter sample). Panels A and C (B and D) plot the estimates and 95% confidence intervals of the coefficients on the event dummies in Eq. (3) using the racially charged Google search index (Twitter post index) as the outcome. Specifications in panels A and C mirror those in column (1) of Table 2, and specifications in panels B and D mirror those in column (1) of Table 2.

similarly to the first local COVID-19 diagnosis. In other words, the negative economic impact of the disease appears to play a relatively weaker role in motivating the initial rise of racial animus. One potential reason is that the long-term economic impact of the pandemic was not well understood at the beginning of the pandemic.

6. Conclusion

Growing racial tension is a serious challenge facing society. Understanding how racial animus forms and spreads is a critical step in addressing the issue. Using evidence from the COVID-19 pandemic, our paper sheds light on how and why negative shocks incite racial animus, types of individuals susceptible to such shocks, and factors that help spread the animus.

We exploit variation in the timing of the first COVID-19 diagnosis across US areas and find that the first local case leads to an immediate increase in local racial animus. This rise in animus specifically targets Asians, implying that the association between this group and the potential geographical origin of the virus likely motivates the animosity. The majority of racist tweets come from users who post the epithet for the first time; these first-time ch-word users are more likely to have expressed animosity against non-Asian minorities in the past, and their interaction with other anti-Asian individuals predicts the timing of their first ch-word tweets. These findings suggest that preconceived notions about minorities and social media network both help in the formation and the spread of racial hatred amid crisis. Moreover, users who list "Trump" in their profiles are more susceptible to the pandemic shock; online animosity and offline hate incidents against Asians both increase when President Trump more frequently links China and COVID-19 in his tweets. These findings underscore the crucial role of public figures in influencing public opinions of a subject matter. Finally, the pandemic-driven racial animus we documented may persist beyond the duration of the pandemic, as most racist tweets do not explicitly mention the virus.

Our findings have practical implications. Careful naming of a shock, debunking claims of any alleged connection between a shock and a group, moderating racist individuals and their interaction with others on social media, and harnessing public figures' opinion-shaping power could all be helpful in curbing animus amid future crises.

This paper also opens up several avenues for future research. While we estimate the effect of pandemics on racial animus, it would be interesting to know the downstream consequences of such crisis-driven animus, for example, on labor market, geographical sorting, and immigration. We characterize the users who are more susceptible to pandemic-induced animus against Asians, and it would be useful to characterize the users who express animosity against minorities in general so as to predict such behaviors and proactively curb the spread of racist content online.

Declaration of Competing Interest

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

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2022.05.014.

References

Anderson, L., Lepore, E., 2013. Slurring words. Noûs 47 (1), 25-48.

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

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

Association, P. B., 2014. Resolution opposing use of racial slurs in business names. https://www.philadelphiabar.org/page/ResolutionOpposingRacialSlurs? appNum=2.

Bartos, V., Bauer, M., Cahlíková, J., Chytilová, J., 2020. COVID-19 crisis fuels hostility against foreigners.

Becker, G.S., 1957. The Economics of Discrimination. University of Chicago Press.

Card, D., Mas, A., Rothstein, J., 2008. Tipping and the dynamics of segregation. Q. J. Econ. 123 (1), 177-218.

Charles, K.K., Guryan, J., 2008. Prejudice and wages: an empirical assessment of becker's the economics of discrimination. J. Polit. Econ. 116 (5), 773–809. Chetty, R., Friedman, J.N., Hendren, N., Stepner, M., et al., 2020. How did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data. Technical Report. National Bureau of Economic Research.

Croucher, S.M., Nguyen, T., Rahmani, D., 2020. Prejudice toward asian americans in the COVID-19 pandemic: the effects of social media use in the United States. Front. Commun. 5, 39.

Deng, H., Hwang, Y., 2021. Structural Analysis of Xenophobia. Available at SSRN 3958536.

Depetris-Chauvin, E., 2015. Fear of Obama: an empirical study of the demand for guns and the US 2008 presidential election. J. Public Econ. 130, 66–79.

Druckman, J.N., Holmes, J.W., 2004. Does presidential rhetoric matter? Priming and presidential approval. Pres. Stud. Q. 34 (4), 755–778.

Egorov, G., Enikolopov, R., Makarin, A., Petrova, M., 2020. Divided We Stay Home: Social Distancing and Ethnic Diversity. Technical Report. National Bureau of Economic Research.

Faulkner, J., Schaller, M., Park, J.H., Duncan, L.A., 2004. Evolved disease-avoidance mechanisms and contemporary xenophobic attitudes. Group Processes Intergroup Relat. 7 (4), 333–353.

Fujiwara, T., Müller, K., Schwarz, C., 2021. The Effect of Social Media on Elections: Evidence from the United States. Technical Report. National Bureau of Economic Research.

Hanes, E., Machin, S., 2014. Hate crime in the wake of terror attacks: evidence from 7/7 and 9/11. J. Contemp. Crim. Justice 30 (3), 247-267.

Healy, J., Searcey, D., 2020. Two Crises Convulse a Nation: A Pandemic and Police Violence. The New York Times.

Ivandic, R., Kirchmaier, T., Machin, S.J., 2019. Jihadi Attacks, Media and Local Hate Crime.

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.

Kaushal, N., Kaestner, R., Reimers, C., 2007. Labor market effects of september 11th on Arab and Muslim residents of the United States. J. Hum. Resour. 42 (2), 275–308.

Lyu, H., Chen, L., Wang, Y., Luo, J., 2020. Sense and sensibility: characterizing social media users regarding the use of controversial terms for COVID-19. IEEE Trans. Big Data.

Müller, K., Schwarz, C., 2019. From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment. SSRN Available at SSRN 3149103.

Mullis, S., Glenn, H., 2020. New site collects reports of racism against Asian Americans amid coronavirus pandemic. https://www.npr.org/sections/ coronavirus-live-updates/2020/03/27/822187627/new-site-collects-reports-of-anti-asian-american-sentiment-amid-coronavirus-pand.

Nguyen, T.T., Meng, H.-W., Sandeep, S., McCullough, M., Yu, W., Lau, Y., Huang, D., Nguyen, Q.C., 2018. Twitter-derived measures of sentiment towards minorities (2015–2016) and associations with low birth weight and preterm birth in the United States. Comput. Hum. Behav. 89, 308–315. Novgorodsky, D., Setzler, B., 2019. Practical guide to event studies.

O'Shea, B.A., Watson, D.G., Brown, G.D., Fincher, C.L., 2020. Infectious disease prevalence, not race exposure, predicts both implicit and explicit racial prejudice across the United states. Soc. Psychol. Personal. Sci. 11 (3), 345–355.

Schaller, M., Neuberg, S.L., 2012. Danger, disease, and the nature of prejudice (s). In: Advances in Experimental Social Psychology, Vol. 46. Elsevier, pp. 1–54.
Schild, L, Ling, C., Blackburn, J., Stringhini, G., Zhang, Y., Zannettou, S., 2020. "Go eat a bat, Chang!": an early look on the emergence of sinophobic behavior on web communities in the face of COVID-19. arXiv preprint arXiv:2004.04046.

Sharma, S., 2015. Caste-based crimes and economic status: evidence from India. J. Comp. Econ. 43 (1), 204-226.

Silver, L.R., 2016. China in the Media: Effects on American Opinion.

Stephens-Davidowitz, S., 2014. The cost of racial animus on a black candidate: evidence using Google search data. J. Public. Econ. 118, 26-40.

Sun, L, Abraham, S., 2020. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. J. Econom.

Tynes, B.M., Giang, M.T., Williams, D.R., Thompson, G.N., 2008. Online racial discrimination and psychological adjustment among adolescents. J. Adolesc. Health 43 (6), 565–569.

Wang, D., 2021. Xiuli Wang, Winning American hearts and minds: China's image building efforts in the 21st century. Int. J. Commun. 15, 4.

Ziems, C., He, B., Soni, S., Kumar, S., 2020. Racism is a virus: anti-asian hate and counterhate in social media during the COVID-19 crisis. arXiv preprint arXiv: 2005.12423.