Journal of Public Economics 249 (2025) 105450

Contents lists available at ScienceDirect

Journal of Public Economics

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

Discrimination and assimilation: Evidence from anti-Chinese sentiments in the United States ☆

Gianandrea Lanzara [a], Sara Lazzaroni [a], Paolo Masella [a], Mara P. Squicciarini [b,*]

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

HIGHLIGHTS

- Discrimination is an increasingly common phenomenon and assimilation decisions of a minority group often result from interactions with the majority group and political leaders.
- We use novel social media data to study the discrimination behavior of White Americans and the assimilation behavior of the Chinese American community in the U.S.
- Anti-Chinese discrimination increased following the COVID-19 outbreak, and Donald Trump's tweet referring to COVID-19 as the "Chinese virus".
- After Trump's tweet, Chinese users responded by asserting more frequently their Americanness and by increasingly distancing themselves from the Chinese Communist Party.

ARTICLE INFO

JEL classification:
D14
Z10

Keywords:
Culture
Immigrants

ABSTRACT

This paper studies the interactions between members of a discriminated minority, members of the majority group, and political leaders. We construct a novel dataset of all tweets posted by "White American" and Chinese users located in the United States from January to August 2020. Using a variety of supervised and unsupervised text-analysis techniques, we show that anti-Chinese discrimination on Twitter significantly increased following (i) the COVID-19 outbreak, and (ii) Donald Trump's tweet referring to COVID-19 as the "Chinese virus." We then study the reaction of the Chinese minority and find that, after Trump's tweet, Chinese users were significantly more likely to (i) tweet assimilation-related content, and (ii) tweet criticism against the Chinese Communist Party. The rise in assimilation-related content is generally stronger for users who were more integrated before the shock.

1. Introduction

Discrimination against minorities has a long history, and the spreading of anti-minority sentiments is increasingly common today. Given the consequences on the social, economic, and political spheres, the literature has extensively investigated the origins of anti-minority sentiments. Economic forces, social media, and political leaders' behavior have been considered among the possible drivers (see Dustmann and Preston, 2001; Mayda, 2006; Becker and Pascali, 2019; Müller and Schwarz, 2020, 2021, among others). However, little is known about how minorities react to discrimination, particularly when it comes from prominent political leaders. We do not know whether they respond by isolating themselves from the majority or by emphasizing their identification with their country of residence.

Exploiting the diffusion of anti-Chinese sentiments in the United States, following the 2020 coronavirus (COVID-19) crisis, the goal of this paper is to empirically analyze the interactions among members

☆ We thank Enrico Cantoni, Simon Görlach, Uta Schönberg, Carlo Schwarz, Vincenzo Scrutinio, and seminar participants at Bocconi University, at the second Text as Data Workshop, at the 2022 Winter Meeting of the Econometric Society, and at the second Bolzano Workshop on Historical Economics for useful comments and suggestions. We thank Caterina Alfonzo, Giacomo Caserta, Costanza De Acutis, Ludovica Massacesi, and Federico Scabbia for excellent assistance throughout the construction of the dataset. 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).

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

https://doi.org/10.1016/j.jpubeco.2025.105450
Received 21 July 2023; Received in revised form 31 May 2025; Accepted 14 July 2025
Available online 18 August 2025

of a discriminated minority, members of the majority group, and political leaders. This setting has two key characteristics. First, given the possible association of the virus with the location where it was initially discovered (Wuhan, China), we interpret the coronavirus outbreak as an exogenous shock to the incentives to discriminate against the Chinese minority. Thus, we identify the effect of the health shock on the discrimination decisions of the majority against the Chinese (which can be defined as "health-related discrimination"), as well as on the assimilation behavior of such minority.[1] Second, the use of anti-Chinese and anti-Asian rhetoric by the (then) U.S. President Donald Trump allows us to investigate the role of political leaders in further exacerbating health-related discrimination – triggering (what can be defined as) "leader-induced discrimination."

To study the health-related and leader-induced discriminatory and assimilation behaviors, we analyze the social media activity of a large sample of Twitter users in the U.S. We focus on two key dates: (i) March 9, 2020, when following the spread of the novel coronavirus, several restrictive measures were adopted in the U.S., such as isolating and quarantining suspected and confirmed cases, canceling public events, and suspending in-person classes at universities and schools (many district- and state-level closures soon followed); and (ii) March 17, when Trump referred to COVID-19 as the "Chinese virus." The majority group is represented by the sample of "White (non-Hispanic) American" users (the White group, henceforth), whereas the minority group is represented by the sample of Chinese users residing in the U.S. As described in detail in Section 3 and Appendix A, we obtain information on users' group identification thanks to the self-descriptions published on their profiles.

Social media data allow us to build several high-frequency measures of both discriminatory and assimilation actions. Past empirical studies on discrimination and assimilation typically relied on either survey responses (see, among others, Aspachs-Bracons et al., 2008; Manning and Roy, 2010), census records (Fouka, 2019, 2020; Abramitzky et al., 2020), or evidence from experimental manipulation of perceptions of the discriminatory context in the lab (see Bertrand and Duflo, 2017, for a review of such experiments). Using supervised and unsupervised text-analysis techniques, we are able to study *actual* individual-level discrimination of White Americans, as well as the reaction of members of the Chinese community. The latter does not capture equilibrium outcomes determined by changes in both the majority and the minority behaviors (e.g., marriage decisions or decisions to participate in the labor market), but it clearly reflects the response of the Chinese community.

To measure discrimination behavior, we follow the literature and track the use of racial slurs and hateful discourse toward the targeted group (see, for example, Lu and Sheng, 2020; Tahmasbi et al., 2021). To measure assimilation behavior, we move one step forward with respect to the economics literature and provide three *original measures* – which build on the work of sociologists studying the assimilation experiences of second-generation Chinese/Asian individuals. In particular, we capture how strongly the tweets either assert Americanness or express distance from the original ethnic group.

Exploiting the timing of the health and leader shocks, we show, by means of a trend-break model and a regression discontinuity design, that both shocks generated a sharp increase in the share of posted discriminatory text. Because China is both the epicenter of the first COVID-19 outbreak and the place of origin of the Chinese community, the cost of discriminating against an individual of Chinese origin decreased; similarly, discriminating against the Chinese community for the virus was a way to hold the minority group accountable for the shock. These effects were ultimately amplified when Trump tweeted about the "Chinese virus." In terms of magnitudes, we find a 1.22 percentage point increase in the daily probability of posting a tweet containing abusive language against the Chinese after March 9, followed by a 14.63 percentage point increase after March 17.

Finally, we turn to the question of whether increased discrimination raises or reduces the likelihood of observing tweets with assimilation content. Following the small increase in health-related discrimination on March 9, the regression discontinuity results point to a non significant reaction from Chinese Twitter users on March 9. However, corresponding to the larger increase in discrimination after March 17, we find a significant increase in tweets with assimilation content: the daily probability of posting a tweet with assimilation content increased by 2.6 percentage points, while the daily average share of text blaming the Chinese Communist Party (CCP) increased by 0.74 percentage points. These magnitudes imply a doubling of the pre-shock probability of tweeting assimilation content and a one third increase in the daily average share of text blaming the CCP, suggesting an important increase in assimilation behavior. The richness of the data allows us to show that the most integrated members of the Chinese community (that is, the Twitter users who were most likely to have posted assimilation content before the health shock) are the ones who most emphatically asserted their Americanness and distanced themselves from the CCP after March 17.

Our results are robust to a variety of robustness checks, e.g., when considering only politicians and activists, and when using alternative samples of users and definitions of our assimilation measures. Taken together, these findings suggest that the Trump's discriminatory behavior triggered a reaction from both the majority and the minority discriminated against. The majority showed higher discriminatory behavior and, in response, the minority asserted their belonging to the majority group or distanced themselves from their ethnic group of origin.

Literature. While most studies analyze the issues of discrimination and integration of minorities separately (see below for specific references), this paper is the first (to the best of our knowledge) to empirically study the interactions among a majority group, a discriminated minority, and political leaders.[2]

By analyzing the assimilation activities of a minority group facing discrimination, we contribute to a growing literature on the cultural and social integration of immigrants. Building on seminal works showing that immigrants' cultural identity significantly explains variation in several socio-economic outcomes (Giuliano, 2007; Bisin et al., 2008; Fernández and Fogli, 2009), researchers have recently investigated the assimilation decisions of immigrant minorities. Some studies focus on the choices of first names for children (Fouka, 2019, 2020; Abramitzky et al., 2020), hosting regions' language adoption/proficiency (Bleakley and Chin, 2010; Avitabile et al., 2013), intermarriage patterns (Bisin and Tura, 2019; Fouka, 2020; Guirkinger et al., 2021), and self-reported national identity (Aspachs-Bracons et al., 2008; Manning and Roy, 2010; Clots-Figueras and Masella, 2013; Abdelgadir and Fouka, 2020). These studies largely exploit the introduction of specific immigration policies and reforms (such as compulsory language laws, citizenship laws, no-fault divorce laws, and veil bans) to achieve causal identification.

Our contribution to this literature is twofold. First, we specifically consider how minorities react to changes in the discrimination environment driven by the behavior of a prominent political leader.

[1] The medical sociology literature has emphasized the ability of epidemic outbreaks to trigger anti-minority sentiments and xenophobia, because "outbreaks create fear, and fear is a key ingredient for racism and xenophobia to thrive" (Devakumar et al., 2020). Anti-minority sentiments have been historically linked to epidemics due to ideological scapegoating, for example in the case of the Black Death, when Jews were accused of well-poisoning in the context of antichrist conspiracy theories (Voigtländer and Voth, 2012; Jedwab et al., 2019), or social stigma of people or places associated with the birth and spread of the virus, as in the case of colonial sites affected by plague, cholera, and yellow fever during colonial times and, more recently, as in the case of Wuhan and China with the COVID-19 epidemic (White, 2020).

[2] See Eguia (2017) and Kim and Loury (2019) for theoretical models in the context of assimilation and discrimination in schools and in worker–employer relationships, respectively.

Second, while Gould and Klor (2016) and Fouka (2019) also exploit an exogenous increase in discriminatory behavior to investigate minority assimilation patterns, we are able to show causal evidence on *both* the discriminatory attitudes of the majority *and* the assimilation reactions of the minority discriminated against.

Showing evidence on *both* discrimination and assimilation in the context of an epidemic also allows us to take a step forward with respect to studies showing increasing anti-minority sentiments after economic shocks (Anderson et al., 2017; Becker and Pascali, 2019; Anderson et al., 2020; Grosfeld et al., 2020) and epidemics (Jedwab et al., 2019), including COVID-19 (Lu and Sheng, 2020; Ziems et al., 2020; Dipoppa et al., 2021; Tahmasbi et al., 2021).

In addition, we contribute to a recent growing literature linking traditional media and social media to the spread of violence against minorities. Regarding traditional media, DellaVigna et al. (2014), Yanagizawa-Drott (2014), and Adena et al. (2015) show that exposure to radio propaganda can contribute to ethnic hatred and violence. Similarly, in the context of digital media, Bursztyn et al. (2019) and Müller and Schwarz (2021) study how social media such as Facebook and VK can foster hatred of minorities. Differently from these studies, we exploit an exogenous trigger of possible discriminatory behavior to show how the majority and the minority respond in terms of discrimination and assimilation behavior.

We also relate to an emerging literature studying how leaders are able to influence the behavior of the population at large by affecting political preferences and mobilizing people in social movements (Dippel and Heblich, 2021; Cagé et al., 2023), and by influencing the societal perception of social norms regarding discrimination (Bursztyn et al., 2020; Grosjean et al., 2020; Müller and Schwarz, 2020). In contrast to these studies, we show how the majority and the minority react directly or indirectly to the discriminatory behavior of a prominent leader.

The remainder of this paper is as follows. Section 2 presents a simple theoretical framework to explore possible interactions between a majority group, a discriminated minority, and a political leader. Section 3 discusses the background and data for the empirical analysis. Section 4 illustrates the empirical strategy, Section 5 and 6 present our main findings. Section 7 concludes.

2. Conceptual framework

We present a simple conceptual framework to guide the interpretation of our empirical findings. Agents belong to one of two social groups: the majority or the minority. Each group consists of a continuum of agents. An agent from the majority group chooses an action d from the set $\{0,1\}$, where $d=1$ if the agent posts discrimination-related content on Twitter, and $d=0$ otherwise. Similarly, an agent from the minority group chooses an action a from the set $\{0,1\}$, where $a=1$ if the agent posts assimilation-related content on Twitter, and $a=0$ otherwise. The incentives to post discrimination- and assimilation-related content depend on the actions of other members of each respective social group. In the empirical application, we interpret the COVID-19 outbreak in the United States as an exogenous positive shock that increases the incentives to discriminate, and we investigate its effect on the share of agents in the majority and minority groups who post, respectively, discrimination- and assimilation-related content.

2.1. The majority group

Let s^M denote the share of agents in the majority group who decide to post a discriminatory tweet, $z \in [0,\bar{z}]$ represent the exogenous circumstances favoring discrimination, and, $\epsilon^M \sim unif(-1,0)$ be an individual idiosyncratic shock. We express the net return to posting discrimination-related content, u^M, as

$$u^M(z,s) = c^M + \alpha^M z + \beta^M s^M + \epsilon^M,$$

where $c^M \in [0,1]$ is a constant, $\alpha^M > 0$, and $0 < \beta^M < 1$. After observing the idiosyncratic shock, a member of the majority group posts a discriminatory tweet if $u^M > 0$. In equilibrium,[3] we have

$$s^M(z) = \frac{c^M}{1-\beta^M} + \frac{\alpha^M}{1-\beta^M} z$$

Intuitively, this equation is saying that a positive shock to z increases the share of agents in the majority group who post discrimination-related content. The term $1-\beta^M$ in the denominator represents the multiplier effect of the equilibrium interactions within the majority group.

2.2. The minority group

Let s^m denote the share of agents in the minority group who post an assimilation tweet, and let $\epsilon^m \sim unif(-1,0)$ be an individual idiosyncratic shock. We express the net return to posting an assimilation-related content, u^m, as:

$$u^m(s^d, s) = c^m + \alpha^m s^M + \beta^m s^m + \epsilon^m,$$

where $c^m \in [0,1]$ is a constant and $0 < \beta^m < 1$. After observing the idiosyncratic shock, a member of the minority group decides to post an assimilation tweet if $u^m > 0$. In equilibrium, we have

$$s^m(z) = \frac{c^m}{1-\beta^m} + \frac{\alpha^m}{1-\beta^m} s^M = \frac{c^m + \alpha^m c^M/(1-\beta^M)}{1-\beta^m} + \frac{\alpha^m}{1-\beta^m}\frac{\alpha^M}{1-\beta^M} z$$

This equation shows that the assimilation behavior of the minority group in response to an exogenous discrimination shock depends on the sign of α^m. Intuitively, the sign of α^m determines whether the assimilation incentives of the minority group become stronger or weaker in a more discriminatory environment.[4]

For $\alpha^m > 0$ ($\alpha^m < 0$), a positive discrimination shock increases (decreases) the share of minority group agents who post assimilation-related content. In both cases, this effect operates through changes in the discriminatory behavior of the majority group. The term $1-\beta^m$ in the denominator represents multiplier effect of the equilibrium interactions within the minority group.

2.3. The leader

In the setting outlined above, the actions of individual agents have no impact on aggregate outcomes, because each agent is infinitely small. However, the actions of certain agents, i.e., leaders, may have a large impact on society. To incorporate this feature, suppose that agents in the majority group make their choices after having observed the action of a leader, d^ℓ, where $d^\ell = 1$ if the leader posts discrimination-related content, and $d^\ell = 0$ otherwise. In our empirical setting, we interpret Trump's tweet on the "Chinese virus" as a discriminatory action on the part of a leader, and we investigate its effect on the share of agents in the majority and minority groups who posted, respectively, discrimination- and assimilation-related content.

The leader's actions affect the discrimination incentives of the majority group. Specifically, we modify the net return to posting a discriminatory tweet, u^M, as follows:

$$u^M(z,s,d^\ell) = c^M + \alpha^M z + \beta^M s^M + \gamma d^\ell + \epsilon^M,$$

with $\gamma > 0$ and $\beta^M + \gamma < 1$. In equilibrium, we have

$$s^M(z,d^\ell) = \frac{c^M}{1-\beta^M} + \frac{\alpha^M}{1-\beta^M} z + \frac{\gamma}{1-\beta^M} d^\ell$$

By influencing the behavior of the majority group, the leader also indirectly affects the behavior of the minority group. Using this expression

[3] To ease the exposition, in the following we always assume that the value of $\bar{z}$ is small enough to ensure that all shares and probabilities are between zero and one. In this specific case, for instance, we are assuming $\bar{z} \le (1-\beta^M - c^M)/\alpha^M$; then, $\bar{z} > 0$ also implies $\beta^M < 1 - c^M < 1$. Similar restrictions can be imposed for the other expressions below.

[4] We remark that $\alpha^m > 0$ is consistent with discrimination having an adverse effect on the (absolute) welfare of the minority group.

for s^M in the previous expression for s^m, we have

$$s^m(z, d^\ell) = \frac{c^m + \alpha^m c^M/(1-\beta^M)}{1-\beta^m} + \frac{\alpha^m}{1-\beta^m}\frac{\alpha^M}{1-\beta^M}z + \frac{\alpha^m}{1-\beta^m}\frac{\gamma}{1-\beta^M}d^\ell$$

Again, the sign of α^m determines how the minority group responds to changes in the aggregate level of discrimination in society, either due to exogenous circumstances or the actions of a leader.

Finally, we can go one step further and derive the leader's optimal behavior. Like the other members of the majority group, the leader's choice to post discrimination-related content depends on the exogenous circumstances, z, on the share of agents from the majority group who post discrimination-related content, s^M, and on an individual idiosyncratic shock, $\epsilon^\ell \sim unif(-1, 0)$. The main difference is that the leader moves first, correctly anticipating the share of agents in the majority group who will post discriminatory tweets following each of his possible choices. Accordingly, we express the leader's net return to posting a discriminatory tweet, u^ℓ, as

$$u^\ell(s, z) = c^\ell + \alpha^\ell z + \beta^\ell s^M(z, 1) + \epsilon^\ell$$

Using the expression for s^M above, we obtain:

$$Pr(d^\ell = 1) = c^\ell + \frac{\beta^\ell c^M}{1-\beta^M} + \left(\alpha^\ell + \beta^\ell\frac{\alpha^M}{1-\beta^M}\right)z + \beta^\ell\frac{\gamma}{1-\beta^M}$$

This equation illustrates that the exogenous circumstances affect the probability that the leader posts a discriminatory tweet through two channels: (i) a direct effect on the leader's own propensity to post discrimination-related content (α^ℓ), and (ii) an indirect effect on the social group's propensity to post discrimination-related content (α^M). Besides the role of z, the third term in the sum illustrates that the leader has a further incentive to post discriminatory tweets, coming from the fact that he is able to affect the payoff of the members of the majority group (γ).

2.4. Summary

In line with the literature on scapegoating, the COVID-19 outbreak fueled xenophobic associations between the spread of the virus and the Chinese minority. Within the simple framework outlined above, we interpret this phenomenon as a positive shock to z, where z represents the exogenous circumstances influencing the incentive to discriminate. Similarly, we interpret Trump's "Chinese virus" tweet as a discriminatory action by a prominent leader, who is able to influence the incentives of the majority group.

Following the U.S. outbreak on March 9, 2020, as well as Trump's tweet on March 17, 2020, our framework predicts (i) an increase in the share of majority group members posting discriminatory content, and (ii) a change in the share of minority group members posting assimilation-related content. This share may increase or decrease depending on the sign of α^m. Because the sign of α^m determines how the assimilation incentives of the minority group are affected by the discriminatory environment, our empirical analysis will provide insights into this key parameter.

3. Background and data

In this section, we discuss the background and data used for the empirical analysis. In Section 3.1, we provide details on the U.S. context during the period from January 6 to August, 27, 2020, when the U.S. population experienced an exogenous change to the returns to discriminatory actions due to (i) the outbreak of a deadly virus that originated in China, and (ii) an unexpected discriminatory action by the U.S. President (the most prominent political leader at that time) toward the Chinese community.[5] Then, Section 3.2 discusses the sample construction and the main dependent variables used in the analysis.

3.1. Background

The 2019 novel coronavirus was first identified in Wuhan, Hubei, China, where a major local outbreak suddenly became a global public health emergency. The first COVID-19 case in the U.S. was reported on January 20 in the state of Washington (Holshue et al., 2020). Italy was the first severely hit country in Western Europe, starting on February 21, 2020. By early March cases had been recorded in more than 100 countries, and after Rhode Island and Ohio declared states of emergency, the United States adopted several restrictive measures on March 9. These included isolating and quarantining several suspected and confirmed cases, canceling public events, and suspending in-person classes at universities. On the evening of March 16, U.S. President Donald Trump referred to COVID-19 as the "Chinese virus" in a tweet regarding economic support for U.S. industries affected by the pandemic. On March 17, National Public Radio announced that all U.S. states had reported COVID-19 cases (NPR, 2020) and Trump again used the expression "Chinese virus" despite media accusations of racism (see Appendix A.1). In line with these accounts, based on data from Twitter (details on the data and the sample are in Section 3.2), the upper line in Fig. 1 displays the evolution of the daily share of Twitter users tweeting keywords such as "virus," "corona," and "covid," and highlights the increasing salience of COVID-related topics in the overall tweeted text after the major events discussed above (signaled by the vertical lines labeled "1st U.S. case" for the first U.S. case, "Italy" for the outbreak in Italy, "U.S." for the implementation of restrictive measures in the U.S., and "Trump" for the discriminatory tweets by Trump). Notably, corresponding to March 9 and March 17, we observe a sharp increase in the share of tweets containing keywords related to China (green line), which is largely driven by tweets containing keywords related to both China and the virus (red and orange lines, the latter restricting the pattern to the specific expression "Chinese virus"). Our analysis will focus precisely on March 9 and March 17 as thresholds at which we may find a discontinuity in both discriminatory behavior and the reaction of the minority discriminated against.

3.2. Data

Our analysis is based on a rich dataset on social media activity in the U.S., including discriminatory attitudes of the White group and related reactions of the Chinese group. In particular, within Twitter, we identify 8,130 White users and 832 Chinese users, and we follow their activity over time. For all users in the sample, we observe the universe of tweets from January 6, 2020 to August 27, 2020 for a total of 13,595,102 tweets.

We selected users based on three criteria: (i) the self-descriptions reported in their profiles had to include keywords that signal their belonging to either the White (non-Hispanic) or the Chinese community; (ii) they had to be likely to have resided in the U.S. during the study period; (iii) they had to be sufficiently active in the Twitter community and likely to be engaged in social, political, and cultural debates (see Appendix A.2 for details on the construction of the dataset).[6]

As in most studies based on social media users, the sample construction may involve some selection, as users tend to cluster (and post content) based on the type of information they are exposed to (see, for instance, the discussion in Schmidt et al., 2017; Sunstein, 2018; Müller and Schwarz, 2021). Moreover, the limited number of Chinese users in our sample, possibly due to a differential propensity to engage on Twitter, may raise concerns about potential selection bias and limit the

[5] Based on the functionalities of the Twitter API accessed in February 2020, we could retrieve only the previous three months of each user's activity. Because some users posted very few tweets before January 6, we chose this as the starting date for our sample. We end our period of analysis on August 27, 2020 when the number of COVID-19 hospitalizations in the United States reached the lowest point after the first two major waves.

[6] To filter out users with minimal level of Twitter activity, we only retain users who posted at least one tweet before December 31, 2019, and at least one tweet after January 1, 2020.

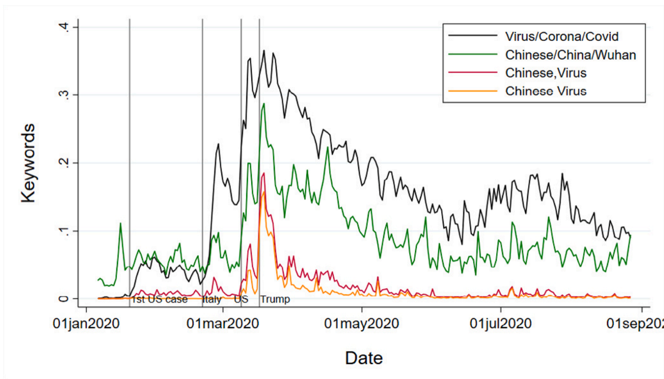


Fig. 1. Daily Share of Users (All Groups) Who Tweeted Selected Keywords. *Notes*: This Figure shows the share of users who tweeted a given keyword or set of keywords on each date from January 6 to August 27, 2020. The black line represents the daily share of users who tweeted "virus," "corona," or "covid." The green line represents the daily share of users who tweeted "chinese," "china," or "wuhan." The red line represents the daily share of users who tweeted "chinese" and "virus" in the same tweet, in any position or order. The orange line represents the daily share of users who tweeted "chinese virus." The gray vertical lines labeled "1st U.S. case" (January 20), "Italy" (February 21), "U.S." (March 9) and "Trump" (March 17) signal the first U.S. case, the outbreak in Italy, the implementation of restrictive measures in the U.S., and the discriminatory tweet by Trump, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

external validity of the results. However, in our case, selection concerns are mitigated by three observations. First, as reported in the Twitter bios, the users in our sample are very heterogeneous in terms of their activities; they range from full-time mothers to entrepreneurs, actors, politicians, and activists. Second, White and Chinese users are also very heterogeneous with respect to their pre-shock share of generalized abusive language (as opposed to directed specifically against the Chinese community) and their pre-shock share of posted assimilation content (see Appendix Figure A5).[7] Third, the focus on alternative samples does not affect the baseline results (see Section 5).

Altogether, these observations help us mitigate selection concerns. In addition, contrary to survey or experimental data, social media data has the clear advantage of providing high-frequency measures of both discrimination and assimilation.

Our goal is to extract information on the discriminatory and assimilation content included in the tweets posted by the users in our sample. To this aim, we adopt three different text-analysis approaches: dictionary-based, supervised machine learning, and unsupervised machine learning. In the first two approaches, a textual unit of analysis, i.e., a "document," is a single tweet. In the third approach, instead, we define a document as containing the entire text tweeted by a single user in a day; thus, if a user created multiple tweets in a day, we paste them together. The reason is that commonly used unsupervised methods, such as the Latent Dirichlet Allocation, tend to underperform with short texts. We now briefly describe each of the three methods. More details can be found in Appendix sections A.3, A.4, and A.5.

We start by performing dictionary-based exercises, that is, we search for specific keywords inside the textual units of analysis. For instance, we look for Chinese slurs (e.g., "ching chong") to detect anti-Chinese discriminatory tweets. Though simple and intuitive, this method bears the limitation that the choice of keywords is subject to human bias and the context in which the keyword occurs is not taken into account.

[7] We consider pre-shock (between January 6 and February 17, 2020) levels of posted abusive language and assimilation content to focus on a period of time when COVID-19 was still not perceived as a direct threat to U.S. society.

Second, we use supervised machine learning. Unlike dictionary-based exercises, supervised methods are based on a training dataset where the features of interest are observed for a certain number of (textual, in our case) examples. This information can then be leveraged to predict the same features in the main dataset of the analysis (Hastie et al., 2009). As far as discrimination content is concerned, we exploit the annotated dataset by Founta et al. (2018), which has been adopted by a recently growing literature on hate–speech detection, and in particular on hate–speech detection on Twitter.[8] Here, tweets are labeled as "hateful," "abusive," "normal," or "spam." Then, we classify our own sample of tweets using the Support Vector Machine (SVM) algorithm, which is widely used in text-analysis applications. In contrast (to the best of our knowledge) no comparable dataset exists for assimilation-related content. We take a different route and follow the methodology of Ash et al. (2021), based on word embeddings. Word embeddings are representations of words in an $\mathbb{R}^K$ vector space, such that each of the K dimensions corresponds to an aspect of meaning. Similar words will be located near each other, and, more generally, relationships between words will follow an internally consistent metric. Documents can also be vectorized as (standardized) sums of the vectorized words that are part of the document. We embed our sample of tweets by Chinese users in a 150-dimensional vector space, and we compute their proximity to the "assimilation dimension" of that space. To locate this dimension, we vectorize a sample of sentences with clear assimilation content, mostly drawn from sociological studies on assimilation, such as Kibria (2000), reporting interviews with second-generation Asian Americans.[9]

Finally, for our unsupervised machine-learning method, we use the Latent Dirichlet Allocation (LDA) algorithm, developed by Blei et al. (2003) (see also Hansen et al., 2018, for an early application in economics), to identify the latent topics in the corpus of documents for a given group (e.g., White American users), and to derive the topic composition of each document.[10]

We preprocess the raw text in several ways. First, we convert the text to lower case and remove URLs, mentions, punctuation, and numbers (except when tagging U.S. bills, e.g., S386: Fairness for High-Skilled Immigrants Act, 2019), plus a number of minor adjustments that we detail in Appendix A.2. Second, when we use the text as an input in the supervised and unsupervised methods, we make it undergo three additional preliminary steps: (i) we remove stopwords; (ii) we replace words with their stems (using the Porter stemmer); and (iii) we apply the term frequency-inverse document frequency (tf-idf) filter, in line with Hansen et al. (2018). Importantly, before the preprocessing phase, all text in Chinese (in the Chinese users' tweets) is translated into English via the DeepL API.

3.2.1. Discrimination content

We now describe our measures of discrimination content and how they are constructed. First, we perform a dictionary-based exercise, selecting a set of 18 keywords that express racial slurs against the Chinese minority.[11] We compute a dummy variable labeled *Chinese Slurs* taking the value one if, on a given day t, the user tweeted at least one of these

[8] Alternative datasets are provided by Waseem and Hovy (2016), Davidson et al. (2017), Vidgen et al. (2020), and Ziems et al. (2020).

[9] Although word embeddings are generated using unsupervised algorithms, we refer to this method as supervised because our classification exercise relies on an external dataset of assimilation sentences.

[10] Three parameters have to be set externally: the number of topics (k), plus the two hyperparameters (α and δ) for the prior Dirichlet distributions. In the following exercises, we set $k = 60$ for the White group and $k = 40$ for the Chinese group, and, following Griffiths and Steyvers (2004) and Hansen et al. (2018), $\alpha = k/50$ and $\delta = 0.1$.

[11] The keywords are the following: chink, chingchong, chinesebat, chinesestudentban, chinesespy, chonk, churka, cokin, coolie, dink, flango, gook, kungfuflu, niakoue, slanteye, slopehead, tingtong, yokel.

keywords. This represents our first discrimination outcome; the unit of observation is at the user-day level.

Second, we perform a supervised machine-learning exercise that relies on the dataset of 99,996 annotated tweets presented in Founta et al. (2018). These tweets were collected via the Twitter Stream API from March 30 to April 9, 2017, and classified into four categories: hateful, abusive, normal, and spam, via the *CrowdFlower* crowdsourcing platform. After deleting the spam tweets, we end up with 4,965 (6 %) hateful tweets, 27,150 (32 %) abusive tweets, and 53,851 (62 %) normal tweets. Using this labeled dataset, we group the hateful and abusive tweets in the same class and predict whether a tweet belongs either to the hateful/abusive category or to the normal speech category. To perform this prediction exercise, we first set aside 30 % of the labeled dataset as a test set to carry out an ex-post evaluation of the performance, and we use the remaining 70 % to train an SVM model (see Hastie et al., 2009, for a reference).[12] On top of word frequencies, we use the total number of letters and the number of punctuation marks as additional predictive features.[13] The model achieves an accuracy (share of correctly classified observations) of 88.2 %.[14] Finally, we use this model to classify the tweets in our sample and compute the dummy variable $Abusive\ Language$ which takes the value one if, on a given day t, a user posted one or more hateful or abusive tweets, including the keyword "Chinese." This is our second discrimination outcome; the unit of observation is at the user-day level.

Finally, we run the LDA algorithm on the corpus of daily tweeted documents of the White users. In our baseline exercise, we use 60 topics (the main topic of interest is also found using more or fewer topics). Figure A3 in Appendix A.5 shows the most representative words for each topic. For our empirical analysis, we focus on a topic that we naturally labeled "Discrimination" (topic 19), as the documents with the largest share of text attributed to this topic clearly contain discriminatory content against the Chinese community (see some examples in Appendix A.5.1). The most relevant keywords for this topic are "Chinese, Wuhan, communist, pandemic, spread, stock, CCP, shut, blame, lab." Fig. 2 shows the wordcloud for this topic. Our third discrimination outcome is the variable $Topic\ Discrimination$, which, for each day t, computes the average of the share of text pertaining to this topic within the entire sample of documents posted by users on that given day.[15] Therefore, the number of observations corresponds to the number of days in the sample.

Fig. 3 reports the evolution of the daily share of White users who tweeted Chinese slurs, the daily share of White users who tweeted expressions with abusive language against the Chinese, as well as the daily average of the share of text associated with the anti-Chinese discrimination topic. At the bottom of each graph, we report the mean and standard deviation of the respective variable. Consistently across all outcome variables, the figure shows a peak in discriminatory language on March 9, 2020, and a subsequent very sizable spike on March 17, 2020.

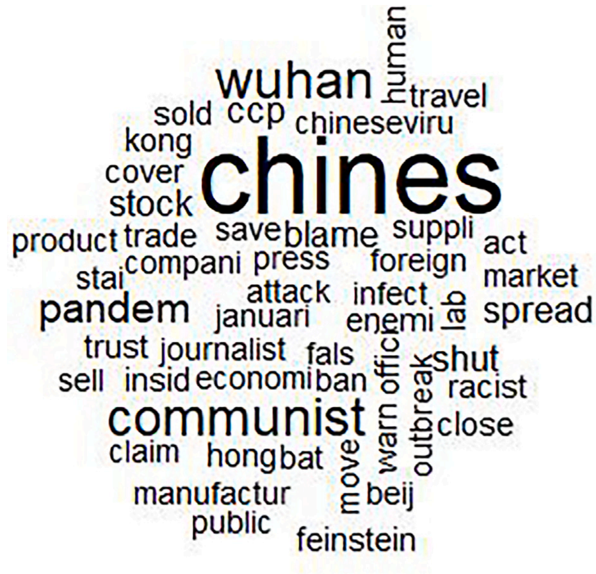


Fig. 2. Wordcloud of the Topic "Discrimination". *Notes*: This wordcloud is based on the LDA on the overall text tweeted by users of the White group from January 6 to August 27, 2020. Larger words are more recurrent. See above and Appendix A.5.1 for details.

3.2.2. Assimilation content

In building a measure of assimilation content, the main challenge we face is that cultural assimilation is a multifaceted concept, difficult to capture using automated techniques. As a result, text analysis algorithms are likely to measure assimilation content with some error, and the sources of measurement error may vary across algorithms. To address this challenge, we construct multiple proxies of assimilation content building on the sociology literature and, specifically, on the literature studying the process of the identity formation of second-generation Asian Americans. In particular, Kibria (2000) examines the dynamics of race and ethnic identity for the Asian ethnic group in U.S. society and emphasizes, among others, two key assimilation behaviors of the second generation: (i) *establishing Americanness* and (ii) *distancing from the reference ethnic group*. We construct a measure of assimilation content based on the first type of behavior, a second measure of assimilation based on the second type of behavior, and a third more-general measure meant to capture both aspects. Starting with the first, Kibria (2000, p. 86) suggests that *establishing Americanness* by stating the ethnic identity (i.e., Chinese or Asian) followed by the nationality (i.e., American), or by expressing adherence to mainstream U.S. culture, norms, and practices, is a key strategy for establishing one's identity as "American," and asserting "the nature of one's ties and relationship to U.S. society and culture." As the first measure of this type of behavior, we rely on the dictionary-based methodology and compute a dummy variable labeled *Chinese/Asian American*, which takes value one if, on a given day t, the user posted one or more tweets with the keyword "Chinese American" or "Asian American." The unit of observation for this measure is at the user-day level.

Then, Kibria (2000, p. 84) emphasizes that *distancing from the reference ethnic group* has traditionally been a way to "cope with racialized hostility." Disidentification is "the use of various clues and signals to distance oneself from the perceived 'problem group'" (Goffman, 2009). This often helps second-generation Chinese and Korean Americans deflect hostility when accused by strangers of the problems caused by various Asian groups (Kibria, 2000).[16] Similarly, in our case, taking distance

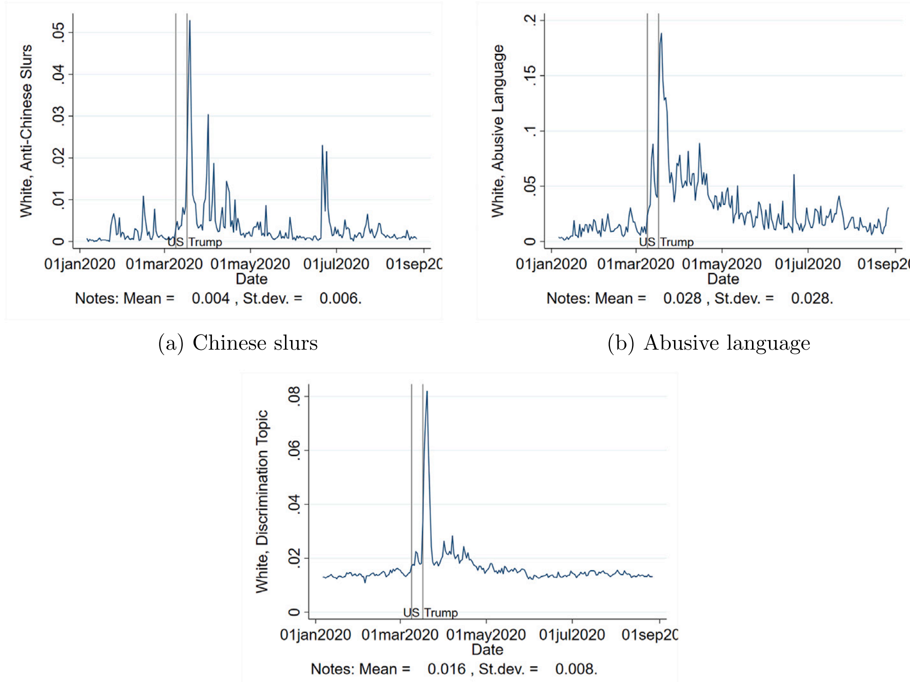
[12] We perform a fivefold cross-validation procedure to tune the SVC cost parameter. Essentially, this procedure performs a grid search over different values of the tuning parameter, then selects the one that achieves the best performance. To measure performance, the algorithm splits the training dataset into n (five, in our case) subsets (folds). Then, we train the model on a dataset comprising $n-1$ folds, and we test its performance over the excluded fold; this step is repeated n times, such that a different fold is excluded in each iteration. At the end, we compute an average measure of performance. We use the following grid: $0.1, 1, 10, 100, 10000$, and the chosen parameter is 0.1.

[13] In the Founta et al. (2018) dataset, the average number of letters is 94.3 in normal tweets, versus 85.4 in hateful tweets and 70.4 in abusive tweets; the average number of punctuation marks is 2.83 in normal tweets, versus 2.35 in hateful tweets and 2.13 in abusive tweets.

[14] The confusion matrix is reported in the Appendix A.3.

[15] Each document contains all the text tweeted by a single user on a given day, as defined in Section 3.2.

[16] "The notion of Asians as 'essential foreigners' has played an important role in organizing and legitimizing hostility toward and discrimination against Asians in the United States [since] the 1882 Chinese Exclusion Act, which banned Chinese



(a) Chinese slurs

(b) Abusive language

(c) Topic "Discrimination"

Fig. 3. Discrimination Keywords, Expressions, and Topics for the White Users. *Notes*: This Figure shows the daily share of White users tweeting Chinese slurs (Panel a), the daily share of White users tweeting abusive language against the Chinese (Panel b), and daily average of the share of text of the anti-Chinese discrimination topic within the sample of White American users tweeting in a given day (Panel c), from January 6 to August 27, 2020. Mean and standard deviation of the variables in the period of interest are reported below each graph.

from some distinctive features of China could be a strategy for facing anti-Chinese/Asian sentiment. Given that the CCP is a defining characteristic of China, and that the Party was particularly involved in the management of the epidemic, disidentification from the CCP seems a natural way to cope with discrimination in the COVID-19 context. To measure this, we use the LDA algorithm and model the daily tweeted documents of the Chinese users into 40 topics.[17] Figure A4 in Appendix A.5 shows the set of words that most closely represent each topic. For our empirical analysis, we focus on topic 13, where the relevant keywords are "ccp, communist, party, ccpvirus, truth, hsk, american, govern, pandem, america." This topic is largely devoted to accusing the CCP of the spread of the virus in China, the U.S., and worldwide. We labeled this topic as *Topic Blame CCP*.[18] Fig. 4 depicts the related wordcloud. In addition, examples of text classified into this topic are reported in Appendix A.5.2. Our second assimilation outcome is the variable $Topic\ Blame\ CCP$, which, for each day t computes the average of the share of text pertaining to this topic within the entire sample of documents posted by users on that day.[19] The number of observations corresponds to the number of days in the sample.

Finally, to build our third, more-general measure of assimilation content, we exploit the transcripts, from nine distinct sources, of interviews

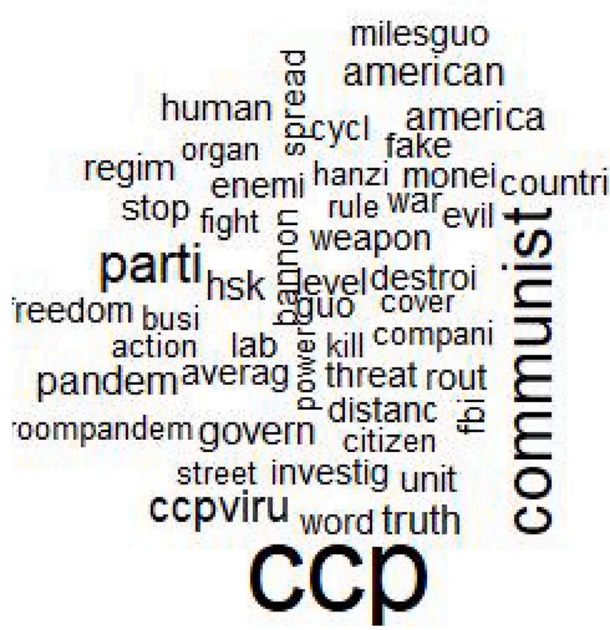


Fig. 4. Wordcloud of the Topic "Blame CCP". *Notes*: This wordcloud is based on the LDA on the overall text tweeted by users of the Chinese group from January 6 to August 27, 2020. Larger words are more recurrent. See above and Appendix A.5.2 for details.

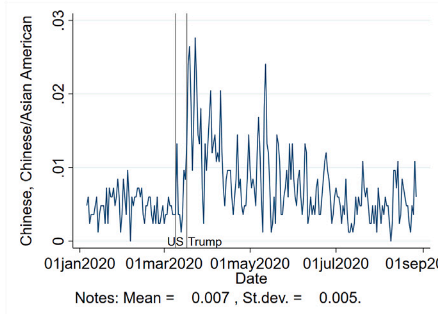
with Chinese/Asian immigrants sharing their assimilation experiences. From these transcripts, we extract 55 sentences with a clear assimilation content—see Table A2 in Appendix A.4 for full text and sources. In these sentences, first- and second-generation Asian Americans express

immigration to the United States. Observers also note how hate crimes against Asians today often focus on their presumed 'foreignness.' […] Thus the second generation felt compelled […] to downplay their distinctive ethnic backgrounds in order to establish themselves as 'American'." (Kibria, 2000, p. 86).

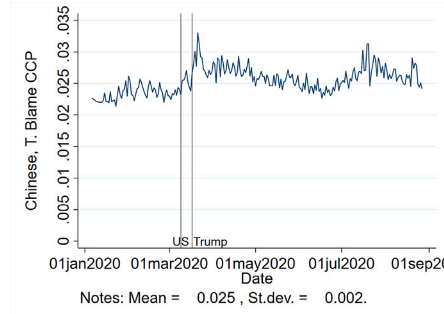
[17] Because the sample of Chinese users is smaller than the sample of White users, we employ a lower number of topics for the unsupervised machine learning exercise on the Chinese group.

[18] This topic can robustly be found also when we increase or decrease the number of topics to be retrieved by the algorithm.

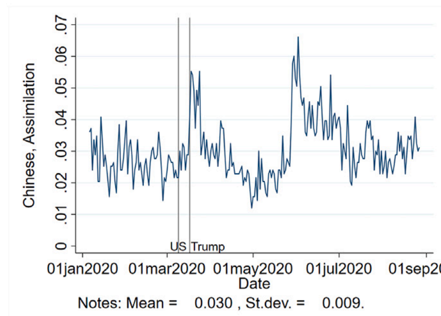
[19] Each document contains all the text tweeted by a single user in a given day, as defined in Section 3.2.



(a) Chinese/Asian American



(b) Topic Blame CCP



(c) Assimilation

Fig. 5. Assimilation Keywords, Expressions, and Topics for the Chinese Users. *Notes*: This Figure shows the daily share of Chinese users who tweeted "Chinese/Asian American" (Panel a), the daily average of the share of text attributed to the topic "blame the CCP" within the sample of Chinese users tweeting in a given day (Panel b), and the daily share of Chinese users who tweeted assimilation content (Panel c), from January 6 to August 27, 2020. Mean and standard deviation of the variables in the period of interest are reported below each graph.

the extent to which they embrace American culture with respect to their ethnicity of origin. For instance, in one of the sentences, an interviewee declares that she considers herself culturally American, while in another sentence another respondent reports feeling that America is her familiar world and China is something she had to learn. Another interviewee reports having deliberately behaved in the exact opposite way with respect to a Chinese classmate in order to be considered a "true" American by her school classmates. Then, we rely on a supervised machine-learning approach. In particular, we train a Word2vec algorithm on the corpus of all Chinese tweets, using it to compute a novel measure of assimilation content in four steps. First, we compute the cosine similarity between each tweet and the 55 assimilation sentences: specifically, we measure similarity to the centroid of the 55 vectorized sentences within the word embedding. Second, we determine the top percentile in the similarity score distribution for tweets posted before February 17, 2020, a period when COVID had not yet become perceived as a direct threat to U.S. society. Third, using this cutoff, we classify tweets as having assimilation content if their similarity score to the centroid of the assimilation sentences exceeds the cutoff.[20] Fourth, and finally, to build a measure at the user-date level, we create the dummy variable *Assimilation* which is equal to one if a user posted at least one tweet with assimilation content on a given day t, and zero otherwise. This is our third measure of assimilation; the unit of observation is at the user-day level.

Fig. 5 reports the evolution of the daily share of Chinese users who tweeted the selected keywords ("Chinese/Asian American"), the daily average share of text associated with blaming the CCP within the sample of documents posted by Chinese users, and the daily share of users who tweeted assimilation content. At the bottom of each graph, we report the mean and standard deviation of the respective variable in the sample. Consistently across all outcome variables, the figure shows a rise in assimilation language after March 9 and March 17, the latter increase being substantially more pronounced.

4. Empirical specification

We now estimate the causal effect of both the health shock and the discrimination shock on discriminatory and assimilation attitudes of the White and Chinese groups. To do so, we adopt two different specifications: (i) a linear trend-break model, and (ii) a regression discontinuity design. While the former allows us to investigate the effects of both shocks within the same empirical framework, the latter helps ensure that the estimated effect of each shock is not confounded by the effect of the other.

We start by considering the following linear trend-break model:

$$\begin{aligned} Y_{it} = {} & \beta_0 + \beta_1 t + \beta_2 FromMarch9 + \beta_3 FromMarch17 \\ & + \beta_4 FromMarch9 \times [t - March9] \\ & + \beta_5 FromMarch17 \times [t - March17] \\ & + \phi X_{it} + \varepsilon_{it} \end{aligned}$$

where Y_{it} is the probability of tweeting a certain keyword or content for user i on day t, or the average share of text about a certain topic for all users of a given group on day t. *FromMarch9* and *FromMarch17* are two dummy variables that take the value of one starting from March 9 and March 17, respectively. The main coefficients of interest are β_2 and β_3, which measure the intercept changes in the relationship between the dependent variable and the time upon occurrence of the health and

[20] Ash et al. (2021) use the cosine similarity directly as a dependent variable; here, we categorize this variable in order to be consistent with the rest of our analysis. Since the share of tweets with either "Chinese American" or "Asian American" is around 0.1 % in our dataset, the 1 % cutoff, albeit arbitrary, seems a reasonable benchmark.

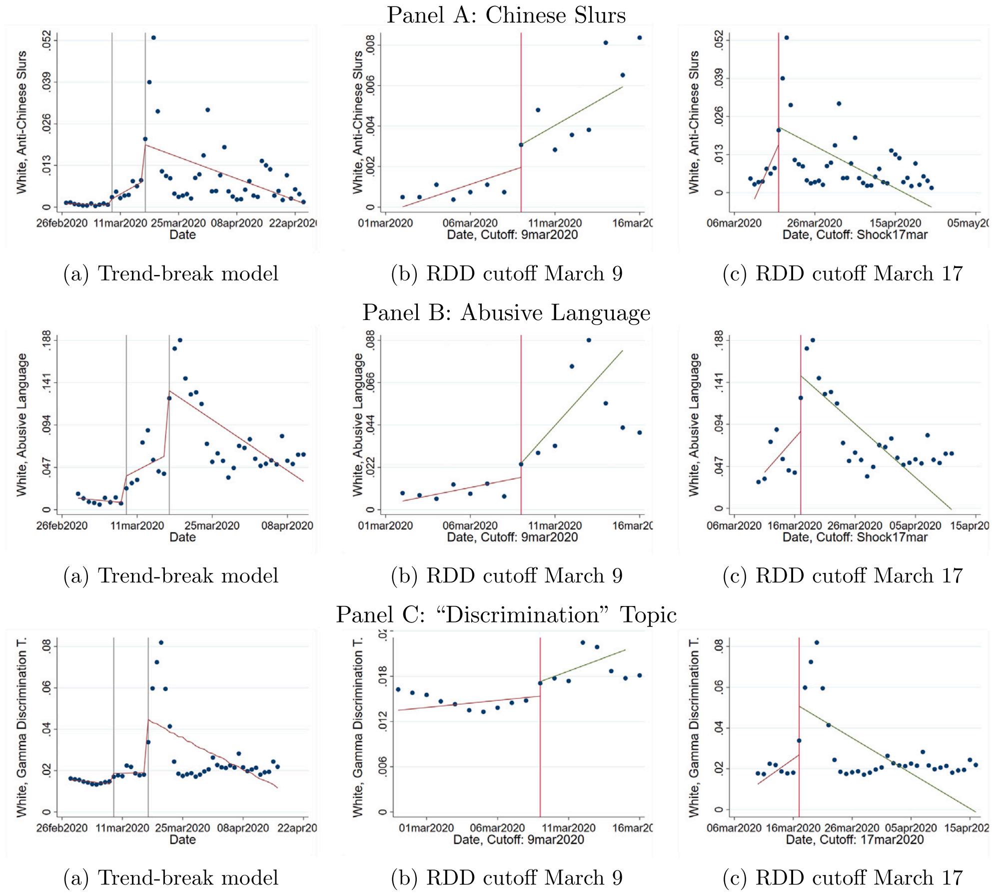


Fig. 6. Tweeting in Time “Chinese Slurs,” “Abusive Language,” and “Discrimination” Topic in Time. Panel A: Chinese Slurs. Panel B: Abusive Language. Panel C: “Discrimination” Topic *Notes*: In this Figure, we consider the sample of White American users. In Panel A, the dependent variable is a dummy taking the value one if the user tweeted Chinese slurs, while in Panel B it is a dummy taking the value one if the user posted a tweet containing both abusive language and the keyword “Chinese.” Dots represent averages of the dependent variable (y-axis) in each day (x-axis), while continuous lines are unconditional linear fits on the panel of user-day observations. In Panel C, the dependent variable is the average of the share of text on the “Discrimination” topic against the Chinese computed within the sample of users who tweeted that day. Dots represent the values of the dependent variable (y-axis) in each day (x-axis), while continuous lines are unconditional linear fits. In each panel, graph (a) depicts trend break model estimates, graph (b) shows RDD results using March 9 as cutoff, and graph (c) shows RDD results using March 17 as the cutoff. See Section 3.2 and Appendix A.2 for details on data construction and sources.

discriminatory shocks. We control for linear time trends, which switch on from March 9 and from March 17. The set of controls X_{it} includes day-of-the-week and month-of-the-year dummies to account for possible day- and month-specific tweeting patterns. We cluster standard errors at the treatment level, that is by date.

Next, using a regression discontinuity design, we estimate the impact of the two shocks using two truncated samples of dates. When focusing on March 9, we consider all days before March 17; when focusing on March 17, we consider all days after March 9. The estimating equations will be:

$$Y_{it} = \beta_0 + \beta_1 FromMarch9 + f(t) + \phi X_{it} + \varepsilon_{it}$$

or

$$Y_{it} = \beta_0 + \beta_1 FromMarch17 + f(t) + \phi X_{it} + \varepsilon_{it}$$

where *FromMarch9* in the first equation is a dummy taking the value one on all dates starting from March 9, and*FromMarch17* in the second equation is a dummy taking the value one on all dates starting from March 17. In both equations, the coefficient of interest is β_1. The forcing variable is t, and $f(t)$ is a polynomial function in the forcing variable with different coefficients on each side of the cutoff dates (March 9 and 17). As in the trend-break model, the set of controls X_{it} includes day-of-the-week and month-of-the-year dummies to account for possible day- and month-specific tweeting patterns. Note that our running variable, date, is somewhat imperfectly measured, since tweets are posted at specific hours and minutes of the day, resulting in a discrete rather than a continuous score. In this case, we follow Lee and Card (2008) and cluster the error term by date in the RD framework as well.[21] Appendix Section B.1 discusses the standard falsification analyses supporting the RDD assumptions based on our samples of White and Chinese Americans.

5. Results

5.1. Increasing discrimination content of white Americans

Fig. 6 illustrates the estimates of the impact of our health and leader-induced shocks on our measures of discriminatory content, both for the trend-break model (graph a) and for the regression discontinuity strategy (graphs b and c). In Panel A, the dependent variable is a dummy equal to one if, on a given date, a user tweeted Chinese slurs; in Panel B, the dependent variable is a dummy equal to one if, on a given date, a user posted a tweet containing both abusive language and the keyword “Chinese”; finally, in Panel C, the dependent variable is the daily average of the share of text related to the “Discrimination” topic,

[21] Results clustering the standard errors by weekdays or by weekday-by-month (available on request) align with our baseline findings.

Table 1
Discrimination: RDD Estimates.

Panel A: White sample, Cutoff March 9, 2020						
Dep. Var.	Anti-Chinese Slurs		Abusive Language		Discrimination T.	
	(1)	(2)	(3)	(4)	(5)	(6)
RD_Estimate	0.0023	0.0020	0.0151	0.0122	0.0032	0.0054
	(0.0006)	(0.0004)	(0.0048)	(0.0016)	(0.0007)	(0.0007)
Robust P-value	0.0001	0.0002	0.2315	0.2943	0.0191	0.0000
Observations Left	89430	65040	73170	65040	10	9
Observations Right	56910	56910	56910	48780	7	7
Polynomial Order	1	1	1	1	1	1
Band. Method	msetwo	msetwo	msetwo	msetwo	msetwo	msetwo
Band. Left	11.628	8.697	9.949	8.520	10.545	9.590
Band. Right	7.000	7.000	7.000	5.051	7.000	7.000
Day and Month Dummies		✓		✓		✓
Panel B: White sample, Cutoff March 17, 2020						
Dep. Var.	Anti-Chinese Slurs		Abusive Language		Discrimination T.	
	(1)	(2)	(3)	(4)	(5)	(6)
RD_Estimate	0.0129	0.0132	0.1146	0.1463	0.0337	0.0431
	(0.0053)	(0.0044)	(0.0153)	(0.0118)	(0.0079)	(0.0055)
Robust P-value	0.0090	0.0001	0.0000	0.0000	0.0000	0.0000
Observations Left	48780	48780	48780	48780	6	6
Observations Right	317070	252030	211380	154470	31	20
Polynomial Order	1	1	1	1	1	1
Band. Method	msetwo	msetwo	msetwo	msetwo	msetwo	msetwo
Band. Left	7.000	7.000	7.000	7.000	7.000	7.000
Band. Right	38.181	30.235	25.888	18.715	30.477	19.211
Day and Month Dummies		✓		✓		✓

Notes: In this table, we consider the sample of tweets of White American users between January 6 and March 16 in Panel A and between March 10 and August 27 in Panel B. In columns 1–2 of Panels A and B, the unit of observation is at user-day level and the dependent variable is a dummy taking the value one if the user tweeted Chinese slurs. In columns 3–4 of Panels A and B, the unit of observation is at the user-day level and the dependent variable is a dummy taking the value one if the user posted a tweet containing both abusive language and the keyword "Chinese." In columns 5–6 of Panels A and B, the unit of observation is at the day level and the dependent variable is the average of the share of text on the "Discrimination" topic computed within the sample of users who tweeted that day. Results are local polynomial estimates using March 9 as the cutoff in Panel A and March 17 as the cutoff in Panel B. Odd columns are unconditional, while even columns control for dummies for weekdays and months of the year. Standard errors, clustered by date in columns 1–4 and robust in columns 5–6, are reported in parentheses. Statistical significance is computed based on the robust p-value. Different bandwidths on each side of the cutoff are derived under the MSE procedure using a linear polynomial and a triangular kernel.

computed within the sample of users tweeting on a given day. Two main results emerge. First, discriminatory behavior, as proxied by all three measures, increased after March 9. Second, discrimination against the Chinese also spiked after Trump tweeted "Chinese virus." Empirically, we find this effect to be even larger than the effect of the health-related shock.

Formal estimates are reported in Table B4 in Appendix B.2.1 for the trend-break model and in Table 1 for the regression discontinuity approach.[22] In both tables, odd columns report unconditional estimates, while even columns report our preferred specifications controlling for dummies for weekdays and months of the year. In particular, focusing on column 4 of Table 1, the probability of tweeting abusive language against the Chinese shows an increase of more than 1 percentage point on March 9 (Panel A), and by almost 15 percentage points on March 17 (Panel b). Similarly, the average share of text related to anti-Chinese discrimination shows a positive jump when the health- and leader-induced shocks took place (+0.5 and +4.3 percentage points, respectively), with a more precisely estimated and larger magnitude on March 17.

Robustness. One possible concern is that our results are driven by specific sets of users included in the analysis. To alleviate this concern, we perturb the White sample in two main ways. First, we check that our results are robust to excluding politicians and activist users based on keywords in their bio (e.g., "congressman," "senator," or "feminist," "dissident," "activist").[23] Table B6 shows that the results generally hold, with slightly higher magnitudes in some cases. Second, in Table B7, we show that coefficients are quite stable and statistically significant when focusing on the subset of users specifically reporting to be politicians or activists in their bio. Thus, it seems that these specific groups of users are not driving the results.

Another possible concern is that our findings are driven by our samples of White users, through the use of specific keywords in their bio. To address this potential issue, we enlarge the sample of users in the White group by considering users who reported the keyword "American" in their bio.[24] Table B11 shows that our results hold both when we focus only on the sample of users reporting "American" in their bio, and when we lump our baseline sample of White Americans together with that of the "American" users. Moreover, results generally align with our baseline estimates if we account for unobserved state-specific characteristics (including differential state-level support for Trump) by controlling for state fixed effects as in Cao et al. (2023) (Table B13). Finally, Table B15 shows that our findings are also robust when we perform estimates using the local randomization approach in place of the continuity-based approach.

[22] RDD estimates are based on the truncated samples as described in Section 4, while trend-break model estimates are computed based on the sample of dates included in the left bandwidth when the cutoff is March 9 and in the right bandwidth when the cutoff is March 17, based on the mean squared error (MSE) procedure allowing for different bandwidths on each side of the cutoff.

[23] See Appendix A.2.2 for the full list of keywords that identify this subsample of users.

[24] This sample was originally retrieved around the same dates as the baseline White and Chinese samples; it contains 6,713 users.

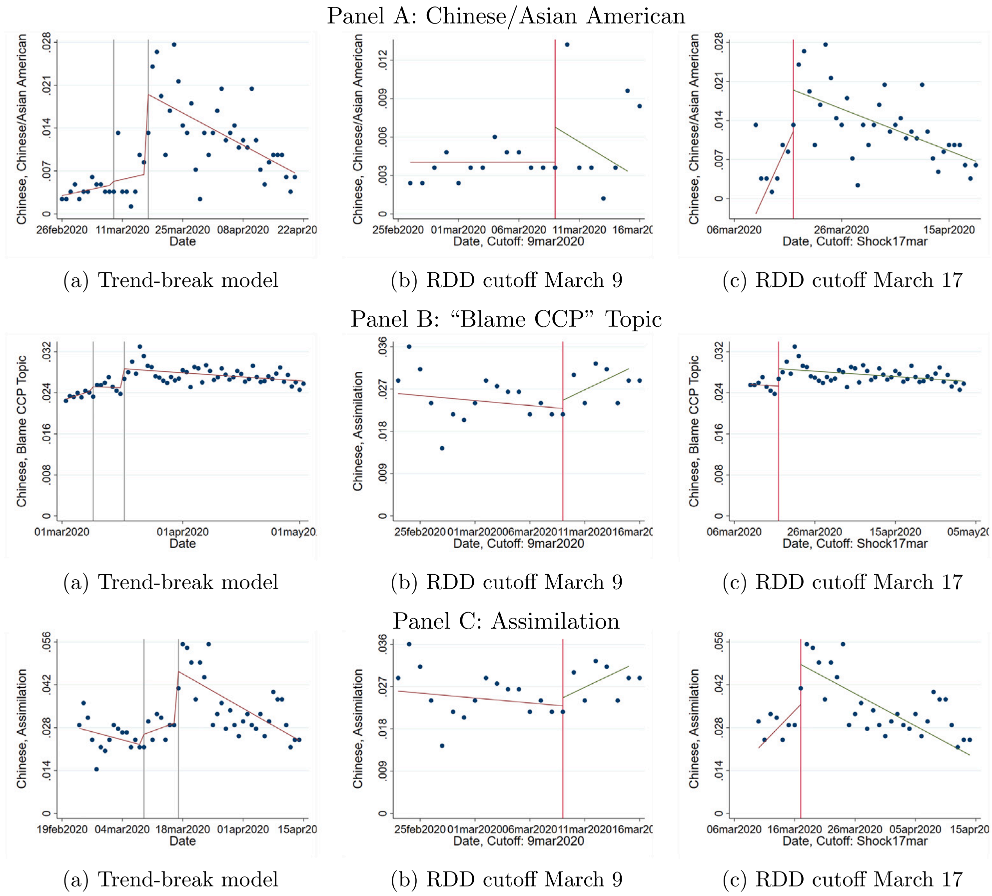


Fig. 7. Tweeting in Time "Chinese/Asian American," the Topic "Blame CCP", and "Assimilation" Content. Panel A: Chinese/Asian American. Panel B: "Blame CCP" Topic. Panel C: Assimilation. *Notes*: In this Figure, we consider the sample of Chinese users. In Panel A, the dependent variable is a dummy taking the value one if the user tweeted the keywords "Chinese American" or "Asian American," while in Panel C, it is a dummy taking the value one if the user tweeted assimilation content. Dots represent averages of the dependent variable (y-axis) in each day (x-axis), while continuous lines are unconditional linear fits on the panel of user-day observations. In Panel B, the dependent variable is the average of the share of text on the "Blame CCP" topic computed within the sample of users who tweeted that day. Dots represent the values of the dependent variable (y-axis) in each day (x-axis), while continuous lines are unconditional linear fits. In each panel, graph (a) depicts trend-break model estimates, graph (b) shows RDD results using March 9 as the cutoff, and graph (c) shows RDD results using March 17 as the cutoff. See Section 3.2 and Appendix A.2 for details on data construction and sources.

Overall, the results in this section point to an increase in discriminatory content by the majority group after both a health shock and a leader-induced shock. In the following section, we assess the reaction of the Chinese minority.

5.2. Reaction of the Chinese minority

We now empirically investigate how the Chinese minority reacted to the worsening discriminatory environment. Fig. 7 depicts the results of the trend-break model (graph a) and regression discontinuity designs (graphs b and c) to estimate the effect of the COVID-19 and leader-induced discrimination on the assimilation content of the Chinese minority. Panel A of Fig. 7 displays results using as the dependent variable a dummy equal to one if, on a given date, the user tweeted the keywords "Chinese American" or "Asian American;" Panel B displays results using the daily average of the share of text devoted to the "Blame CCP" topic, computed within the sample of users tweeting on a given day; finally, Panel C focuses on a dummy equal to one if, on a given date, the user tweeted assimilation content. None of the dependent variables (except for the "Blame CCP" topic) shows a significant jump when the COVID-19 shock occurred (March 9), while there is a positive and significant increase on March 17, following Trump's "Chinese virus" tweet.

Table B5 in Appendix B.2.1 and Table 2 provide the formal estimates of the linear trend-break model and of regression discontinuity, respectively. Odd columns report unconditional estimates, while even columns report our preferred specifications controlling for dummies for weekdays and months of the year. None of the dependent variables shows a significant jump on March 9, while there is a positive and significant increase on March 17, following Trump's "Chinese virus" tweet.[25] In particular, the political-discrimination shock is associated with a 1.18 percentage point increase in the probability of tweeting "Chinese/Asian American" (column 2), a 0.74 percentage point increase in the average share of text taking distance from the Chinese Communist Party as a major cause for the spread of the virus (column 4), and a 2.6 percentage point increase in the probability of tweeting assimilation content (column 6). In terms of magnitudes, these effects are important as they amount to a doubling of the probability of tweeting "Chinese/Asian American" and assimilation content, and to a one third increase in the daily average share of text blaming the CCP.

Robustness. We present three sets of exercises showing the robustness of our results to using alternative measures of assimilation content,

[25] This is consistent with the substantially larger increase in discrimination we find on March 17 than on March 9.

Table 2
Assimilation: RDD Estimates.

Panel A: Chinese sample, Cutoff March 9, 2020						
Dep. Var.	Chinese/Asian American		Blame CCP Topic		Assimilation	
	(1)	(2)	(3)	(4)	(5)	(6)
RD_Estimate	0.0025	−0.0004	−0.0001	0.0013	0.0011	0.0025
	(0.0034)	(0.0013)	(0.0008)	(0.0004)	(0.0030)	(0.0008)
Robust P-value	0.5278	0.8915	0.0086	0.0508	0.8274	0.5852
Observations Left	9984	8320	8	9	12480	3328
Observations Right	5824	5824	7	7	5824	5824
Polynomial Order	1	1	1	1	1	1
Band. Method	msetwo	msetwo	msetwo	msetwo	msetwo	msetwo
Band. Left	12.303	10.152	8.907	9.198	15.494	4.992
Band. Right	7.000	7.000	7.000	7.000	7.000	7.000
Day and Month Dummies		✓		✓		✓
Panel B: Chinese sample, Cutoff March, 17 2020						
Dep. Var.	Chinese/Asian American		Blame CCP Topic		Assimilation	
	(1)	(2)	(3)	(4)	(5)	(6)
RD_Estimate	0.0090	0.0118	0.0054	0.0074	0.0205	0.0260
	(0.0026)	(0.0028)	(0.0007)	(0.0008)	(0.0033)	(0.0033)
Robust P-value	0.0370	0.0001	0.0000	0.0000	0.0000	0.0000
Observations Left	4992	4992	6	6	4992	4992
Observations Right	29120	24960	47	28	24128	18304
Polynomial Order	1	1	1	1	1	1
Band. Method	msetwo	msetwo	msetwo	msetwo	msetwo	msetwo
Band. Left	7.000	7.000	7.000	7.000	7.000	7.000
Band. Right	34.807	29.237	46.665	27.461	28.076	21.931
Day and Month Dummies		✓		✓		✓

Notes: In this table, we consider the sample of tweets by Chinese users between January 6 and March 16 in Panel A and between March 10 and August 27 in Panel B. In columns 1–2 of Panels A and B, the unit of observation is at the user-day level and the dependent variable is a dummy taking the value one if the user tweeted the keywords "Chinese American" or "Asian American." In columns 3–4 of Panels A and B, the unit of observation is at the day level and the dependent variable is the average of the share of text on the topic "Blame CCP" computed within the sample of users who tweeted that day. In columns 5–6 of Panels A and B, the unit of observation is at user-day level and the dependent variable is a dummy taking the value one if the user posted a tweet containing assimilation content. Results are local polynomial estimates using March 9 as the cutoff in Panel A and March 17 as the cutoff in Panel B; odd columns are unconditional, while even columns control for dummies for weekdays and months of the year. Standard errors, clustered by date in columns 1–2 and 5–6 and robust in columns 3–4, are reported in parentheses. Statistical significance is computed based on the robust p-value. Different bandwidths on each side of the cutoff are derived under the MSE procedure using a linear polynomial and a triangular kernel.

alternative samples of Chinese users, and alternative estimation strategies. For all robustness checks – as the results in Fig. 7 and Table 2 are mostly not significant on March 9 – we focus on the shock that occurred on March 17.

First, Table 3 shows results using alternative definitions of our assimilation measures. In columns 1–5, we focus on our first measure, i.e., the dummy tracking whether the user tweeted "Chinese" or "Asian" combined with "American." A careful reading of the tweets reveals that when expressing particularly strong assimilation content users also included the pronoun "we" together with "Chinese American" or "Asian American." Therefore, in column 1 of Table 3, our dummy *Chinese/Asian American* is equal to one when the keyword "we" is also included in the tweet. The result is robust – although the coefficient is lower in magnitude with respect to column 2 of Table 2 (Panel B). Next, one possible concern is that our results were driven by tweets using "Chinese American" or "Asian American" in the context of reported acts of discrimination against the Chinese and/or broader Asian community. Thus, in columns 2–5 we recode to zero the dependent variable when the tweet includes different sets of keywords signaling reported discrimination. In particular, in column 2 we set our dummy to zero if the tweet includes one keyword within the following list: "racist, racism, xenophob, report, hate, spit, yell, incident, harass, anti-Asian, blame, discriminate, affirmative action, hatred, attack, scapegoat." Afterward, we further recode to zero our dummy if the tweet includes one keyword within an expanded list. This contains the keywords in the previous list and: (i) "Chinesevirus," which could specifically account for cases commenting on the content of Trump's tweet (column 3); or (ii) "Trump," possibly accounting for tweets commenting on the president's behavior (column 4); or (iii) both "Chinesevirus" and "Trump" (column 5). The results in columns 2–5 suggest that our findings are robust when accounting for possible misinterpretations. Then, in columns 6–7 of Table 3, we assess the robustness of our third measure of assimilation, based on the supervised-machine-learning exercise. To rule out that tweets containing the keyword "feel" (and variants) might report a general feeling rather than a sense of "Americanness," we set to zero the dependent variable if the text also included the keyword "feel" (and variants) (column 6). The result is virtually unchanged with respect to column 6 of Table 2 (Panel B). Moreover, in column 7 we perturb the set of assimilation sentences at the basis of the word2vec algorithm to consider only sentences that include the keyword "American." The result remains positive, significant, and comparable in magnitude with respect to the baseline.[26]

Second, to mitigate the concern that our results are driven by specific sets of users, we perform two main exercises. In the first exercise, we show the robustness of results to different subsamples of the Chinese users. In columns 1–3 of Table 4, we present estimates excluding politicians and activist users. All results still hold, with only small changes in magnitudes. The results are also qualitatively similar if we focus on the sample of users who report being politicians or activists (columns 4–6). These results suggest that politicians and activists are not driving our findings. In the second exercise, to further alleviate the concern that our results stem from selecting Chinese users by specifically

[26] Our results are also robust if we expand the baseline set of sentences to include 36 tweets with assimilation content and when we use a large language model (GPT4o) to detect posted assimilation content in the tweets of the Chinese Americans. See Appendix Section B.2.6.

Table 3
Robustness Chinese/Asian American and Assimilation: RDD Estimates, Cutoff March 17.

Dep. Var.	Chinese/Asian American					Assimilation	
Variant	Recoded 0 if NOT We	Report	Report,CV	Report,Trump	Report,CV,Trump	Recoded 0 if Feel	Sentences American
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
RD_Estimate	0.0059	0.0074	0.0049	0.0031	0.0035	0.0259	0.0261
	(0.0012)	(0.0021)	(0.0019)	(0.0014)	(0.0016)	(0.0033)	(0.0042)
Robust P-value	0.0000	0.0011	0.0143	0.0133	0.0367	0.0000	0.0000
Observations Left	4992	4992	4992	4992	4992	4992	4992
Observations Right	30784	24960	28288	27456	28288	16640	19136
Polynomial Order	1	1	1	1	1	1	1
Band. Method	msetwo	msetwo	msetwo	msetwo	msetwo	msetwo	msetwo
Band. Left	7.000	7.000	7.000	7.000	7.000	7.000	7.000
Band. Right	36.583	29.650	33.671	32.848	33.984	19.446	22.134
Day and Month Dummies	✓	✓	✓	✓	✓	✓	✓

Notes: In this table, we consider the sample of tweets from Chinese users between March 10 and August 27. In all columns, the unit of observation is at the user-day level. The dependent variable is a dummy taking the value one if the user tweeted the keywords "Chinese/Asian American" and "we" in column 1. In column 2, the dependent variable is a dummy taking the value one if the user tweeted the keywords "Chinese/Asian American" but did not tweet one keyword within the following list of keywords signaling reported discrimination: "racist", "racism", "xenophob", "report", "hate", "spit", "yell", "incident", "harass", "anti-Asian", "blame", "discriminate", "affirmative action", "hatred", "attack", and "scapegoat." Column 3 adds to the list the keyword "chinesevirus," column 4 adds to the list the keyword "Trump," and column 5 adds to the list both "chinesevirus" and "Trump." In column 6, the dependent variable is a dummy taking the value one if the user tweeted assimilation content but did not tweet the keyword "feel." In column 7, the dependent variable is a dummy taking the value one if the user tweeted assimilation content based on the subset of assimilation sentences including the keyword "American." Results are local polynomial estimates using March 17 as the cutoff, controlling for dummies for weekdays and months of the year. Standard errors clustered by date are in parentheses, and statistical significance is computed based on the robust p-value. Different bandwidths on each side of the cutoff are derived under the MSE procedure using a linear polynomial and a triangular kernel.

Table 4
Robustness Assimilation: RDD Estimates, Cutoff March 17.

Sample	Excluding politicians and activists			Only Politicians and activists		
Dep. Var.	Chin./Asian Amer.	Blame CCP T.	Assimilation	Chin./Asian Amer.	Blame CCP T.	Assimilation
	(1)	(2)	(3)	(4)	(5)	(6)
RD_Estimate	0.0122	0.0056	0.0247	0.0145	0.0134	0.0201
	(0.0032)	(0.0007)	(0.0032)	(0.0051)	(0.0013)	(0.0121)
Robust P-value	0.0000	0.0000	0.0000	0.1209	0.0000	0.5360
Observations Left	4410	6	4410	582	6	582
Observations Right	19110	44	14700	3104	26	3395
Polynomial Order	1	1	1	1	1	1
Band. Method	msetwo	msetwo	msetwo	msetwo	msetwo	msetwo
Band. Left	7.000	7.000	7.000	7.000	7.000	7.000
Band. Right	25.240	43.554	19.351	31.757	25.463	34.203
Day and Month Dummies	✓	✓	✓	✓	✓	✓

Notes: In this table, we consider the sample of tweets by Chinese users, between March 10 and August 27. Columns 1–3 focus on the subsample of users whose Twitter bio does not include keywords related to politics and activism and columns 4–6 focus on the subsample of users whose Twitter bio includes keywords related to politics and activism. In columns 1 and 4, the unit of observation is at the user-day level and the dependent variable is a dummy taking the value one if the user tweeted the keywords "Chinese American" or "Asian American." In columns 2 and 5, the unit of observation is at day level and the dependent variable is the average of the share of text on the topic "Blame CCP" computed within the sample of users who tweeted that day. In columns 3 and 6, the unit of observation is at the user-day level and the dependent variable is a dummy taking the value one if the user posted a tweet containing assimilation content. The results are local polynomial estimates using March 17 as the cutoff, controlling for dummies for weekdays and months of the year. Standard errors, clustered by date in columns 1, 3, 4, and 6, and robust in columns 2 and 5 are reported in parentheses. Statistical significance is computed based on the robust p-value. Different bandwidths on each side of the cutoff are derived under the MSE procedure using a linear polynomial and a triangular kernel.

considering those stating their ethnicity in their bio and already on Twitter before December 31, 2019, we further add to our baseline sample three additional sets of Chinese users. These are: (i) users who accessed Twitter for the first time after December 31, 2019; (ii) users with the most common Chinese names and/or surnames from an ancillary search of users with the keyword "American" in their bio; and (iii) the set of "American" users who posted tweets on the Chinese New Year (after manually checking their ethnicity). Table B12 shows the robustness of our results. These exercises suggest that the rise in assimilation behavior after March 17 is not driven by our initial set of users.

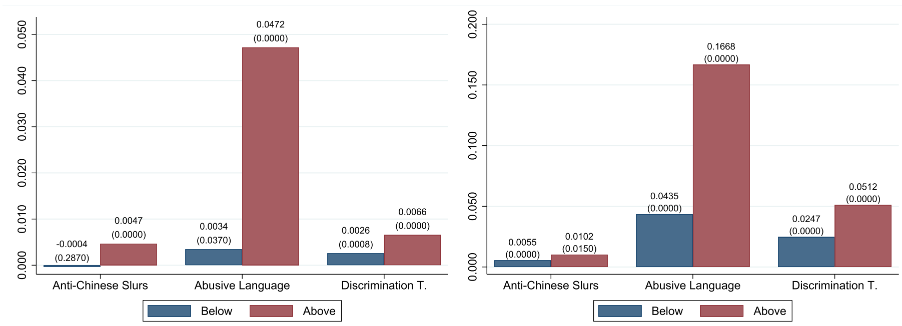
Third, we show the robustness of our findings by accounting for unobserved state-level characteristics using state fixed effects (Table B14) and by performing estimates using the local randomization approach (Table B16).

Finally, the results of the unsupervised text-analysis exercise show that *Blaming the CCP* was, among all other topics, the one that experienced the highest increase after Trump's "Chinesevirus" tweet (Figure B2).

Altogether, the findings in this section point to an increase in assimilation content by the minority as a response to a worsening discriminatory environment, triggered by a political leader: in particular, minorities tend to assert that they belong to the majority group and to distance themselves from their original Chinese identity.

6. Heterogeneity of results

We now consider possible heterogeneous effects stemming from pre-shock individual levels of discrimination and assimilation in U.S. society.



(a) Cutoff: 9 March 2020 (b) Cutoff: 17 March 2020

Fig. 8. Discrimination: Heterogeneity by Levels of (General) Abusive Language Before February 17, 2020. *Notes*: The bars represent the point estimates of separately replicating even specifications in Panels A and B of Table 1 on the subset of users with pre-shock levels of generalized abusive language below the median (blue) and above the median (maroon). The dependent variables are indicated on the x-axis. The magnitude of the coefficients and the robust p-values are reported in parentheses above the bars. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

6.1. Heterogeneity of discrimination behavior of the white group

We start by investigating whether the discrimination effects found within the White group depend on a generic initial individual propensity to discriminate. To measure this, we calculate the share of days a user posts at least one tweet using generalized abusive language from January 6 (the first day available in our sample) to February 17 (up to one month before Trump's "Chinese virus" tweet). Therefore, we consider all instances of abusive language, and not only the ones directed against the Chinese community. Panels (a) and (b) of Fig. 8 report conditional estimates for all our proxies of discriminatory content, comparing results from separate regressions on the subsample of White users with pre-shock shares of generalized abusive language below vs. above the median using as cutoff dates March 9 and 17, respectively. Across all proxies of discriminatory content, we find an increase in the discriminatory behavior for both groups as a result of both shocks; the increase triggered by both shocks is, however, larger and statistically significant for the set of users whose share of pre-shock generalized abusive language was above the median.[27]

6.2. Heterogeneity of assimilation behavior of the Chinese minority

We now explore potential heterogeneous effects stemming from pre-shock individual levels of assimilation in U.S. society. In particular, we investigate whether the change in assimilation content after Trump's "Chinese virus" tweet found in the baseline analysis depends on the Chinese users' use of assimilation content before February 17 (one month before the most important shock). Fig. 9 reports conditional estimates for our proxies of assimilation content from separate regressions for the subsamples of Chinese users never posting assimilation content and those posting assimilation content at least once before February 17. March 17 is used as the cutoff date. Across all proxies, we find a rise in the assimilation content of both groups; however, the increase is much

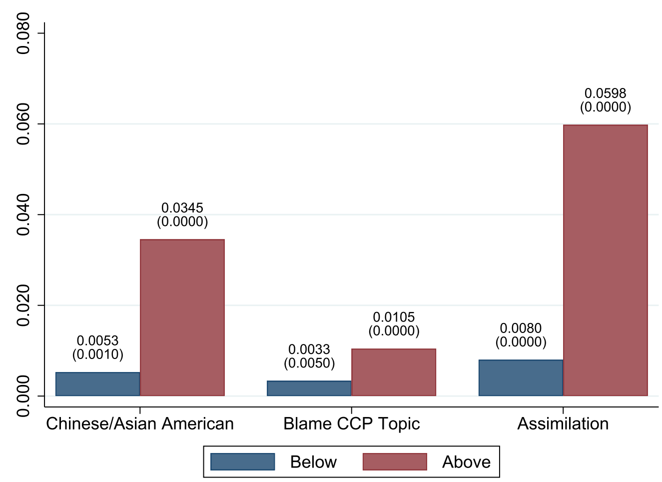


Fig. 9. Assimilation: Heterogeneity by Levels of Assimilation Tweets before February 17, 2020. *Notes*: The bars represent the point estimates of separately replicating even specifications in Panel B of Table 2. The results are based on the subset of users who never posted assimilation content before the shocks (blue) and users who previously posted assimilation content (maroon). The dependent variables are indicated on the x-axis. The magnitude of the coefficient and the robust p-value in parentheses are reported above the bars. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

more pronounced for the set of users who previously posted assimilation content.

7. Conclusion

Immigrants' assimilation into the host country has always been a crucial issue in the U.S. and other major receiving countries. Crucially, assimilation decisions of a minority group are not taken in isolation; they often result from dynamic interactions with the majority group and political leaders in a context with varying levels of discrimination. We exploit novel social media data to study the discriminatory behavior of White Americans and the assimilation behavior of the Chinese American community in the U.S. To do this, we leverage two shocks to the discriminatory environment: (i) the COVID-19 outbreak on March 9, 2020, and (ii) Trump's reference to COVID-19 as the "Chinese virus" on March 17, 2020.

Three major results stand out. First, we show that users of the White group tended to post stronger discriminatory content after both shocks,

[27] We also perform a heterogeneity analysis focusing on the subsamples of users who never posted anti-Chinese abusive language before February 17 vs. users who posted anti-Chinese abusive language at least once before February 17. Appendix Figure B1 displays the results. Across all discrimination proxies, following both shock dates, we find a stronger increase in the discrimination behavior for the set of users that had already posted anti-Chinese abusive language prior to the shocks. However, also for the set of users who had never posted anti-Chinese abusive language, the probability to discriminate significantly increases after March 17. This suggests that, besides unleashing greater discriminatory behavior in those who were already engaged in some anti-Chinese abusive language, Trump's behavior also engendered discriminatory behavior from users who were not used to discriminate the Chinese group.

but the effect was stronger after Trump's discriminatory words. Second, we find that Chinese users significantly responded to the rise in discriminatory content following Trump's "Chinese virus" tweets by asserting more frequently their Americanness and by increasingly distancing themselves from the Chinese Communist Party, a distinctive feature of China in general, and specifically related to the management of the COVID-19 epidemic. Third, both sets of results are generally stronger when we consider users with higher pre-shock levels of discrimination and assimilation.

Although our sample of Chinese Americans on Twitter is not very large, possibly due to sample selection issues, our results are in line with the qualitative findings of Kibria (2000), emphasizing the perceived need of Chinese (and other Asian) Americans to counteract the strong connotation of "foreignness" associated with the Asian race. Moreover, our evidence is consistent with the empirical findings of Fouka (2019) and Saavedra (2021), focusing on the Americanization of German immigrants in the U.S. during World War I and the Americanization of Japanese immigrants in the U.S. after the bombing of Pearl Harbor, respectively. Future empirical work should study how other discriminatory shocks affect the behavior of other minorities.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.jpubeco.2025.105450.

Data availability

Data will be made available upon request, subject to the terms and conditions of our agreement with Twitter.

References

Abdelgadir, A., Fouka, V., 2020. Political secularism and muslim integration in the west: assessing the effects of the French headscarf ban. Am. Polit. Sci. Rev. 114 (3), 707–723.
Abramitzky, R., Boustan, L., Eriksson, K., 2020. Do immigrants assimilate more slowly today than in the past? Am. Econ. Rev. insights 2 (1), 125–141.
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.
Ash, E., Chen, D., Naidu, S., 2021. Ideas Have Consequences: The Impact of Law and Economics on American Justice.
Aspachs-Bracons, O., Clots-Figueras, I., Costa-Font, J., Masella, P., 2008. Compulsory language educational policies and identity formation. J. Eur. Econ. Assoc. 6 (2–3), 434–444.
Avitabile, C., Clots-Figueras, I., Masella, P., 2013. The effect of birthright citizenship on parental integration outcomes. J. Law Econ. 56 (3), 777–810.
Becker, S.O., Pascali, L., 2019. Religion, division of labor, and conflict: anti-semitism in Germany over 600 years. The Am. Econ. Rev. 109 (5), 1764–1804.
Bertrand, M., Duflo, E., 2017. Field experiments on discrimination. Handb. Econ. Field Exp. 1, 309–393.
Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2008. Are muslim immigrants different in terms of cultural integration? J. Eur. Econ. Assoc. 6 (2–3), 445–456.
Bisin, A., Tura, G., 2019. Marriage, fertility, and cultural integration in italy. Technical report, National Bureau of Economic Research.
Bleakley, H., Chin, A., 2010. Age at arrival, English proficiency, and social assimilation among us immigrants. Am. Econ. J. Appl. Econ. 2 (1), 165–192.
Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet Allocation. J. Mach. Learn. Res. 3 (Jan), 993–1022.
Bursztyn, L., Egorov, G., Enikolopov, R., Petrova, M., 2019. Social media and xenophobia: evidence from Russia. Technical report, National Bureau of Economic Research.
Bursztyn, L., Egorov, G., Fiorin, S., 2020. From extreme to mainstream: the erosion of social norms. Am. Econ. Rev. 110 (11), 3522–3548.
Cagé, J., Dagorret, A., Grosjean, P.A., Jha, S., 2023. Heroes and villains: the effects of heroism on autocratic values and nazi collaboration in France. Am. Econ. Rev. 113 (7), 1888–1932.
Cao, A., Lindo, J.M., Zhong, J., 2023. Can social media rhetoric incite hate incidents? evidence from trump's "chinese virus" tweets. J. Urban Econ. 137, 103590.
Clots-Figueras, I., Masella, P., 2013. Education, language and identity. Econ. J. 123 (570), F332–F357.
Davidson, T., Warmsley, D., Macy, M., Weber, I., 2017. Automated hate speech Detection and the problem of offensive language. In: Proceedings of the Eleventh International AAAI Conference on Web and Social Media, ICWSM 2017, pp. 4.
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.
Devakumar, D., Shannon, G., Bhopal, S.S., Abubakar, I., 2020. Racism and discrimination in covid-19 responses. Lancet. 395 (10231), 1194.
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
Dippel, C., Heblich, S., 2021. Leadership in social movements: evidence from the "forty-eighters" in the civil war. The Am. Econ. Rev. 111 (2), 472–505.
Dustmann, C., Preston, I., 2001. Attitudes to ethnic minorities, ethnic context and location decisions. Econ. J. 111 (470), 353–373.
Eguia, J.X., 2017. Discrimination and assimilation at school. J. Public Econ. 156, 48–58.
Fernández, R., Fogli, A., 2009. Culture: an empirical investigation of beliefs, work, and fertility. Am. Econ. J.: macroecon. 1 (1), 146–177.
Fouka, V., 2019. How do immigrants respond to discrimination? the case of germans in the US during world war I. Am. Polit. Sci. Rev. 113 (2), 405–422.
Fouka, V., 2020. Backlash: the unintended effects of language prohibition in us schools after world war i. Rev. Econ. Stud. 87 (1), 204–239.
Founta, A.-M., Djouvas, C., Chatzakou, D., Leontiadis, I., Blackburn, J., Stringhini, G., Vakali, A., Sirivianos, M., Kourtellis, N., 2018. Large scale crowdsourcing and characterization of twitter abusive behavior. In: 11th International Conference on Web and Social Media, ICWSM 2018, AAAI Press.
Giuliano, P., 2007. Living arrangements in western europe: does cultural origin matter? J. Eur. Econ. Assoc. 5 (5), 927–952.
Goffman, E., November 2009. Stigma: Notes on the Management of Spoiled Identity, Simon and Schuster.
Gould, E.D., Klor, E.F., 2016. The long-run effect of 9/11: terrorism, backlash, and the assimilation of muslim immigrants in the west. Econ. J. 126 (597), 2064–2114.
Griffiths, T.L., Steyvers, M., April 2004. Finding scientific topics. Proc. Natl. Acad. Sci. USA 101 (suppl 1), 5228–5235.
Grosfeld, Irena, Sakalli, S.O., Zhuravskaya, E., 2020. Middleman minorities and ethnic violence: anti-jewish pogroms in the Russian empire. Rev. Econ. Stud. 87 (1), 289–342.
Grosjean, P.A., Masera, F., Yousaf, H., 2020. Whistle the racist dogs: political campaigns and police stops. UNSW Bus. Sch. Res. Pap. Forthcom..
Guirkinger, C., Platteau, J.-P., Wahhaj, Z., 2021. Behind the Veil of Cultural Persistence: Marriage and Divorce in a Migrant Community.
Hansen, S., McMahon, M., Prat, A., May 2018. Transparency and deliberation Within the FOMC: A Computational Linguistics Approach*. Q. J. Econ. 133 (2), 801–870.
Hastie, T., Tibshirani, R., Friedman, J. ed), 2009. The Elements of Statistical Learning, 2nd edition ed. Springer..
Holshue, M.L., DeBolt, C., Lindquist, S., Lofy, K.H., Wiesman, J., Bruce, H., Spitters, C., Ericson, K., Wilkerson, S., Tural, A., et al, 2020. First case of 2019 novel coronavirus in the United States. N. Engl. J. Med..
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.
Kibria, N., March 2000. Race, ethnic Options, and ethnic Binds: identity Negotiations of second-generation Chinese and korean Americans. Sociol. Perspect. 43 (1), 77–95.
Kim, Y.-C., Loury, G.C., 2019. To be, or not to be: stereotypes, identity choice and group inequality. J. Public Econ. 174, 36–52.
Lee, D.S., Card, D., 2008. Regression discontinuity inference with specification error. J. Econom. 142 (2), 655–674.
Lu, R., Sheng, Y., 2020. From fear to hate: how the covid-19 pandemic sparks racial animus in the United States. arXiv preprint arXiv:2007.01448.
Manning, A., Roy, S., 2010. Culture clash or culture club? national identity in Britain.
Mayda, A.M., 2006. Who is against immigration? a cross-country investigation of individual attitudes toward immigrants. The Rev. Econ. Stat. 88 (3), 510–530.
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. Econ. Assoc. 19 (4), 2131–2167.
NPR, 2020. Coronavirus: all 50 states report cases; south America has nearly 1, 000 cases. Accessed Sep. 9, 2021 [Online], available at https://www.npr.org/sections/health-shots/2020/03/17/817096232/coronavirus-radical-change-to-life-as-covid-19-reaches-152-countries?t=1631182083547
Saavedra, M., 2021. Kenji or kenneth? pearl harbor and Japanese-American assimilation. J. Econ. Behav. & Organ. 185, 602–624.
Schmidt, A.L., Zollo, F., Del Vicario, M., Bessi, A., Scala, A., Caldarelli, G., Stanley, H.E., Quattrociocchi, W., 2017. Anatomy of news consumption on facebook. Proc. Natl. Acad. Sci. USA 114 (12), 3035–3039.
Sunstein, C.R., 2018. # Republic: Divided Democracy in the Age of Social Media, Princeton University Press.
Tahmasbi, F., Schild, L., Ling, C., Blackburn, J., Stringhini, G., Zhang, Y., Zannettou, S., 2021. "Go eat a bat, chang!": on the emergence of sinophobic behavior on web communities in the face of covid-19. In: Proceedings of the Web Conference 2021, pp. 1122–1133.
Vidgen, B., Hale, S., Guest, E., Margetts, H., Broniatowski, D., Waseem, Z., Botelho, A., Hall, M., Tromble, R., November 2020. Detecting east Asian Prejudice on social media.

In: Proceedings of the Fourth Workshop on Online Abuse and Harms, Association for Computational Linguistics, Online, pp. 162–172.

Voigtländer, N., Voth, H.-J., 2012. Persecution perpetuated: the medieval origins of anti-semitic violence in nazi Germany. The Quarterly Journal Of Economics 127 (3), 1339–1392.

Waseem, Z., Hovy, D., June 2016. Hateful symbols or hateful People? predictive Features for hate speech Detection on twitter. In: Proceedings of the NAACL Student Research Workshop, Association for Computational Linguistics, San Diego, California, pp. 88–93.

White, A.I.R., 2020. Historical linkages: epidemic threat, economic risk, and xenophobia. Lancet 395 (10232), 1250–1251.

Yanagizawa-Drott, D., 2014. Propaganda and conflict: evidence from the rwandan genocide. Q. J. Econ. 129 (4), 1947–1994.

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