Perceived Popularity and Online Political Dissent: Evidence from Twitter in Venezuela

Juan S. Morales*

Abstract

On October 31st, 2013, thousands of Twitter accounts, automated to actively retweet President Nicolas Maduro of Venezuela, were unexpectedly closed by the social media platform. I exploit this event to study the relationship between perceived popularity on social media (amplified through the use of bot accounts) and online political expression. The analysis uses over 200,000 tweets spanning six months around the event and employs a quasi-experimental empirical framework. Following the closure of the accounts, the volume of tweets mentioning the president increased by an estimated 33 percent, with a differential increase for critical messages. Relative to tweets by government leaders, the number of likes for tweets by opposition leaders increased by an estimated 21 percent. Consistent with the presence of a spiral of silence in online political expression, the results suggest that the change in the perceived popularity of Maduro led to an increase in users' willingness to express both criticism of the president and support for the opposition. While previous studies have documented how autocratic governments engage in manipulative online campaigns, this paper provides evidence of their effectiveness and highlights an important mechanism through which they can influence behaviour.

Keywords: political expression, social media, bots, spiral of silence, event study
Word count: 8,000

*Collegio Carlo Alberto and University of Turin.
Piazza Arbarello, 8, 10122, Torino, TO, Italy. E-mail: juan.morales@carloalberto.org

Acknowledgements: I thank the editor, Cristian Vaccari, and four anonymous referees for their careful reviews and helpful suggestions. I also thank Krzysztof Krakowski, Ignacio Monzón, and Jean-William Laliberté for productive discussions. Nicolas Bohorquez and Patricia Garza provided excellent research assistance. All errors are my own.

Introduction

Campaigns to manipulate information are pervasive in online social networks (Ferrara et al, 2016; Allcot and Gentzkow, 2017; Lazer et al, 2018, Allcot et al, 2019). Inflated popularity statistics through the use of fake followers, or bots, are one common form of social media manipulation (investigations have found that between 29 and 60 percent of President Donald Trump's Twitter followers are not real).[1] Though media reports on how such fake accounts affect political dynamics are common,[2] studies investigating these causal relationships using quasi-experimental design methods are not. At the same time, several studies have documented how governments engage in various forms of strategic online campaigns (King et al, 2017; Keller et al, 2017; Field et al, 2018), but our understanding of their effectiveness remains limited. Can coordinated online campaigns help governments abate dissent?

This study exploits the exogenous closure of thousands of Twitter accounts, programmed to retweet a prominent political leader, to study political engagement on social media. The mechanism I propose by which the retweeting activity mattered, through changes in perceived popularity that can induce or dissolve online 'spirals of silence', is an important way in which bots are likely to affect political expression, and especially so in autocratic regimes.

On October 31st, 2013, over six thousand Twitter accounts, all of which actively supported President Nicolas Maduro of Venezuela by retweeting his messages, were unexpectedly closed by the social media platform. Though there was no official announcement from Twitter, the evidence suggests these bot accounts violated Twitter's rules by being automatically programmed to retweet the president's own tweets. The accounts represented less than 0.5 percent of Maduro's total followers but their closure led to an 81 percent drop in the president's average number of retweets. I study the relationship between perceived popularity and political expression on social media through an empirical investigation of this event. In particular, I use approximately 220,000

[1] https://www.thetimes.co.uk/article/donald-trump-has-15m-fake-followers-but-tweets-still-hit-target-trc07kqhf and https://qz.com/1422395/how-many-of-donald-trumps-twitter-followers-are-fake/.

[2] See for instance: https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html; https://www.scientificamerican.com/article/how-twitter-bots-help-fuel-political-feuds/; and https://www.bbc.com/news/technology-40344208

tweets, published within a six-month window around the date of the event, to examine patterns in political expression in a quasi-experimental framework.

The research design exploits the accounts' cancelation in both pre/post event study and difference-in-differences analyses to investigate the extent to which social media engagement was affected. I use three main datasets collected from Twitter public feeds: tweets published by @NicolasMaduro; tweets that mention @NicolasMaduro; and tweets by approximately 150 selected users, including a combination of prominent opposition and government leaders, in addition to a set of Venezuelan sports-related accounts (which is used as a control group). I then look at measures of social media engagement and tweet volume in a narrow time-window around the date of the event, comparing outcomes 15 days after the cancelations, relative to 15 days before.

I document three main results. Following the closure of the retweeting accounts: users become more willing to reply to tweets by @NicolasMaduro; the volume of tweets mentioning @NicolasMaduro increases by an estimated 33 percent, with a differential increase for anti-Maduro tweets, relative to pro-Maduro tweets; and tweets from opposition leaders receive on average 21 percent more likes relative to tweets by government leaders.

To formalize the analysis, I propose a theoretical framework based on the presence of a spiral of silence in online political expression. The spiral of silence theory, introduced in Noelle-Neumann (1974), proposes that individuals' perceptions of dominating public opinions decrease their willingness to express minority views. Several studies have documented findings consistent with these dynamics in online settings, including in news websites (Soffer and Gordoni, 2018; Wu and Atkin, 2018), and Facebook (Gearhart and Zhang, 2013, Liu et al, 2017; Zerback and Fawzi, 2017).[3] The event I study allows me to examine hypotheses about how perceptions of dominating public opinions, as inferred by retweet activity, affect political expression on social media. The evidence I present is consistent with the presence of a spiral of silence in online political expression, and it highlights the relevance of the framework for studies of politics and social media in modern autocracies.

[3] See Matthes et al (2018) for a recent review and meta-analysis of research in the spiral of silence.

The article contributes to a rapidly emerging literature on social media and politics (reviewed in Woolley and Howard, 2016; and Tucker et al, 2018). Several studies have documented relationships between social media use and social media derived-outcomes and other measures of: political participation (Gil de Zúñiga et al, 2014; Boulianne, 2015; Vaccari et al, 2015; Skoric et al, 2016), public support (O'Connor, 2010; DiGrazia et al, 2013; Ceron et al, 2014; Barberá, 2016) and political ideology (Colleoni et al, 2014; Barberá, 2015; Halberstam and Knight, 2016). The scope of the analysis focuses on Twitter engagement in the form of posts and likes. Understanding the determinants of these "tiny acts of political participation" is important for understanding larger political dynamics (Margetts et al, 2016). Closest to the aim of examining these online dynamics in a quasi-experimental empirical framework is Gorodnichenko et al (2018), who find that, during both the 2016 US and the Brexit campaigns, bots had a significant effect on the tweeting activity of ideologically-similar humans.

Several studies have highlighted both the promise and downsides of social media as a tool for transforming politics and advancing democracy (Tufekci and Wilson, 2012; King et al, 2013; Tucker et al, 2017; Qin et al, 2016; Enikolopov et al, 2017; Suhay et al, 2018; Jost et al, 2018). Previous work has documented that world leaders adopt social media in times of political pressure and social unrest (Barberá and Zeitzoff, 2018), and portraying a distorted image of popular opinion constitutes an important component of the social media strategy of autocratic governments (King et al, 2017). Waisboard and Amado (2017) highlight that, despite its promise, social media in Latin America did not transform the way that politicians communicate with citizens. Instead, social media has extended the traditional broadcast structure of leader communications and has been used by populists, including Maduro, to harass critics. Munger et al (2018) document how the Venezuelan government hoped to distract the public by tweeting about a wide range of topics in times of political tension, aiming to make collective action, and protests in particular, more difficult. This article contributes to this literature by presenting evidence on a previously undocumented case of information distortion in Venezuela, but which has broader implications for understanding the determinants of online political dissent and the mechanisms through which manipulative online campaigns can affect behaviour.

Background

Nicolas Maduro became leader of Venezuela on March 5th, 2013, after the death of the then president, Hugo Chavez. Nicolas Maduro's first tweet was published twelve days later (Figure 1). The account quickly amassed a large number of followers.[4] Figure 2 plots the number of retweets and the number of likes for each of Maduro's tweets. A tweet is represented in the scatter plot by two points, circles, which measure retweets, and crosses, which measure likes. The x-axis represents the date in which the tweet was published. Two features of the graph are worth noting. First, the number of retweets and the number of likes rapidly diverged. Second, a sharp discontinuity in the number of retweets is observed at the end of October.

[Figure 1 here] [Figure 2 here]

On October 31st, 2013, thousands of Twitter accounts were canceled by the company, without prior announcement or subsequent comment, leading to a sharp reduction in the average number of retweets that Maduro received. The exact number of accounts affected is not known, but Venezuela's Minister of Communication, Delcy Rodriguez, argued that almost 6,600 accounts were closed.[5] The government and the opposition both highlighted the event. Maduro called the event "a massive attack from the international right".[6] Other pro-Maduro and government leaders reinforced this line, while the opposition accused the government of focusing on the wrong things and declared that these were fake accounts (Figure A1 shows a small sample of tweets referring to the event).

The closed accounts represented a very small fraction of the overall number of Maduro's followers (less than 0.5 percent), but they represented a very large share of Maduro's retweets (around 81 percent).[7] The evidence suggests these accounts had been programmed to automatically retweet the president, violating Twitter's rules, and that a large share of these were

---

[4] Table A1 shows the number of followers of @NicolasMaduro in 2013.

[5] See *Maduro denuncia "ataque masivo" de Twitter Inc*. Retrieved from https://www.youtube.com/watch?v=ioTJQzTuPaU

[6] See https://www.semana.com/mundo/articulo/maduro-denuncia-ataque-de-twitter/363243-3

[7] The average number of retweets for @NicolasMaduro was 8,521 in October 2013, and 1,601 in November 2013. A regression of retweets on a post dummy indicator using Maduro's tweets confirms these estimates.

fake accounts.[8] The behaviour is also consistent with what is known as *astroturf* political campaigning: "politically-motivated individuals and organizations that use multiple centrally-controlled accounts to create the appearance of widespread support for a candidate or opinion" (Ratkiewicz et al, 2011).

Conceptual framework

I frame the study through the lens of the spiral of silence theory of public opinion. Noelle-Neumann (1974) proposed that individuals who perceive their views to be in the minority are generally less likely to express them publicly. I examine whether this observation extends to political dissent on social media. For this case study of Venezuela, I can derive specific predictions by taking advantage of the cancellation of the retweeting accounts, highlighting how this event changed individuals' "picture of the distribution of opinion in their social environment and of the trend of opinion" (Noelle-Neumann, 1974, p. 45).

The sharp drop in retweets for Maduro was a signal that the president did not have as much support as portrayed by his Twitter account. If one were to assess what the *dominating view* was among Venezuelan Twitter users by this metric, and comparing retweet counts with the leader of the opposition Henrique Capriles, the picture shifted from clear greater support for Maduro (Maduro retweet average pre-event: 8,607, Capriles retweet average pre-event: 835) to a much less clear picture (Maduro retweet average post-event: 1,501, Capriles retweet average post-event: 1,456).[9] The event was widely discussed by government and opposition leaders, news organizations,[10] as well as on Twitter by users of the platform, all of these contributing to users

[8] Though I can not state with certainty that all of the retweeting accounts were fake (ie. created in bulk and not belonging to a single individual), the behaviour of the accounts was such that they fit the definition of bots, as they were programmed to retweet Maduro through an automated script. See also https://help.twitter.com/en/rules-and-policies/twitter-rules; https://hipertextual.com/2017/06/por-que-twitter-pudo-haber-bloqueado-cuentas-en-venezuela; and https://panampost.com/panam-staff/2015/08/05/twitter-bots-make-maduro-worlds-third-most-retweeted-celebrity/.

[9] The averages reported are for the 30-day window around October 31st. Figure A3 shows that the pattern of increased retweets persisted for Capriles during the window of analysis.

[10] In the supplementary information file, I discuss the role of traditional media.

updating their perceptions of popular opinion. This belief updating would decrease the *fear of isolation* of users who held views previously thought of as being "minority".

The negative change in the perceived popularity of Maduro, and in particular the shift in what the dominating view was, would then lead to an increase in user's willingness to express these minority views. Therefore, I expect that *immediately* after the reduction in retweet activity in favour of Maduro, both criticism of the president (**H1a**) and support for the opposition (**H1b**) will shift upwards.[11]

As users became more willing to engage politically on the platform, the dynamics are likely to have reinforced themselves by both a change in individuals' own behaviour (Cho et al, 2016), as well as by observed changes in the behaviour of others. The dynamics, therefore, would lead to an unraveling of an existing spiral of silence and a shift in trends. This "unraveling" is similar to an informational cascade as formalized in Lohmann (1994). These informational cascades are of particular relevance for revolutions and regime transitions (Ellis and Fender, 2011), and communication technologies can play an important role in these dynamics (Little, 2016); which in turn explains why autocratic regimes engage in *astroturf* political campaigning and the manipulation of information (Guriev and Treisman, 2018).

Whereas the first hypotheses refer to a static shift in political expression, the second set of hypotheses proposes a dynamic shift, that is, a shift in the rate of change of political expression (or what Noelle-Neumann considers *the trend of opinion*). In particular, I expect that *over time* after the reduction in retweet activity in favour of Maduro, both criticism of the president (**H2a**) and support for the opposition (**H2b**) will increase.

Finally, I hypothesize that Twitter followers of users who frequently mention Maduro will be differentially aware of the event and thus be more likely to update their beliefs about public support for the president. Therefore, these users will experience a differential change in their public support, as measured by their follower engagement. In particular, support for opposition users who more frequently mention Maduro will increase differentially (**H3**).

---

[11] Though these two factors are closely linked in the context of Venezuela, empirically they are tested separately.

The data and empirical strategies are outlined below. I first use a graphical analysis to assess both H1 and H2. In particular, I study patterns in online engagement around the time of the event. I then formally evaluate H1 using both a pre/post research analysis and a difference-in-differences research design which restricts the sample to a short window of time around the event. Finally, in the supplementary information file, I evaluate H3 using a difference-in-differences framework which allows for heterogeneous effects depending on leaders' propensity to mention Maduro.

Data and graphical analysis

I collect three distinct datasets of tweets using Twitter's *Advanced Search* tools and the Twitter API, using both keywords and usernames. One important limitation of the collection methodology is that only the last 20 tweets per keyword/user per day can be obtained. This limitation binds (mostly) for the dataset of tweets mentioning @NicolasMaduro, which is collected using keywords. Since most users do not tweet more than twenty times in one day, the users (or leaders) dataset is substantially less affected by this limit. This limitation is discussed in detail below. Additional details on the data collection methodology are presented in the supplementary information file.

Two variables are used to measure follower engagement as outcomes: number of likes and number of replies.[12] Number of likes is an unambiguously positive signal of support. Number of replies is a measure of engagement that could be negative, as criticism, opposing views, and "trolling" are often found in tweet replies (Theocharis et al, 2016). The measures are log-transformed using a log(x+1) function because of both a large number of zeroes (many tweets with no engagement) and a long tail (tweets with thousands of likes).[13] The text of all tweets is converted to lowercase for analysis. In addition, I measure tweet volume using the time of publication of a subset of the tweets in the database (discussed at length below).

[12] Since retweets are more likely to be manipulated through the use of automated scripts, I do not use them as an outcome. The main results, however, are very similar for likes and for retweets. Evidence that the Venezuelan opposition also engaged in the use of Twitter bots is presented in Forelle et al (2015).

[13] In the supplementary information file I present results for a proxy measure of negative engagement constructed using the ratio of replies and likes.

This section describes each of the datasets collected and presents a graphical analysis of the quasi-experiment. The plots shown include scatter points, each representing either tweets, or average-daily values for a group of tweets, as well as kernel-weighted local polynomials that aim to fit the data from the tweets. I plot a six-month window of time around the date of the event to evaluate the relative magnitude of the changes in engagement following the event.

Maduro's tweets

The first dataset contains @NicolasMaduro's tweets starting from the date of his account opening until January of 2014. The total number of tweets is 1,820. During this period, Maduro tweeted at least 20 times per day on 5 occasions (thus there are potentially a few tweets missing), but none of these days occurred in October or November. Figure 3 shows three measures of engagement, retweets, likes, and replies, during a six-month window around the date of the event. The left panel highlights again the cancelation of the retweeting accounts. The event led to a sharp drop in retweets for the president. The middle panel shows that around the date of the event the number of likes Maduro receives is on an upward trend, there is a very small drop just at the time of the event, and the upward trend continues afterward. The right panel shows an increase in the number of replies in the days just after the cancelation of the accounts. In particular, it shows a slope change in the trend, suggesting that the upward trend in the number of replies accelerated, consistent with hypothesis H2a. Overall, users become more willing to reply to Maduro's tweets in the days following the cancelation of the retweeting accounts.

[Figure 3 here]

Tweets that mention @NicolasMaduro

The second dataset contains tweets which mention president Maduro's Twitter username, @NicolasMaduro, in the six-month window around the event. These are significant because they address or refer to the president, since he is "tagged", they are potentially visible to him and others who follow him. The overall volume of tweets mentioning @NicolasMaduro is very large and the collection methodology allows me to only capture a fraction of these. In particular, the last 20 tweets for each day. I expand the dataset by pairing the search with keywords. In particular, I first searched for tweets containing both @NicolasMaduro and specific keywords: two neutral keywords, *venezuela* and *pueblo* (the people), as well as *twitter*, which is relevant for the case study. After the initial data collection, the tweets were examined and ten more keywords were selected which indicate a bias either for or against (and even offensive towards) Maduro. As proxy keywords for support I use *revolucion* (revolution), *comandante* (commander), *camarada* (comrade), *chavez*, *victoria* (victory); and as proxy keywords for opposition I use *regimen* (regime), *ilegitimo* (illegitimate), *escasez* (scarcity), *maldito* (damned/cursed), and *ladron* (thief).[14] The final dataset consists of 42,631 tweets, including the last 20 tweets per day that contain both @NicolasMaduro and each of the thirteen selected keywords.

I am interested in measuring the volume of tweets but because the sample of tweets is a selected subsample resulting from the collection methodology, there are many keyword-day pairs with exactly 20 tweets. Since these are the *last* 20 tweets of the day, as a proxy measure for the overall volume of tweets I can use the exact times at which these were published. The intuition behind the idea can be illustrated with a simple example. Suppose that tweets are published uniformly throughout the day. On a day in which 20 (or less) tweets are published, all of them are observed, such that on average, the publication time of tweets in the sample would be 12:00. On a day in which 40 tweets are published matching the search criteria, since only the latest 20 are observed, the publication time would be, on average, 18:00. Higher tweet volume implies later publication times for the last 20 tweets of the day. Note that for this to work the distribution does not need to be uniform, I only need to assume that the distribution of times at which people tweet does not change after the event. There are no particular reasons for why the accounts' cancelations would affect the time at which people tweet.[15] I use the natural log of the seconds to the end of the day as the proxy measure for tweet volume.

Using the third dataset, tweets by political leaders (discussed below), for which this sampling constraint does not bind, I can measure the strength of the correlation between the volume of tweets and the publication times of the last 20 tweets. In particular, I look for all tweets by the selected users (including government, opposition, and sports-related accounts) which

[14] The choice of keywords was analyzed by manually coding a random selection of 300 tweets. The results of this process and the potential biases that may arise due to measurement error are discussed in the supplementary information file.
[15] And indeed, in the user-based sample for political leaders (discussed below) for which the 20-tweets-per-day limit does not generally bind, there are no changes in tweeting behaviour with respect to time of tweets after the event.

mention a particular keyword. Since this dataset is collected at the username level (not the keyword level), there can be more than 20 tweets in a day which contain these keywords. I look at three keywords (*maduro*, *venezuela*, and *pueblo*), and find that indeed this measure is strongly correlated with the number of tweets published (shown in Figure A4). The relationship between number of tweets and publication times of the last 20 tweets is strongly statistically significant ($p<0.001$). In the results section, I use this analysis to quantify the changes in tweets volume.

Having established that publication times are a good proxy for tweet volume, Figure 4 shows the publication time (measured in time to the end of the day) for tweets containing the @NicolasMaduro username, as well as for tweets containing both @NicolasMaduro and the *twitter* keyword. The first graph shows that the publication time of tweets increases (seconds to the end of the day decreases), revealing a higher number of tweets mentioning @NicolasMaduro after the cancelation of the retweeting accounts. A sharper drop (in time to the end of the day) is observed for tweets mentioning @NicolasMaduro which also contain the keyword *twitter* (right panel), perhaps not surprising given the event under study, which reveals the event was widely discussed by users of the platform while addressing Maduro. In the following section, I evaluate the statistical significance of the patterns outlined here as well as the extent to which the increase in tweet volume is different for anti-Maduro keywords, relative to pro-Maduro keywords.

[Figure 4 here]

Tweets from Political Leaders + Control Group

The third dataset contains tweets by Venezuelan political leaders. I first collected account usernames from http://www.twven.com/, a website which archives and documents Twitter users in Venezuela. The top accounts were selected, ranked by number of followers, using two categories: politics and government. With this resource, I compiled and selected a resulting list of users containing 50 government-affiliated users and 50 opposition users. I remove media, government agencies, and other news accounts when selecting the list of users. An additional 50 accounts were compiled from the sports category, which is used as a control group in some of the empirical exercises below.[16]

I then collect tweets from these users during the six-month window around the date of the event. The final dataset contains 184,855 tweets.[17] Figure 5 shows the daily average number of likes for tweets from the opposition (circles), from the government (squares), and the control-group accounts (triangles), during the period of study. The vertical line indicates the day after the accounts' cancelations, October 31st, 2013. After the event, there is a differential increase in tweet likes for opposition accounts, relative to both the government and the control-group accounts. There is both a discontinuity in the number of likes, as seen in the level change on the date of the event, as well as a trend break, observed as a change in the slope of the kernel-weighted polynomial (consistent with hypothesis H2b). The graph also reveals no differential trends before the event for any of the subgroups. The empirical exercises below aim to precisely measure the gap in relative engagement and to determine its' statistical significance.

[Figure 5 here]

Research design and empirical framework

To analyze the impact of the closure of the retweeting accounts on the outcomes of interest, I present evidence from two types of empirical exercises. The first exercise is a pre/post analysis which captures the change in the outcomes for the days after the event, relative to the days before. The second type of analysis is a difference-in-differences, in which I estimate the relative change in outcomes across groups.

The research design exploits the cancelation of the accounts, which was arguably unanticipated by social media users and political elites. The figures shown above reveal no

[16] The complete list of accounts selected is listed in Table A2. Users initially selected who did not tweet during the period of study are excluded from the analysis.

[17] The 20 tweets-per-day constraint binds for about 23% of user-day pairs (17% for government leaders, 22% for opposition leaders, and 35% for the control-group accounts). I do not anticipate any particular bias to arise due to excluded tweets, but this is discussed in more detail below.

changes in political expression in the days before the event. Given the high-frequency of the data, the event's impact can be assessed using a narrow window of time around the date it took place.[18] There exists an important trade-off in choosing the time window of analysis. Longer time windows allow for the analysis of longer-run relationships, which is important considering the potential dynamic aspect of the effect, by which a change in online political expression propagates and intensifies over days as users observe and learn from the behaviour of others (as highlighted in the theory and evidenced in the graphical analysis above). In addition, statistical precision increases by including more data. On the other hand, the longer the window, the higher the probability of other unobserved shocks affecting the outcomes, cautioning against the interpretation of these estimates as a *causal* effect. The chosen specification looks at a 30-day window around the event date, comparing outcomes 15 days before the event relative to outcomes 15 days after. This 30-day window is long enough to allow for the dynamic component of the effect to unravel and has sufficient statistical power, but also excludes two potentially confounding events which occurred in the weeks after the cancelation event: the National Assembly's decision to grant Maduro emergency powers to rule by decree on November 19th,[19] and the Venezuelan municipal elections on December 8th.

Pre/post analysis

To study whether patterns of engagement change in the days just before the event, relative to the days just after, I restrict the sample to the 30-day window around the event, and estimate regression equations of the following form:

$$Y_{iut} = \beta \cdot Post_t + X_{iut} \cdot \theta + \alpha_u + \varepsilon_{iut} \quad (1)$$

where Y_{iut} is the outcome of interest, for tweet i by user u at time t. $Post_t$ is a dummy variable equal to 1 if the tweet occurred after the accounts were canceled; and X_{iut} is a vector of controls depending on the dataset. These include dummy variables for whether the tweet starts with the

[18] Similar methodological approaches exploiting high-frequency data and short-horizons are common in financial studies (Kothari and Warner, 2007) and environmental impact analyses.
[19] See https://www.reuters.com/article/us-venezuela-maduro-powers/venezuelas-congress-approves-decree-powers-for-maduro-idUSBRE9AI16L20131119

"rt" keyword (indicating a retweet), contains a hashtag ("#"), an @ symbol (indicating a mention), or a url (with the keyword "http"), a cubic polynomial for the time of day at which the tweet was published, and day of the week fixed effects.[20] The user-level analyses include also user fixed effects α_u (to capture time-invariant characteristics of the users, such as overall popularity) and the monthly number of followers as a control.[21] The coefficient of interest, β, captures the change in engagement in the days after the event, relative to the days before the event.

Difference-in-differences analysis

To study whether there are differential changes in engagement across subgroups in the data, I estimate difference-in-differences regressions of the following form:

$$Y_{iut} = \beta_1 \cdot Post_t + \beta_2 \cdot Opposition_{iu} + \beta_3 \cdot Post_t \cdot Opposition_{iu} + X_{iut} \cdot \beta + \alpha_u + \varepsilon_{iut} \qquad (2)$$

on selected subsamples of interest. In particular, I evaluate whether anti-Maduro tweets received differential increased in volume (as measured by time of publication), and whether opposition leaders received differential increase in engagement outcomes (likes and replies). In some specifications of the difference-in-differences analysis, I include day fixed effects α_t (and remove the collinear $Post_t$ dummy) and include user/keyword fixed effects α_u. The coefficient of interest, β_3, captures the differential change in the outcome in the days after the event for a treated group (in most specifications coded as the opposition), relative to a control group (either the neutral keywords/sports accounts, or the government, depending on the specification).

The difference-in-differences specification generally considers one unaffected group (in the baseline specifications, the sports-related accounts), as a control group, to assess how the outcome would have evolved for the treated group in the absence of treatment. The identification assumption is that, in the absence of treatment, the groups would have followed "parallel trends".

[20] Some of these controls are potentially "bad controls", since users may have reacted to the event by, for instance, publishing tweets with a #hashtag, and this may be arguably part of the "effect" of the event. However, they also control for the fact that tweets with these features have different engagement patterns, potentially leading to spurious relationships when evaluating the effect (replies, for instance, tend to generate less engagement on average), due to sampling error. I present results from models both with and without controls.

[21] Shown in Figure A5. An overall increase in the number of followers in November of 2013 is observed for both opposition and government accounts, suggesting that the differential increase in engagement is not driven by new followers. This data is collected from twven.com, as historical follower statistics are not available from Twitter. Because historical data is not available for some accounts, I impute missing values based on predictions from a simple regression model of followers on the measures of engagement (likes, retweets and replies), using the accounts for which the data is available.

Figure 5 shows that, before the event, all of the groups seemed to be on roughly parallel paths. Under the same assumption, I use this analysis to test whether, in the case of two "treated" groups (government and opposition), one of the groups was affected differentially. Additional details on the empirical framework and a test for differential pre-trends are presented in the supplementary information file.

Results

@NicolasMaduro's tweets

Table 1 examines engagement (measured in retweets, likes, and replies) for @NicolasMaduro. I discuss results from the more conservative specification, which includes a set of tweet-level controls (shown in even-numbered columns). The first two columns reveal the magnitude of the event in question, and in particular reveal a sharp drop in the log number of retweets (β=-1.837, p<0.001). Despite the sharp drop in retweets, the number of likes is relatively unaffected (β=0.106, p=0.28). On the other hand, the number of replies increases (β=0.269, p=0.022), suggesting that users are more likely to engage with tweets by the president in this form. Considering that replies can represent negative engagement, this result is partially consistent with hypothesis H1a (increased criticism of the president). In the following subsection, I look at tweets with specific keywords to further assess this hypothesis.

[Table 1 here]

Tweets mentioning @NicolasMaduro

The analysis of tweet volume, which uses the publication times of tweets as a proxy outcome, is presented in Tables 2 and 3. Table 2 shows the results from the pre/post analysis. I present evidence for the unrestricted sample, as well as for the *twitter*, *venezuela*, and *pueblo* keywords. Observations for keywords identified as being anti-Maduro and pro-Maduro are pooled into groups. The results suggest a significant increase in the volume of tweets mentioning the president in the days after the event (as revealed by a decrease in the log number of seconds until the end of the day, for the selected subsample of the last twenty tweets in the day). The coefficients for the unrestricted sample (β=-0.651, p=0.002) and for the neutral *venezuela* (β=-0.558, p=0.018) and *pueblo* (β=-0.805, p=0.021) keywords are relatively similar. For tweets that mention both @NicolasMaduro and the *twitter* keyword, the suggested increase in volume is much larger (β=-1.485, p<0.001), which is not surprising given the event under study, but reveals that indeed the cancelation of the accounts was widely discussed on the platform. Finally, the coefficients for anti-Maduro keywords (β=-0.845, p<0.001) and pro-Maduro keywords (β=-0.423, p=0.001) both suggest increased activity, but a larger increase for anti-Maduro tweets.

[Table 2 here]

Table 3 presents results from the difference-in-differences regressions with publication times as the outcome of interest. The preferred specification, which includes day fixed-effects and keywords fixed-effects, reveals earlier publication times for pro-Maduro keywords relative to neutral keywords (column 2, β=0.326, p=0.039), suggesting a relative decrease in volume. There is also a significant difference between publication times of tweets with anti-Maduro keywords relative to those with pro-Maduro keywords (β=-0.423, p=0.002), suggesting that the increase in volume of tweets mentioning @NicolasMaduro was significantly larger for those critical of the president relative to those supportive of him.

[Table 3 here]

To assess the magnitude of these coefficients, I estimated the relationship between the proxy measure (log number of seconds to the end of the day) and tweet volume (log number of tweets), using the sample of tweets for political leaders (as illustrated in Figure A4). The estimated coefficient suggests that a -1 unit change in the log of seconds to the end of the day is associated with an increase in tweet volume of 48 percent (β=-0.485, p<0.001; not shown in tables).[22] A simple calculation then suggests that in the days after the event, the overall number of tweets including the @NicolasMaduro username increased by an estimated 32 percent (-0.651×-0.485); tweets including @NicolasMaduro and the keyword *twitter* increased by an estimated 72 percent (-1.485×-0.485); and that anti-Maduro tweets increased by a differential 21 percent, relative to pro-Maduro

[22] Using the sample of tweets by Venezuelan leaders, I pool the keywords selected (*venezuela*, *pueblo*, *maduro*), and regress the log daily number of tweets on the log of seconds to the end of the day of the last 20 tweets, for the six-month long sample (N=518).

tweets (-0.423×-0.485). Consistent with hypothesis H1a, users were more willing to express criticism of the president following the event. In the supplementary appendix, I repeat this analysis on an alternative dataset of tweets including the "Maduro" keyword.

Tweets by Political Leaders

This section evaluates whether engagement for political leaders changed after the cancellation of the accounts. Table 4 reports the results from the pre/post analysis. The estimated coefficients (including controls and fixed-effects) for the engagement outcomes, likes for the government (β=0.132, p<0.001), replies for the government (β=0.142, p<0.001), likes for the opposition (β=0.372, p<0.001), and replies for the opposition (β=0.242, p<0.001), all reveal increased engagement.

[Table 4 here]

To further investigate these patterns, I use a set of sports-related accounts which should have been unaffected by the event, but capture broader trends in Twitter adoption and usage in Venezuela, as a control group, in a difference-in-differences analysis. In addition, I evaluate the relative difference between opposition and government tweets with the same approach. These results are shown in Table 5. The analysis reveals that the increase in likes for the government is not statistically significant relative to the control group (β=0.064, p=0.118), but it is for the number of replies (β=0.170, p=0.001). For the opposition, both number of likes (β=0.282, p<0.001), and number of replies (β=0.218, p<0.001), are significantly greater than for the control group. Finally, the relative changes between opposition and government show 21 percent increased likes for the opposition (β=0.214, p<0.001) and an insignificant difference for replies (β=0.058, p=0.242). These results support hypothesis H1b, increased willingness to express support for the opposition. In the supplementary information file, I use the ratio of replies to likes as a proxy outcome for negative engagement and show that this increased differentially more for government accounts.

[Table 5 here]

Discussion

I have presented evidence from a previously undocumented event to study the effects of social media bots on political dynamics. Though much has been speculated about these potential effects, few studies exist which aim to establish causal relationships using large scale data on political behaviour.

The analysis revealed that in the days after the cancelation of thousands of Twitter accounts, which actively retweeted Nicolas Maduro's tweets, both criticism of the president and support for the opposition substantially increased.[23] As the picture of the dominating view of Venezuelan Twitter users sharply changed, users who held minority views became more willing to express their political dissent. It should also be noted that the shift in perceived popularity did not affect government supporters (support for Maduro and other government leaders did not significantly change), presumably because these users were in the previously perceived majority. After the accounts were canceled there was no clear dominating view, and the fear of isolation would have been weaker for users supporting the government to begin with.

The graphical analyses revealed a shift in the static picture of support, at the time of the event, and a dynamic change in the "trend of opinion", as apparent by the slope changes in the estimated outcomes. These findings are consistent with an unraveling of an existing spiral of silence in online political expression. The initial shift in public opinion led to a cascade over the following days, likely through users both observing the behavioral shift in others' increased willingness to express their views, as well as own previous expressions reinforcing their behaviour. The graphical analysis also reveals that some of these changes may have persisted even months after the event, in particular individuals' willingness to support the opposition (Figure 5). Though a long-run *causal* relationship is not warranted by the empirical design, I present a formal analysis looking at the six-month window around the event and discuss the results (and cautious interpretations) in the supplementary information file.

[23] The two separate analyses (using different datasets and outcomes) suggest a relative increase in opposition support, and anti-Maduro criticism, of 21 percent.

One important limitation of the study stems from the data collection restrictions faced for the keyword-based tweets, as well as the potential measurement error that arises from using a selected set of keywords. The results showing that the volume of anti-Maduro tweets increased relative to that of pro-Maduro tweets should be interpreted with these caveats in mind. However, the proposed solutions and robustness tests, and in particular the use of a proxy for tweet volume using tweets' publication times, could be used by future studies facing similar data restrictions. A second limitation is the inability to disentangle the precise mechanism through which the effects arise. Though the results support the proposed framework (and other studies have documented the presence of spirals of silence in online settings), an alternative explanation is that it was the bad press Maduro received due to the event which led to the shift in users' behaviour. A fully experimental setting is one possible avenue to disentangle these mechanisms in future work. Finally, I do not investigate the extent to which the observed online political behaviour mapped into offline political behaviour.

Political leaders continuously strive for popularity and broad public support. As autocrats' hold on power largely depends on a shared perception of weak opposition to their leadership, it is especially useful for them to amplify their perceived popularity. The thousands of accounts which actively retweeted Maduro's messages, increasing his retweet statistics by over 81 percent, accomplished precisely this. Autocratic environments, in which both public perceptions are distorted and the fear of isolation is particularly acute, present conditions well-suited for a spiral of silence in online political expression to arise. On the other hand, such dynamics would be less likely to develop in democratic environments with a strong culture of free speech. Future research should examine under which conditions spirals of silence in online political expression emerge, and aim to understand the circumstances under which they unravel.

References

Allcott, Hunt, and Matthew Gentzkow (2017). Social Media and Fake News in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211-36.

Allcott, Hunt, Matthew Gentzkow, and Chuan Yu (2019). Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, *6*(2).

Barberá, Pablo (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*, *23*(1), 76-91.

Barberá, Pablo (2016). Less is More? How Demographic Sample Weights Can Improve Public Opinion Estimates Based on Twitter Data. *Working Paper.*

Barberá, Pablo, and Thomas Zeitzoff (2018). The new public address system: Why do world leaders adopt social media? *International Studies Quarterly*, *62*(1), 121-130.

Boulianne, Shelley (2015). Social media use and participation: A meta-analysis of current research. *Information, Communication & Society*, 18, 524-538.

Ceron, Andrea, Luigi Curini, Stefano M. Iacus, and Giuseppe Porro (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society*, *16*(2), 340-358.

Chadwick, Andrew (2017). *The hybrid media system: Politics and power*. Oxford University Press.

Cho, Jaeho, Saifuddin Ahmed, Heejo Keum, Yun Jung Choi, and Jong Hyuk Lee (2018). Influencing myself: Self-reinforcement through online political expression. *Communication Research*, 45(1), 83-111.

Colleoni, Elanor, Alessandro Rozza, and Adam Arvidsson (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication*, *64*(2), 317-332.

DiGrazia, Joseph, Karissa McKelvey, Johan Bollen, and Fabio Rojas (2013). More Tweets, More Votes: Social Media as a Quantitative Indicator of Political Behavior. *PLoS ONE* 8(11): e79449.

Ellis, Christopher J., and John Fender (2011). Information cascades and revolutionary regime transitions. *The Economic Journal*, *121*(553), 763-792.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova (2017). Social media and protest participation: Evidence from Russia. *Available at SSRN 2696236*.

Ferrara, Emilio, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96-104.

Ferraz, Claudio, and Frederico Finan (2008). Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, *123*(2), 703-745.

Field, Anjalie, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov (2018). Framing and Agenda-setting in Russian News: a Computational Analysis of Intricate Political Strategies. *arXiv preprint arXiv:1808.09386*.

Forelle, Michelle, Phil Howard, Andrés Monroy-Hernández, and Saiph Savage (2015). Political bots and the manipulation of public opinion in Venezuela. *arXiv preprint arXiv:1507.07109*.

Gearhart, Sherice, and Weiwu Zhang (2014). Gay bullying and online opinion expression: Testing spiral of silence in the social media environment. *Social science computer review*, 32(1), 18-36.

Gil de Zúñiga, Homero, Logan Molyneux, and Pei Zheng (2014). Social media, political expression, and political participation: Panel analysis of lagged and concurrent relationships. *Journal of Communication*, 64(4), 612-634.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera (2018). Social media, sentiment and public opinions: Evidence from #Brexit and #USElection. *NBER Working Paper 24631.*

Guriev, Sergei M., and Daniel Treisman (2018). Informational autocrats. *Available at SSRN 3208523*.

Halberstam, Yosh, and Brian Knight (2016). Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. *Journal of Public Economics*, *143*, 73-88.

Jost, John T., Pablo Barberá, Richard Bonneau, Melanie Langer, Megan Metzger, Jonathan Nagler, Joanna Sterling, and Joshua A. Tucker (2018). How Social Media Facilitates Political Protest: Information, Motivation, and Social Networks. *Political Psychology*, *39*, 85-118.

Keller, Franziska B., David Schoch, Sebastian Stier, and JungHwan Yang (2017). How to manipulate social media: Analyzing political astroturfing using ground truth data from South Korea. *Proceedings of the Eleventh International AAAI Conference on Web and Social Media.*

King, Gary, Jennifer Pan, and Margaret E. Roberts (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, 107(2), 326-343.

King, Gary, Jennifer Pan, and Margaret E. Roberts (2017). How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *American Political Science Review*, 111(3), 484-501.

Kothari, Sagar P., and Jerold B. Warner (2007). Econometrics of event studies. *Handbook of empirical corporate finance*, 1, 3-36.

Lazer, David M.J., Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, Jonathan L. Zittrain (2018). The Science of Fake News. *Science,* 359(6380): 1094-1096.

Lohmann, Susanne (1994). The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-91. *World Politics*, 42-101.

Little, Andrew T (2016). Communication technology and protest. *The Journal of Politics*, *78*(1), 152-166.

Liu, Yu, Jian Raymond Rui, and Xi Cui (2017). Are people willing to share their political opinions on Facebook? Exploring roles of self-presentational concern in spiral of silence. *Computers in Human Behavior*, *76*, 294-302.

Margetts, Helen, Peter John, Scott Hale, and Taha Yasseri (2016). *Political Turbulence: How Social Media Shape Collective Action.* Princeton, New Jersey: Princeton University Press.

Matthes, Jörg, Johannes Knoll, and Christian von Sikorski (2018). The "spiral of silence" revisited: a meta-analysis on the relationship between perceptions of opinion support and political opinion expression. *Communication Research*, 45(1), 3-33.

Munger, Kevin, Richard Bonneau, Jonathan Nagler, and Joshua A. Tucker (2018). Elites tweet to get feet off the streets: Measuring regime social media strategies during protest. *Political Science Research and Methods*, 1-20. doi:10.1017/psrm.2018.3.

Noelle-Neumann, Elisabeth (1974). The spiral of silence: a theory of public opinion. *Journal of Communication*, *24*(2), 43-51.

O'Connor, Brendan, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith (2010). From tweets to polls: Linking text sentiment to public opinion time series. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, 122–129.

Qin, Bei, David Strömberg, and Yanhui Wu (2017). Why does China allow freer social media? Protests versus surveillance and propaganda. *Journal of Economic Perspectives*, *31*(1), 117-40.

Ratkiewicz, Jacob, Michael D. Conover, Mark Meiss, Bruno Gonçalves, Alessandro Flammini, and Filippo Menczer (2011, July). Detecting and tracking political abuse in social media. *Fifth international AAAI conference on weblogs and social media*.

Skoric, Marko M., Qinfeng Zhu, Debbie Goh, and Natalie Pang (2016). Social media and citizen engagement: A meta-analytic review. *New Media & Society*, 18, 1817-1839.

Soffer, Oren, and Galit Gordoni (2018). Opinion expression via user comments on news websites: analysis through the perspective of the spiral of silence. *Information, communication & society*, *21*(3), 388-403.

Suhay, Elizabeth, Emily Bello-Pardo, and Brianna Maurer (2018). The polarizing effects of online partisan criticism: evidence from two experiments. *The International Journal of Press/Politics*, *23*(1), 95-115.

Theocharis, Yannis, Pablo Barberá, Zoltán Fazekas, Sebastian Adrian Popa, and Olivier Parnet (2016). A bad workman blames his tweets: the consequences of citizens' uncivil Twitter use when interacting with party candidates. *Journal of Communication*, *66*(6), 1007-1031.

Tucker, Joshua A., Yannis Theocharis, Margaret E. Roberts, and Pablo Barberá (2017). From liberation to turmoil: social media and democracy. *Journal of Democracy*, 28(4), 46-59.

Tucker, Joshua A., Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *Hewlett Foundation Report.*

Tufekci, Zeynep, and Christopher Wilson (2012). Social media and the decision to participate in political protest: Observations from Tahrir Square. *Journal of Communication*, *62*(2), 363-379.

Vaccari, Cristian, Augusto Valeriani, Pablo Barberá, Rich Bonneau, John T. Jost, Jonathan Nagler, and Joshua A. Tucker (2015). Political expression and action on social media: Exploring the relationship between lower- and higher-threshold political activities among Twitter users in Italy. *Journal of Computer-Mediated Communication*, *20*(2), 221-239.

Waisbord, Silvio, and Adriana Amado (2017). Populist communication by digital means: presidential Twitter in Latin America. *Information, Communication & Society, 20(9),* 1330-1346.

Howard, Philipp N., and S. C. Woolley (2016). Political communication, computational propaganda, and autonomous agents. *International Journal of Communication, 10(2016).*

Wu, Tai-Yee, and David J. Atkin (2018). To comment or not to comment: Examining the influences of anonymity and social support on one's willingness to express in online news discussions. *New Media & Society*, 20(12), 4512-4532.

Zerback, Thomas, and Nayla Fawzi (2017). Can online exemplars trigger a spiral of silence? Examining the effects of exemplar opinions on perceptions of public opinion and speaking out. *New Media & Society*, *19*(7), 1034-1051.

Figures

Figure 1: Nicolas Maduro's first tweet

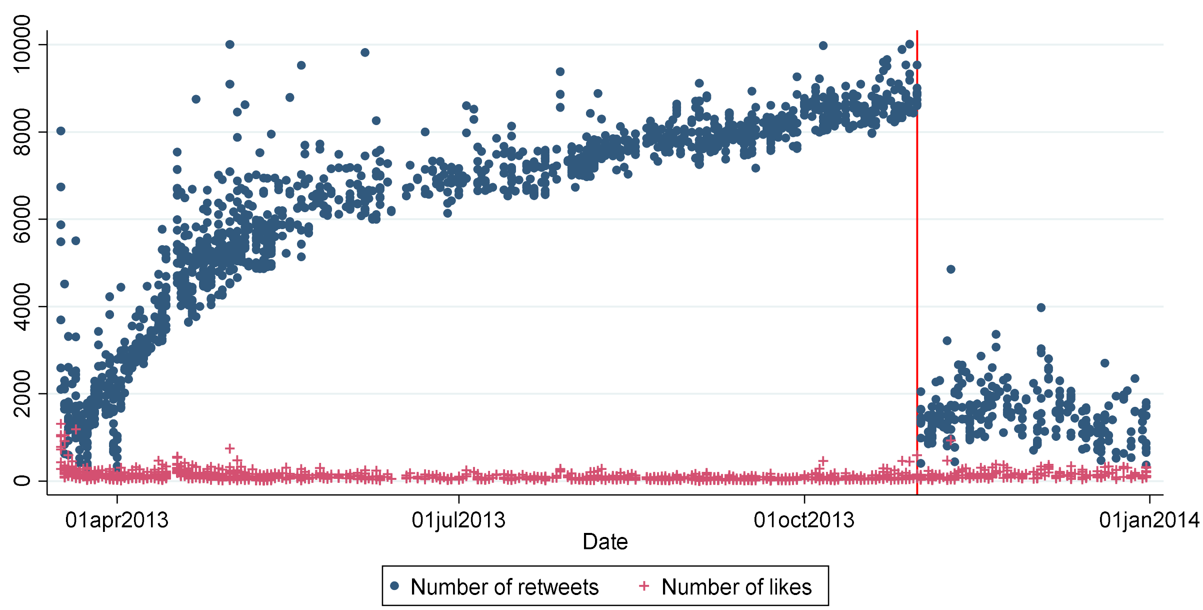

Figure 2: @NicolasMaduro retweets and likes

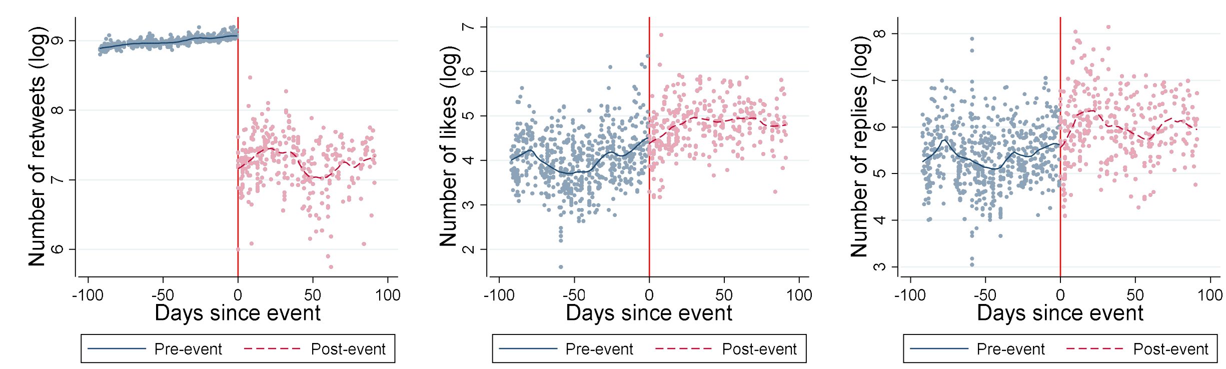

Figure 3: Engagement for @NicolasMaduro's tweets

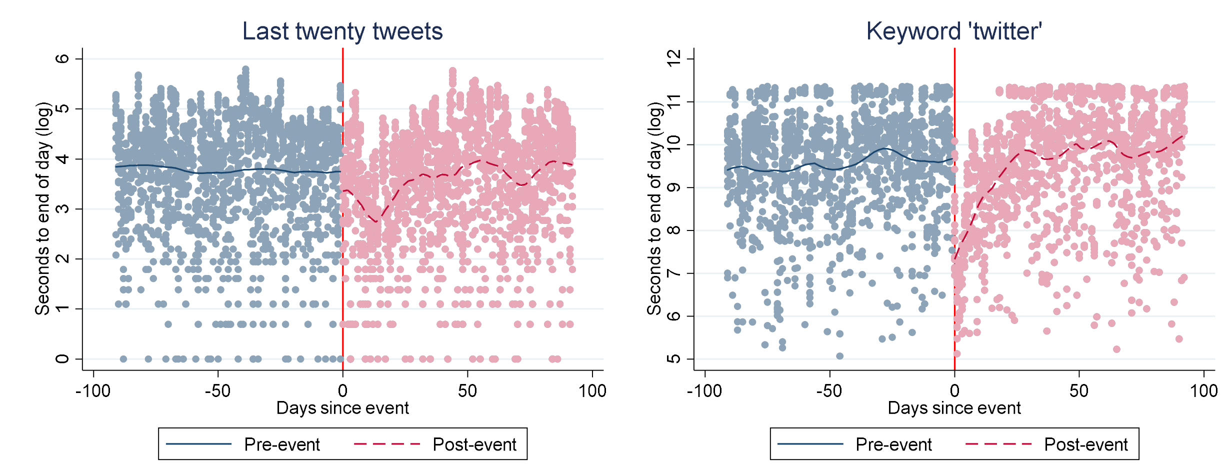

Figure 4: Time of tweets for last 20 tweets mentioning @NicolasMaduro, and last twenty 20 mentioning @NicolasMaduro and containing the 'twitter' keyword

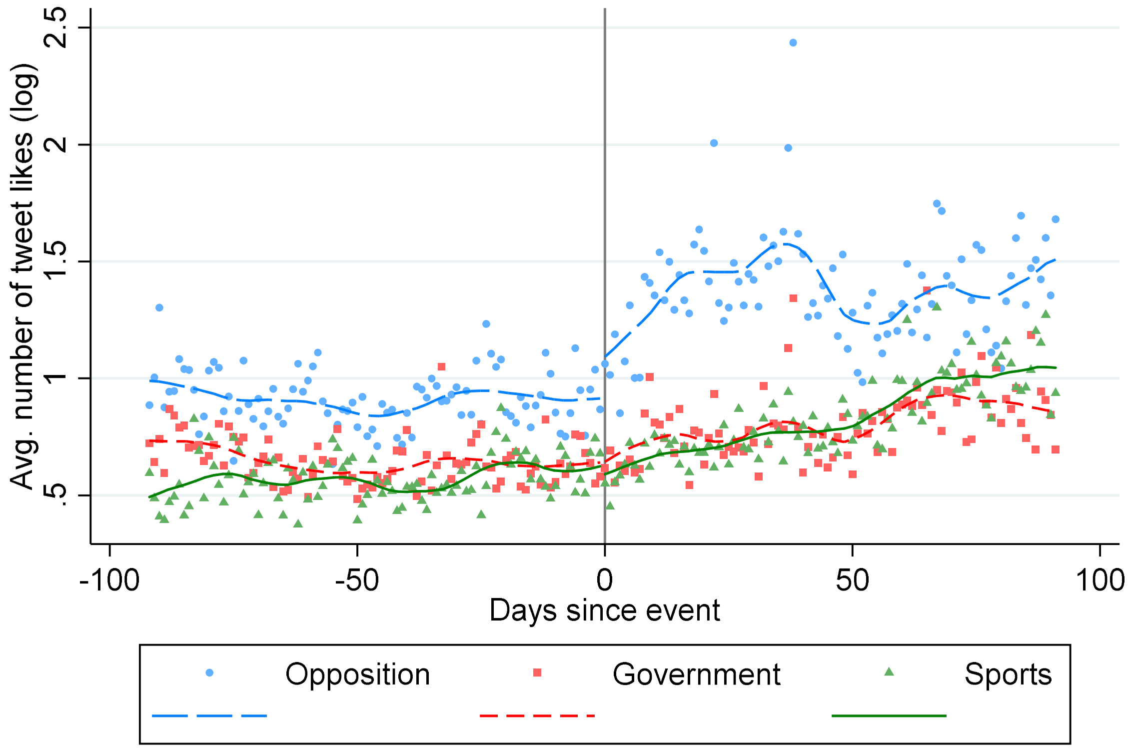

Figure 5: Average tweet likes for selected accounts

Tables

Table 1: Relationship between accounts' cancellation and engagement for @NicolasMaduro's tweets

Dependent variable:	Log retweets		Log likes		Log replies	
	(1)	(2)	(3)	(4)	(5)	(6)
Post	-1.811***	-1.837***	0.15	0.106	0.345***	0.269**
	(0.037)	(0.049)	(0.093)	(0.098)	(0.109)	(0.117)
N	183	183	183	183	183	183
Controls	No	Yes	No	Yes	No	Yes

Notes: Sample includes tweets by @NicolasMaduro in a 30-day window around October 31, 2013. Controls include dummy variables for whether the tweet starts with the "rt" keyword (indicating a retweet), contains a hashtag ("#"), an @ symbol (indicating a mention), or a url (with the keyword "http"), a cubic polynomial for the time of day at which the tweet was published, and day of the week fixed effects. Robust standard errors in parenthesis. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table 2: Relationship between accounts' cancellation and publication times for tweets mentioning @NicolasMaduro

Keyword sample:	None (last 20 overall)	*venezuela*	*pueblo*	*twitter*	Anti-Maduro (pooled)	Pro-Maduro (pooled)
	(1)	(2)	(3)	(4)	(5)	(6)
Post	-0.651***	-0.558**	-0.805**	-1.485***	-0.845***	-0.423***
	(0.194)	(0.223)	(0.330)	(0.314)	(0.126)	(0.122)
N	600	600	600	584	2,602	2,712
N-clusters	30	30	30	30	150	150
Keyword fixed-effects	No	No	No	No	Yes	Yes

Notes: Outcome measured is time of publication, measured in log seconds to the end of the day, as a proxy for tweet volume. Less time to the end of the day indicates greater tweet volume. All of the specifications include day of the week fixed effects. Standard errors clustered at the day level for columns 1-4, and clustered at the keyword-day level for columns 5-6. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table 3: Differential change in publication times for tweets mentioning @NicolasMaduro

Keyword sample:	Pro-Maduro vs. Neutral		Anti-Maduro vs. Neutral		Pro-Maduro vs. Anti-Maduro	
	(1)	(2)	(3)	(4)	(5)	(6)
Post	-0.728***		-0.685***		-0.424**	
	(0.206)		(0.209)		(0.171)	
Pro-Maduro	0.857***					
	(0.139)					
Post * Pro-Maduro	0.296	0.326**				
	(0.267)	(0.157)				
Anti-Maduro			3.187***		2.299***	
			(0.123)		(0.127)	
Post * Anti-Maduro			-0.127	-0.117	-0.391*	-0.423***
			(0.250)	(0.152)	(0.219)	(0.137)
N	3,697	3,697	3,738	3,738	5,288	5,288
N-clusters	210	210	210	210	300	300
Day of the week fixed-effect	Yes	No	Yes	No	Yes	No
Keyword fixed-effects	No	Yes	No	Yes	No	Yes
Day fixed-effects	No	Yes	No	Yes	No	Yes

Notes: Outcome measured is time of publication, measured in standardized time until the end of the day, as a proxy for tweet volume. Less time to the end of the day indicates greater tweet volume. Standard errors clustered at the keyword-day level. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table 4: Relationship between accounts' cancellation and engagement for political leaders

Sample:	Government				Opposition			
Dependent variable:	Log likes		Log replies		Log likes		Log replies	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	0.088**	0.132***	0.064	0.142***	0.330***	0.372***	0.177***	0.242***
	(0.041)	(0.030)	(0.063)	(0.035)	(0.052)	(0.056)	(0.056)	(0.048)
N	7,666	7,666	7,666	7,666	11,480	11,480	11,480	11,480
N-clusters	44	44	44	44	48	48	48	48
Controls	No	Yes	No	Yes	No	Yes	No	Yes
User fixed-effects	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Sample includes tweets by political leaders in a 30-day window around October 31, 2013. Controls include dummy variables for whether the tweet starts with the "rt" keyword (indicating a retweet), contains a hashtag ("#"), an @ symbol (indicating a mention), or a url (with the keyword "http"), a cubic polynomial for the time of day at which the tweet was published, day of the week fixed effects, and number of followers (monthly). Standard errors clustered at the user (leader) level in parenthesis. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table 5: Differential change in engagement for tweets by political leaders

Sample:	Government vs. Sports		Opposition vs. Sports		Opposition vs. Government	
Dependent variable (log):	likes	replies	likes	replies	likes	replies
	(1)	(2)	(3)	(4)	(5)	(6)
Post * Government	0.064	0.170***				
	(0.040)	(0.051)				
Post * Opposition			0.282***	0.218***	0.214***	0.058
			(0.049)	(0.052)	(0.050)	(0.049)
N	20,830	20,830	24,644	24,644	19,146	19,146
N-clusters	90	90	94	94	92	92
Controls (not shown)	Yes	Yes	Yes	Yes	Yes	Yes
Day fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
User fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Sample includes tweets by political leaders (government and opposition) and by popular sports accounts in a 30-day window around October 31, 2013. Standard errors clustered at the user level in parenthesis. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Supplementary Information: For Online Publication

This online appendix presents discussions about Twitter use in Venezuela, the role of traditional media during the event, an extended explanation of the data collection methodology, an extended discussion of the difference-in-differences research design, and a series of empirical extensions to the main analyses.

Twitter use in Venezuela

Twitter gained popularity in Venezuela after being adopted by Hugo Chavez in 2010 as "a tool for government", encouraging citizens to tweet concerns directly to him and employing 200 people to respond to citizens' messages.[24,25] In the early 2010s, Venezuela ranked among the top 20 countries in number of Twitter users[26] and ranked among the top 5 in terms of Twitter penetration[27] (with an estimate for 2012 of 21 percent). A Pew survey conducted in 2013 in Venezuela revealed that 73% of those aged 18-29, 52% of those aged 30-49, and 15% of those 50 or older, used social networking sites. Of these users, 49% reported using the sites to share views about politics, and 74% reported they had learned about others' political beliefs from something they posted on a social networking site (the highest rate among the 22 developing countries in the study).[28] Twitter has been used extensively by both government and opposition leaders, and has played an important role in the country's political developments.[29] Other sources regarding Twitter and social media use in Venezuela can be found in Forelle et al (2015) and Munger et al (2018).

Traditional media

Though the analysis is centered on Twitter, traditional media also played a role in diffusing information related to the cancellation of the retweeting accounts. As previously highlighted,

---

[24] https://www.theguardian.com/world/2010/aug/10/hugo-chavez-twitter-venezuela
[25] His approach to social media made Chavez the second most popular head of state on Twitter in 2012, only behind Barack Obama (see http://www.digitaldaya.com/admin/modulos/galeria/pdfs/69/156_biqz7730.pdf).
[26] https://semiocast.com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US
[27] https://www.comscore.com/Insights/Press-Releases/2011/4/The-Netherlands-Ranks-number-one-Worldwide-in-Penetration-for-Twitter-and-LinkedIn
[28] https://www.pewglobal.org/2014/02/13/emerging-nations-embrace-internet-mobile-technology/
[29] https://www.theatlantic.com/international/archive/2014/02/why-venezuelas-revolution-will-be-tweeted/283904/

Maduro made a public announcement of the "attack" on national TV.[30] Using the *wayback machine*, one can observe too that the event was on the front page of the *TeleSUR* website, one of the main TV networks sponsored by the Venezuelan government, on November 2nd (see Figure A6).[31] In the leading article, Maduro proposes the creation of a new social network to combat Twitter's bias against his government. On the other hand, coverage of the event on private media (which is more sympathetic to the opposition), including newspaper *El Universal* and television station *Globovisión*, focused on Twitter's policies and possible reasons for the forced closure of the accounts, including the use of "phantom" users and behaviour consistent with "spam".[32] This coverage is consistent with the position of the opposition, and likely reinforced the effect of the event on citizens' perceptions of Maduro's popularity.

Though a comprehensive analysis of how the event was portrayed in older media forms is outside of the scope of this study, this snapshot reveals patterns consistent with those observed on Twitter, and is congruent with the context of a hybrid media system (Chadwick, 2017). It also suggests that traditional media was likely to have contributed to users' updating their beliefs about the popularity of Maduro, and was one of the mechanisms supporting the unraveling of the existing 'spiral of silence'.

Data collection methodology

This section presents additional details on the collection methodology used to gather the dataset. I code a crawling tool using the python programming language and the 'requests' library (http://docs.python-requests.org). The tool makes calls to Twitter's advanced search engine, which are generally formatted in the URL as this:

https://twitter.com/search?**f=tweets**&vertical=default&q=twitter%20%40NicolasMaduro%20since%3A2013-11-01%20until%3A2013-11-02&src=typd&lang=en

---

30 A clip of which is available here: https://www.youtube.com/watch?v=ioTJQzTuPaU

31 The wayback machine is a digital archive of the Internet which stores web content at different points in time. The first date at which the TeleSUR website was archived after the event was November 3rd: https://web.archive.org/web/20131103001549/https://www.telesurtv.net/

32 See https://web.archive.org/web/20131104004753/http://globovision.com/articulo/sabe-usted-por-que-twitter-elimina-cuentas and http://www.eluniversal.com/nacional-y-politica/131102/estiman-que-twitter-elimino-seguidores-fantasmas-de-maduro

This page returns the last 20 tweets posted on November 1st, which contain the @NicolasMaduro username and the keyword "twitter".[33] The tool makes calls for each date and each selected keyword and stores the text of the site (using the functionality of requests), from which one can then retrieve the tweet-id and the number of replies for each result. Depending on whether one wants tweets containing a username, or tweets written by a specific user, the format varies slightly (these can be verified by examining the results from [Twitter's advanced search interface](https://twitter.com/search-advanced)). The tweets' unique IDs are then used to make calls to the Twitter API and "hydrate" the tweets dataset. The "**f=tweets**" argument tells Twitter to show the "latest" tweets, as opposed to those which the platform deems most important, since this could lead to sampling bias; whereas the last twenty tweets can be expected to be more representative of all tweets. Doing this also allows me to proxy tweet volume using the time of publication of these last twenty tweets, as discussed in the text.

Network analysis of selected accounts

Figure A7 shows the Twitter network structure of the selected accounts of the political leaders, which can be used to validate the account classification.[34] Each node represents one of the accounts and an edge is drawn between two nodes if either of them follows the other (though the underlying network is directed, I represent it as undirected for simplicity). Government accounts, shown as red squares, are clustered strongly together. Opposition accounts, shown as blue circles, are also clustered together, though more weakly so. Finally, the control-group sports-related accounts, are shown as green triangles. The bigger node shown as a brown square represents @NicolasMaduro, who is not only central in the government cluster, but also in the network overall.[35]

---

[33] Twitter search is not case or accent sensitive.

[34] Twitter network structure has been shown to be a predictor of political ideology (Colleoni et al, 2014; Barberá, 2015; Halberstam and Knight, 2016).

[35] There is one government outlier account that appears in the opposition cluster, which is that of Luisa Ortega Díaz (@lortegadiaz). She was the Prosecutor General appointed by Chavez and was loyal to Nicolas Maduro until 2017. By 2018 (when the network data was collected) she was a Maduro critic and, as the network graph suggests, had made "new friends" among the opposition. Since the tweets used are for 2013, I keep her as a government account. A qualitative review of her tweets during the period of study reveal mostly neutral positions. In addition, the results do not change in any meaningful way when removing her tweets from the analysis.

Keyword analysis of political leaders

In addition to the analysis on engagement in the main text, Table A3 reports keyword counts during the event study window for this dataset. It is worth highlighting that, for opposition tweets, the frequency with which the keyword *maduro* appears increases significantly (as is the case for the keyword *venezuela*), while the frequency of the username of Henrique Capriles (@hcapriles; the main leader of the opposition), and the opposition promoted hashtag #quenadatedetenga, decreases. For tweets by government leaders, the use of both the *maduro* keyword and Maduro's username @NicolasMaduro, increase. These patterns are consistent with the patterns documented for the other datasets and further show that the event was widely discussed by Venezuelan political leaders.

Validating the selection of pro-Maduro and anti-Maduro keywords

Together with a Research Assistant, we manually coded a subsample of tweets into "Pro-Maduro", "Neutral" and "Anti-Maduro" categories to validate the keyword choices for the tweets which mention @NicolasMaduro. Thirty tweets were selected (using a random number generator) for each of the 10 selected keywords. We then read each of these tweets and manually classified them across the categories. For the selected pro-Maduro keywords (*revolucion, camarada, comandante, chavez* and *victoria*), 65 percent of these are indeed supportive of Maduro, 29 percent are neutral, and 7 percent are critical of the president. On the other hand, for the selected anti-Maduro keywords (*regimen, ilegitimo, escasez, ladron,* and *maldito*), 92 percent are critical of Maduro, 7 percent are neutral, and 1 percent are supportive of the president. The complete results of this exercise, by keyword, are shown in Figure A8. Though the keywords are not perfectly predictive of the political stance of the tweets, they are strongly correlated with it.[36]

This measurement error in the political stance of tweets can lead to bias in the estimates of the effect of the accounts' closures on anti-Maduro and pro-Maduro tweet volume.[37] The estimates suggest that the volume in tweets using the pro-Maduro keywords increase by about 21 percent

[36] The correlation with the intended political stance is 0.83 coding the categories as -1 (anti-maduro), 0 (neutral), and 1 (pro-maduro).
[37] Note importantly that this is not "classical measurement error", the measure is instead an upper bound of the true increase in tweets of the intended sentiment.

(from Table 2, column 6, multiplied by the estimated publication time to volume factor of 0.485) and those using the anti-Maduro keywords increase by about 41 percent (Table 2, column 5); a differential increase of 20 percent. If only a share of this increase actually captures the intended sentiment, as suggested by the manual validation exercise, then the true increase in volume may be lower. A "back of the envelope" calculation using the results from the validation exercise would suggest that the increase in pro-Maduro tweets from these keywords is actually of 14 percent (21×0.65) and for anti-Maduro tweets of 38 percent (41×0.92); suggesting a differential increase of 24 percent more anti-Maduro tweets. Note importantly that since the anti-Maduro keywords capture the intended sentiment more precisely than the pro-Maduro keywords, this measurement error will tend to bias the estimates of the differential increase for anti-Maduro tweets downwards, suggesting that the estimates of the increase of anti-Maduro tweets relative to pro-Maduro tweets is a lower bound, or a conservative estimate for the actual change. [38]

Discussion of the difference-in-differences empirical methodology

This section presents additional details regarding the main identification strategy, as well as a placebo experiment, that help to illustrate the intuition behind the analysis.

The basic idea behind the research design can be illustrated using 2x2 tables. Table A4 shows the means of log(likes) for different periods, separately for the groups of interest. In panel A the periods coincide with the quasi-experiment of interest, comparing outcomes 15 days before the cancellation of the accounts, relative to 15 days after. In both of these time windows, opposition leaders (column 1) received on average more likes than the government (column 2, and the difference is presented in column 3). In addition, for both groups, log (likes) increased after the cancellation of the accounts (last row of panel A). However, the increase was larger for the opposition. Analogously, the gap between opposition and government was larger after the cancellation of the accounts. The difference in the differences of these means therefore captures the differential increase in likes that the opposition received, following the cancellation of the

[38] For this exercise, an even more conservative calculation can use the normal approximation of the binomial distribution for the proportion of correctly assigned tweets (note the anti-Maduro sample of 150 tweets with 0.08 failure rate is just large enough to allow this). In that case, one could use the upper bound of the 95% confidence interval for the pro-Maduro sample (0.7), and the lower bound for the anti-Maduro sample (0.89), this would suggest a 15 percent increase for pro-Maduro tweets, 36 percent increase in anti-Maduro tweets, and a 21 percent differential increase.

accounts. In particular, the estimates presented here suggest that opposition leaders received 0.242 more log (likes) relative to government leaders (column 3, last row of panel A, marked in bold).

One potential concern with these estimates could be that opposition leaders may have been on a faster growth path, even before the cancellation of the accounts. If this was the case, similar difference-in-differences estimates may be observed, but only because of these different pre-trends, not because of the cancellation of the accounts. In other words, the identification assumption of "parallel trends" would be violated. Figure 5 suggests that this was not the case, but a placebo experiment can provide additional evidence. Panel B repeats the exercise above but looking at estimates moving one 15-day period further back (October 2 to October 16). Note first that the gap between opposition and government persists for this period, but it is of a similar magnitude than the gap during the October 17 to October 31 period. Note too that there are no significant differences in engagement between this period and the next 15 days for either group (last row). Therefore, and reassuringly, the difference-in-differences estimate is statistically insignificant for this placebo experiment (-0.05).

Another potential bias affecting these difference-in-differences estimates could potentially arise if a very popular opposition leader (who on average gets lots of engagement) decides to tweet much more after the event, relative to less popular leaders. The preferred specification of the difference-in-differences estimator which presented in the main text includes user fixed-effects, which would account for this possible "selection into tweeting" mechanism. Therefore, the coefficients presented are within-user differences in these changes, that is, the difference is relative to engagement for the same user before the event. The estimates from the preferred specification which includes user fixed-effects, day fixed-effects and a series of controls, are smaller (0.214, Table 5, column 5), but similar in magnitude to the raw estimates presented here.

Lastly, Column 4 of Table A4 presents the means of log likes for the control group used in some specifications, the sports-related accounts. There are no statistically significant changes in the measure of engagement for these accounts, in neither the period of interest, nor the placebo experiment of the preceding time window.

Analysis of tweet volume in alternative keyword dataset

I analyze the volume to tweets which mention @NicolasMaduro to evaluate whether users' willingness to mention and criticize the president increased after the cancellation of the retweeting accounts. Directly addressing the president using his username @NicolasMaduro is an important act, as it can be viewed by him and others who follow him. Alternatively, however, one could simply look for tweets which contain the "maduro" keyword, regardless of whether they "tag" the president or not. Here I replicate the exercise of tweet volume using this alternative dataset. I collected tweets with the "maduro" keyword, as well as tweets with both "maduro" and each of the selected keywords (*venezuela, pueblo, twitter, revolucion, camarada, comandante, chavez, victoria, regimen, ilegitimo, escasez, ladron,* and *maldito*). The same restrictions as with the other datasets apply. The total number of tweets in this alternative dataset is 44,999.

I replicate tables 2 and 3 from the paper in this new dataset. These are presented in tables A5 and A6. The publication times suggest greater volumes of tweets overall for this new dataset, but the estimated changes in logs are generally similar in magnitude to those observed in the @NicolasMaduro dataset. The results from this alternative dataset suggest, relative to tweets including the @NicolasMaduro username, an even greater increase in the volume of tweets which mention the keyword *twitter* (Table A5, column 3, β=-2.119, p<0.001), which can be estimated to correspond to a 102 percent increase (-2.119×-0.485), and an even greater increase in tweets with anti-Maduro keywords (column 5, β=-0.918, p<0.001), or an estimated 45 percent (-0.918×-0.485), whereas there is a smaller and statistically insignificant increase in the volume of pro-Maduro tweets (column 6, β=-0.918, p<0.001). The results in table A6 suggest that the increase in anti-Maduro tweets relative to pro-Maduro tweets was statistically significant (column 6, β=-0.720, p<0.001), and represents an estimated differential increase of 35 percent.

The evidence presented in this section again confirms hypothesis H1a (increased criticism of the president). In fact, the support for the hypothesis is even stronger in this alternative dataset, which suggests even larger increases in the volume of anti-Maduro tweets, both overall and relative to pro-Maduro tweets.

Proxy measure of negative engagement

I use a proxy measure of negative engagement as an alternative outcome for the analysis. I construct the measure using log(replies + 1) – log(likes + 1). This measure considers the idea that tweets with a stronger negative reaction tend to receive more replies, but because many replies can also be positive, I use the number of likes to compensate for the strength of the positive feedback. This is illustrated with an example from current events using two recent tweets by Donald Trump in Figure A2. The tweets have similar number of likes and retweets but differ substantially in the number of replies. The more controversial tweet, which is subject to a strong negative reaction, has substantially more replies. The ratio of replies to likes is thus informative about negative reactions to a tweet.[39]

I replicate the main analyses with this measure as an outcome, which can help assess the statistical significance of the negative reaction. The results are presented in Table A7. For tweets by @NicolasMaduro, the proxy measure for negative engagement increases significantly after the cancelation of the retweeting accounts (column 2, β=0.163, p=0.047). For tweets by political leaders, there is an increase of about 11 percent in negative engagement for government leaders relative to sports-related accounts (column 3, β=0.106, p=0.059), and 16 percent greater negative engagement for government leaders relative to the opposition leaders (column 5, β=-0.156, p<0.001). The negative measure of engagement is not statistically different for the opposition relative to the sports-related accounts (column 4, β=-0.065, p=0.222). These findings are consistent suggest increased criticism of the government (congruent with H1a).

Heterogeneity across users' propensity to mention Nicolas Maduro

An additional empirical exercise studies whether the estimated effects are heterogeneous across users depending on their tendency to mention Maduro in their Twitter feeds. If users are more willing to express their relative support for the opposition because the president now appears less popular, then users who frequently mention Maduro may differentially benefit from the closure of the accounts (H3). I study this using both a difference-in-differences specification with a continuous treatment variable, as well as a triple-interaction framework as follows:[40]

[39] See also the definition of *#ratioed* (https://www.merriam-webster.com/words-at-play/words-were-watching-ratio-ratioed-ratioing).
[40] For an example of a triple-interaction design see for instance the analogous specification in Ferraz and Finan (2008), which examines whether government audits differentially affected municipalities with more local AM radio stations.

$$Y_{iut} = \beta_1 \cdot Post_t \cdot Opposed_u + \beta_2 \cdot Post_t \cdot MentionsMaduro_u$$

$$+ \beta_3 \cdot Post_t \cdot Opposed_u \cdot MentionsMaduro_u + X_{iut} \cdot \theta + \alpha_u + \alpha_t + \varepsilon_{iut}$$

where $MentionsMaduro_u$ is a variable that measures the frequency with which user u brings up the president in his or her tweets. The coefficient of interest, β_3, captures the differential engagement for users who mention Maduro more frequently, after the cancelation event, when coming from an opposition user. I present results both with and without fixed effects.

The analysis of heterogeneity across users' propensity to mention Maduro reveals that opposition leaders who mention Maduro more frequently experienced a larger relative increase in their number of likes (Table A8) relative to their peers (β=1.755, p=0.011), but the same is not true for government leaders. The results also reveal that the heterogeneity is differentially significant for opposition leaders relative to government leaders (β=1.502, p=0.034, for the preferred specification in column 6).

The coefficients suggest that an opposition political leader who mentions Maduro in ten percent of his tweets experienced a differential increase in tweet likes of around 17 percent after the accounts' cancelation, relative to an opposition leader who never mentions Maduro (1.755×0.1), and a differential increase of 29 percent relative to a government leader who also mentions Maduro in ten percent of his tweets (0.14 + 1.502×0.1). Figure A9 shows these marginal effects. The left panel shows the estimated change in log (likes) after the cancellation event based on the preferred fixed effects specification (Table A8, column 6). In addition, I show the results from a random effects model (right panel) that allows the estimation of the predicted log of likes, both before and after the event.[41] Before the cancellation of the accounts, higher propensity to mention Maduro was negatively associated with engagement for members of the opposition, but this relationship flips in the days after the event. On the other hand, there is no significant relationship between engagement and users' tendency to mention the president for government leaders, neither before or after the event.

[41] Note that the *post* indicator is collinear with day-fixed effects in the preferred specification shown in Table A8, column 6, for this reason, I show the alternative model as well.

That opposition users who more frequently mentioned Maduro on the platform benefitted differentially from the cancelation of the accounts suggests that the change in the perception of Maduro's popularity was stronger for followers of these users, and therefore they experienced larger gains in political support, as measured by their follower engagement. This finding is consistent with hypothesis H3 and suggests that opposition followers who were differentially aware of Maduro's popularity on the platform reacted more strongly after the accounts were closed by Twitter.

Long-run analysis

As discussed in the main text, using a short window of time allows me to get closer to being able to infer a "causal relationship" between the cancellation of the accounts and the empirical facts documented. Many of the patterns in the medium and long-run can be viewed in the figures, but I formally replicate the main empirical exercises here using a 6-month long window (instead of a 30-day window), such that I compare outcomes in the 3-months after the cancellation of the accounts, relative to the 3-months before. Given that many other events can confound the analysis in the long-run, the exercise presented here should be viewed as descriptive. I replicate only the main results, and when possible, prioritize the difference-in-difference specifications (the pre/post analysis is more problematic due to seasonal trends, including the presence of Christmas and New Year's during the long-run window).

The results are presented in Table A9. In contrast with the short-run evidence, there is no observed increased differential volume for anti-Maduro tweets in the @NicolasMaduro dataset (column 3), or increased differential support for opposition users who more frequently mention @NicolasMaduro (column 7, this estimate is very imprecise but could be explained by the unraveling dynamics, such that the effects may initially be concentrated in these users but spread more broadly in the longer run). On the other hand, the patterns of increased replies for @NicolasMaduro (column 2), increased differential support for the opposition (column 5), and increased volume of anti-Maduro tweets in the alternative dataset (column 4, using "Maduro" keyword), are also present in the long-run.

Appendix Tables and Figures

Table A1: @NicolasMaduro Followers

Date	Number of followers
March 2013	553,064
April 2013	983,123
May 2013	1,144,893
June 2013	1,207,997
July 2013	1,257,782
August 2013	1,320,013
September 2013	1,376,534
October 2013	1,418,953
November 2013	1,513,680
December 2013	1,575,707
January 2014	1,651,921

Source: twven.com

Table A2: Selected accounts for political leaders' dataset

Government	Opposition	Control
NicolasMaduro	hcapriles	MeridianoTV
dcabellor	leopoldolopez	SaschaFitness
jaarreaza	MariaCorinaYA	CarolinaPadron
JauaMiranda	liliantintori	leones_cbbc
TareckPSUV	ComandoSB	Pastormaldo
luislopezPSUV	alcaldeledezma	caroguillenESPN
jorgerpsuv	JJRENDON	greivisvasquez
taniapsuv	CarlosOcariz	LVBP_Oficial
JacquelinePSUV	hramosallup	Magallanes_bbc
gestionperfecta	HenriFalconLara	SalvadorPerez15
lortegadiaz	ramonmuchacho	6cichero6
HugoCabezas78	JulioBorges	la_grulla5
DanteRivasQ	luisvicenteleon	OzzieGuillen
ConCiliaFlores	VoluntadPopular	GarbiMuguruza
AmeliachPSUV	unidadvenezuela	BobKellyAbreu
DrodriguezVen	JuanRequesens	VizquelOmar13
NestorReverol	MiguelContigo	EUDeporte
garcesfrancisco	Diego_Arria	caraquistas
jdavidcabello	Pr1meroJusticia	porlagoma
Adan_Coromoto	rociosanmiguel	emanuelatleta
PartidoPSUV	GerardoBlyde	Magallanes_News
maperezpirela	ismaelprogreso	LuisAlvarez_1
psuvaristobulo	delsasolorzano	salorondon23
JuventudPSUV	carlosvecchio	aguilasdelzulia
blancaePSUV	PadreJosePalmar	Caracas_FC
ErikaPSUV	jcsosazpurua	Tomapapa
PanchoArias2012	PabloPerezOf	Lavinotintocom
irisvarela	FreddyGuevaraC	Alex_candal
FreddyBernal	VicenteDz	rubenoszki
HectoRodriguez	GenPenaloza	MaxCordaro
IzarraDeVerdad	DavidUzcategui	FerAlvarez
WalterDossier	Simonovis	CardenalesDice
vladimirpadrino	alfredoromero	RichardGol_espn
RobertSerraPSUV	manuelrosalesg	hturinese
jesusfariaPSUV	dsmolansky	OficialTigres
jchacon2021	antonioriverog	DIRECTVSportsVE
PedroCarreno_e	AndresVelasqz	cuantoacuanto
IsisPSUV	ENZOSCARANO	ElvisandrusSS1
jorgeamorin	BMarmoldeLeon	SeleVinotinto
AndreinaTarazon	alFranceschi	Arango_18
anat5	EvelingTrejo	milenagimon
danicabello11	juanjosemolina	JoseAltuve27
durancandanga	HimiobSantome	DvoTachira
tongorocho	plomoparejo	adriana_donghia
nicmaduroguerra	humbertotweets	JuanPaGalavis
PatriciaDorta40	orlandourdaneta	fpetrocelli
RALDAHIR	Gral_Vivas_P	GatoradeVzla
Cosole_Roja	mferreiratorres	DeportivoLara
MQuevedoF	MarceloNolla	dtcesarfarias
Marlenycdc	manocompa	MecheCelta

Table A3: Most frequently used keywords in tweets of political leaders

Opposition (period / num tweets)				Government (period / num tweets)			
Pre / 5,531		*Post / 5,949*		*Pre / 3613*		*Post / 4053*	
frequency	**keyword**	**frequency**	**keyword**	**frequency**	**keyword**	**frequency**	**keyword**
406	quenadatedetenga	476	venezuela	307	chavez	419	pueblo
382	hcapriles	455	maduro	286	forocandanga	368	chavez
310	venezuela	347	quenadatedetenga	283	nicolasmaduro	344	nicolasmaduro
310	maduro	319	hcapriles	238	venezuela	297	forocandanga
309	gobierno	314	gobierno	229	pueblo	237	hoy
299	pueblo	279	pueblo	212	psuv	211	venezuela
239	unidad	252	pais	152	maduro	207	psuv
238	pais	226	unidad	140	patria	196	maduro
237	caracas	164	venezolanos	131	revolucion	189	presidente
178	cambio	159	vecinos	130	gobierno	179	gobierno
161	vecinos	155	gracias	127	apoyoanicolasmaduro	175	contra
150	gracias	151	cambio	118	presidente	165	patria
136	progreso	133	contra	115	zulia	147	tropa
125	seguridad	132	plan	112	vivachavezcarajo	143	revolucion
125	baruta	127	regimen	112	bolivar	136	durancandanga
122	candidatos	125	ahora	111	simulacro	134	ubch
121	regimen	124	gente	105	contra	131	victoria
117	dias	121	dia	103	fotos	121	apoyoanicolasmaduro
117	contra	121	caracas	103	estado	119	fotos
110	bastaya	116	diputado	98	victoria	115	vivachavezcarajo
12	twitter	62	twitter	11	twitter	34	twitter

Notes: The table shows keyword counts for tweets by political leaders before and after the accounts' cancellations. Sample includes tweets by political leaders (government and opposition) in a 30-day window around October 31, 2013.

Table A4: Illustration of the difference-in-differences research design

	Mean of log (likes)			
	Opposition	Government	Difference (O-G)	Sports
	(1)	(2)	(3)	(4)
Panel A: Quasi-experiment of interest				
November 1 to November 15	1.233***	0.719***	0.514***	0.648***
	(0.134)	(0.104)	(0.168)	(0.075)
October 17 to October 31	0.903***	0.631***	0.273*	0.606***
	(0.127)	(0.091)	(0.156)	(0.085)
Difference (post - pre)	0.330***	0.088**	**0.242*** **	0.042
	(0.052)	(0.041)	**(0.066)**	(0.035)
Panel B: Placebo experiment				
October 17 to October 31	0.903***	0.631***	0.273*	0.606***
	(0.127)	(0.091)	(0.156)	(0.085)
October 2 to October 16	0.960***	0.638***	0.322**	0.643***
	(0.121)	(0.086)	(0.147)	(0.103)
Difference (post - pre)	-0.057	-0.007	**-0.05**	-0.038
	(0.039)	(0.032)	**(0.050)**	(0.039)

Notes: Sample includes tweets by political leaders (and sports-accounts) in the specified date ranges. Outcome measured is log(likes). Standard errors clustered at the user level in parentheses. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table A5: Relationship between accounts' cancellation and publication times for tweets containing "Maduro"

Keyword sample:	None (last 20 overall)	*venezuela*	*pueblo*	*twitter*	Anti-Maduro (pooled)	Pro-Maduro (pooled)
	(1)	(2)	(3)	(4)	(5)	(6)
Post	-0.718***	-0.575**	-0.522	-2.119***	-0.929***	-0.217
	(0.196)	(0.255)	(0.425)	(0.366)	(0.151)	(0.135)
N	600	600	600	600	2,746	2,784
N-clusters	30	30	30	30	150	150
Keyword fixed-effects	No	No	No	No	Yes	Yes

Notes: Outcome measured is time of publication, measured in log seconds to the end of the day, as a proxy for tweet volume. Less time to the end of the day indicates greater tweet volume. All of the specifications include day of the week fixed effects. Standard errors clustered at the day level for columns 1-4, and clustered at the keyword-day level for columns 5-6. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table A6: Differential change in publication times for tweets containing "Maduro"

Keyword sample:	Pro-Maduro vs. Neutral		Anti-Maduro vs. Neutral		Pro-Maduro vs. Anti-Maduro	
	(1)	(2)	(3)	(4)	(5)	(6)
Post	-0.617**		-0.610**		-0.265	
	(0.260)		(0.257)		(0.234)	
Pro-Maduro	1.882***					
	(0.234)					
Post * Pro-Maduro	0.358	0.416*				
	(0.348)	(0.240)				
Anti-Maduro			3.308***		1.392***	
			(0.196)		(0.177)	
Post * Anti-Maduro			-0.301	-0.294	-0.657**	-0.720***
			(0.298)	(0.216)	(0.280)	(0.171)
N	3,839	3,839	3,889	3,889	5,484	5,484
N-clusters	210	210	210	210	300	300
Day of the week fixed-effect	Yes	No	Yes	No	Yes	No
Keyword fixed-effects	No	Yes	No	Yes	No	Yes
Day fixed-effects	No	Yes	No	Yes	No	Yes

Notes: Outcome measured is time of publication, measured in standardized time until the end of the day, as a proxy for tweet volume. Less time to the end of the day indicates greater tweet volume. Standard errors clustered at the keyword-day level. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table A7: Effects on proxy for negative engagement (log((replies+1)/(likes+1)))

Sample:	Tweets by @NicolasMaduro		Government and Sports	Opposition and Sports	Opposition and Government
	(1)	(2)	(3)	(4)	(5)
Post	0.195**	0.163**			
	(0.086)	(0.082)			
Post * Government			0.106*		
			(0.056)		
Post * Opposition				-0.065	-0.156***
				(0.052)	(0.041)
N	183	183	20,830	24,644	19,146
N-clusters			90	94	92
Controls	No	Yes	Yes	Yes	Yes
Day fixed-effects	No	No	Yes	Yes	Yes
User fixed-effects	No	No	Yes	Yes	Yes

Notes: All specifications use the proxy for negative engagement as the dependent variable. Sample includes tweets by @NicolasMaduro (columns 1 and 2) and tweets by prominent leaders (government, opposition and sports accounts; columns 3-5) in a 30-day window around October 31, 2013. Columns 1-2 use the pre/post regression specification and columns 3-5 use a difference-in-differences specification. Robust standard errors (columns 1-2) and standard errors clustered at the user level (columns 3-5) in parenthesis. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table A8: Differential change in tweet likes by propensity of political leaders to mention Maduro

Sample:	Government		Opposition		Opposition vs. Government	
Model:	Difference-in-differences				Triple-difference	
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.024		0.240***		0.024	
	(0.054)		(0.056)		(0.054)	
MentionsMaduro	-0.699		3.281**		-0.699	
	(0.641)		(1.400)		(0.637)	
Post * MentionsMaduro	0.436	0.291	1.603*	1.755**	0.436	0.256
	(0.361)	(0.189)	(0.894)	(0.661)	(0.358)	(0.185)
Opposition					-0.037	
					(0.179)	
Post * Opposition					0.216***	0.140**
					(0.077)	(0.063)
Opposition * MentionsMaduro					3.980**	
					(1.532)	
Post * Opposition * MentionsMaduro					1.167	1.502**
					(0.959)	(0.699)
N	7,666	7,666	11,480	11,480	19,146	19,146
N-clusters	44	44	48	48	92	92
Controls (not shown)	No	Yes	No	Yes	No	Yes
Day fixed-effects	No	Yes	No	Yes	No	Yes
User fixed-effects	No	Yes	No	Yes	No	Yes

Notes: Dependent variable is natural log of number of likes for all columns. Sample includes tweets by political leaders in a 30-day window around October 31, 2013. Standard errors clustered at the user level in parenthesis. Significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Table A9: Descriptive long-run analysis (6-month window)

Sample:	Tweets by @NicolasMaduro		Tweets that mention @NicolasMaduro	Tweets that contain "Maduro" keyword	Tweets by opposition and government leaders		
Dependent variable:	Log retweets	Log replies	Log seconds to end of day		Log likes	Log replies	Log likes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post	-1.736***	0.599***					
	(0.021)	(0.046)					
Post*Anti-Maduro			-0.055	-0.302***			
			(0.058)	(0.071)			
Post*Opposition					0.319***	0.215***	0.338***
					(0.057)	(0.066)	(0.073)
Post*Opposition*MentionsMaduro							-0.795
							(1.092)
N	932	932	30,813	32,243	107,795	107,795	107,795
N-clusters			1,839	1,840	94	94	94
Original specification	Table 1, Column 2	Table 1, Column 6	Table 3, Column 6	Table A6, Column 6	Table 5, Column 5	Table 5, Column 6	Table A8, Column 6

Notes: The table shows the replication exercises for the main results in the 6-month long window. Refer to original tables for details on specification, controls, and clustering of standard errors. Results significant at (*) 90 percent, (**) 95 percent, (***) 99 percent confidence levels.

Figure A1: Political leaders' reactions to the closure of the accounts

Figure A2: Example of reply counts for controversial (right) vs non-controversial (left) tweet

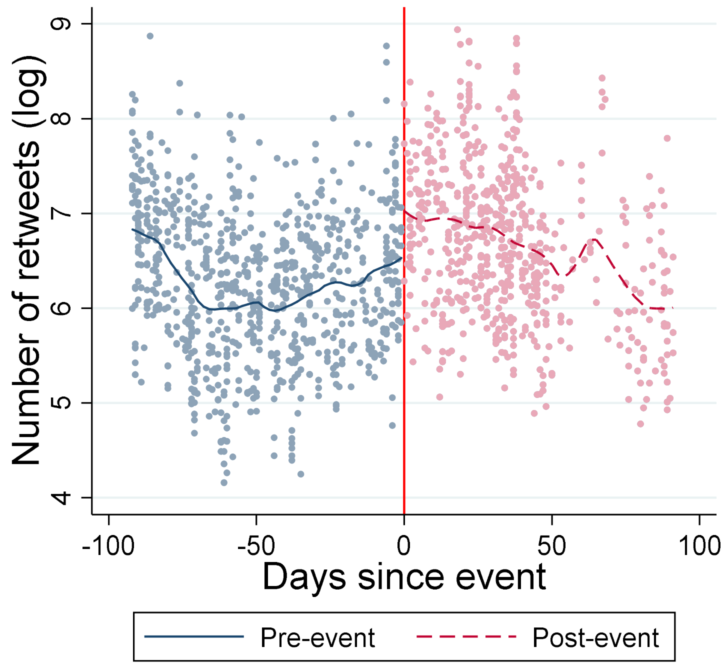

Figure A3: Retweets for the leader of the opposition Henrique Capriles (@hcapriles)

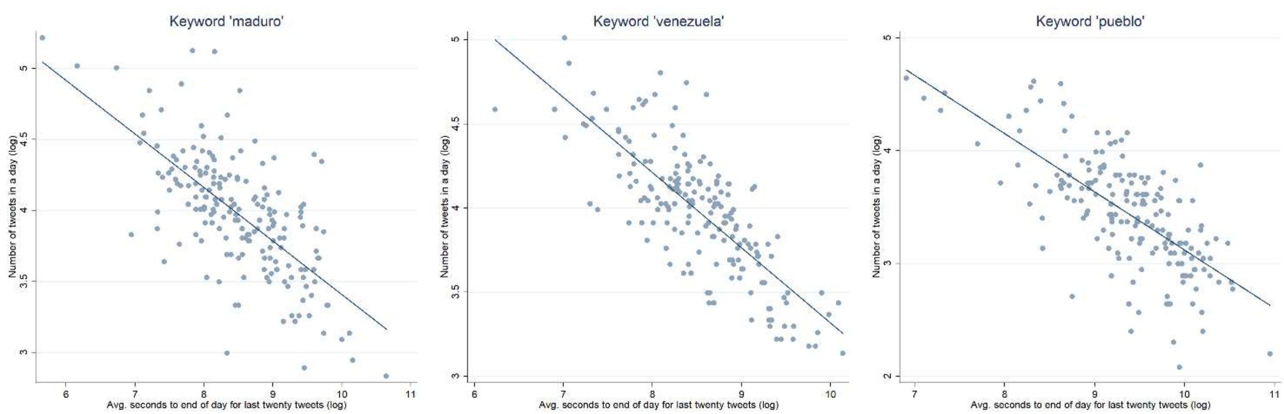

Figure A4: Relationship between number of tweets and time of publication of last twenty tweets at the daily level (each scatter point is one day)

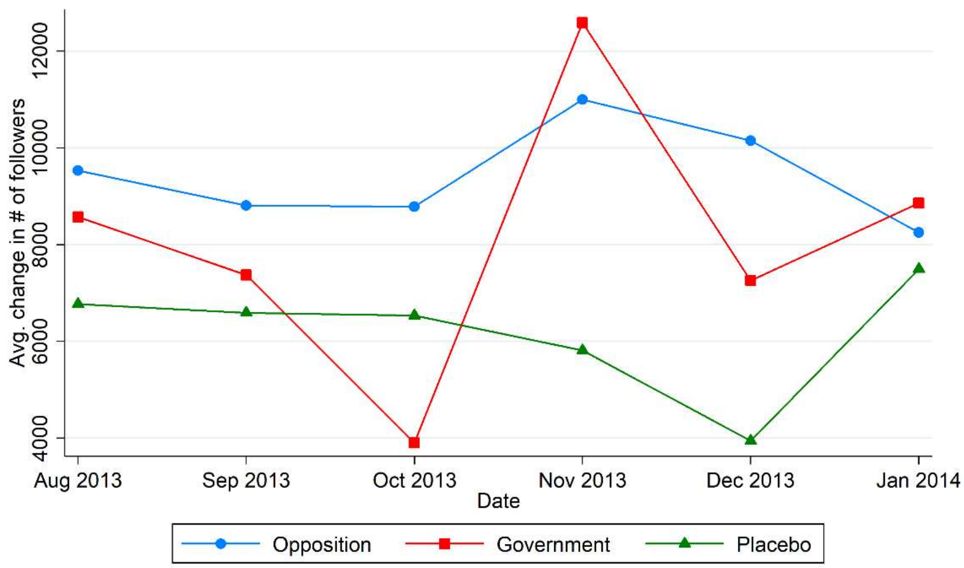

Figure A5: Average change in number of followers for selected accounts

Figure A6: Front page of the TeleSUR website on November 2nd/3rd (the website is dated November 2nd though the archive date is November 3rd)

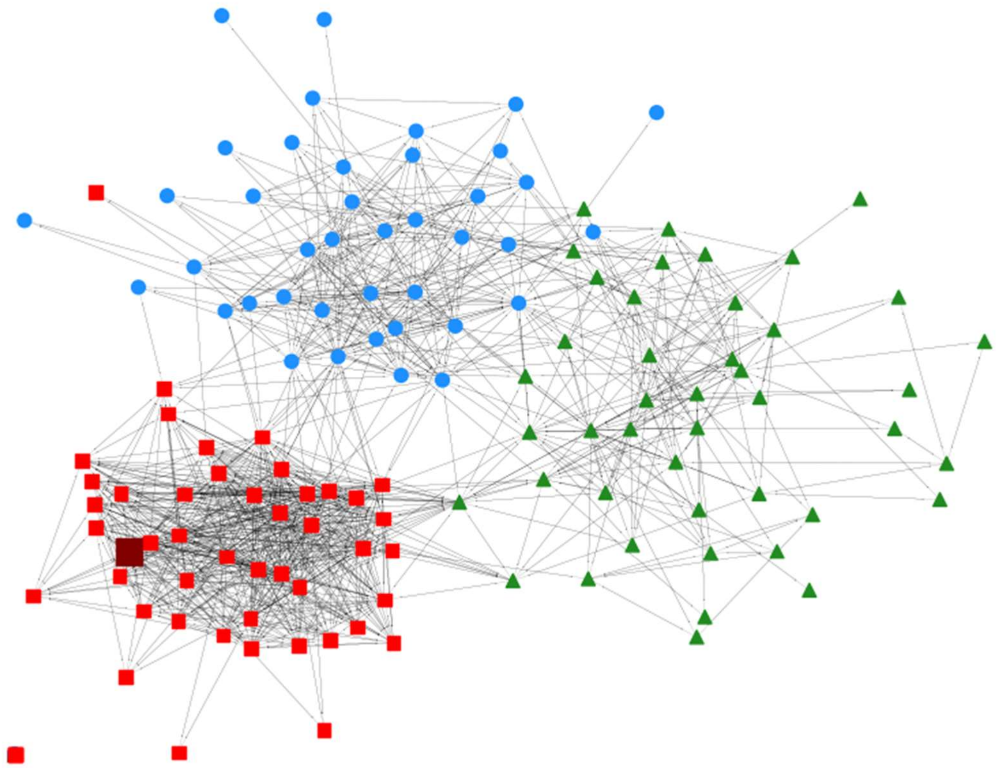

Figure A7: Twitter network of selected accounts

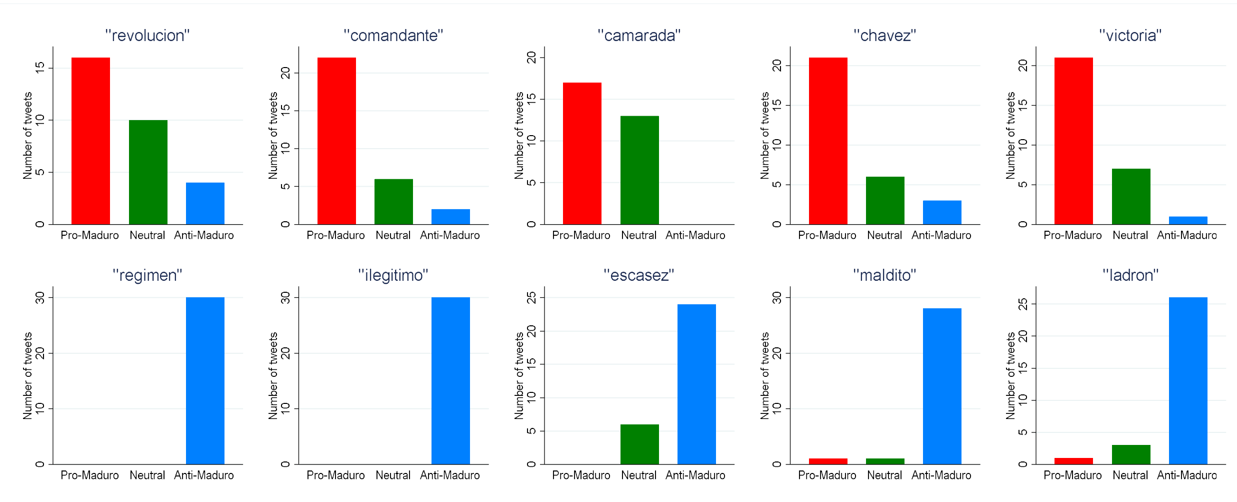

Figure A8: Results from manually coding a subsample of 300 randomly selected tweets which mention @NicolasMaduro, by keyword

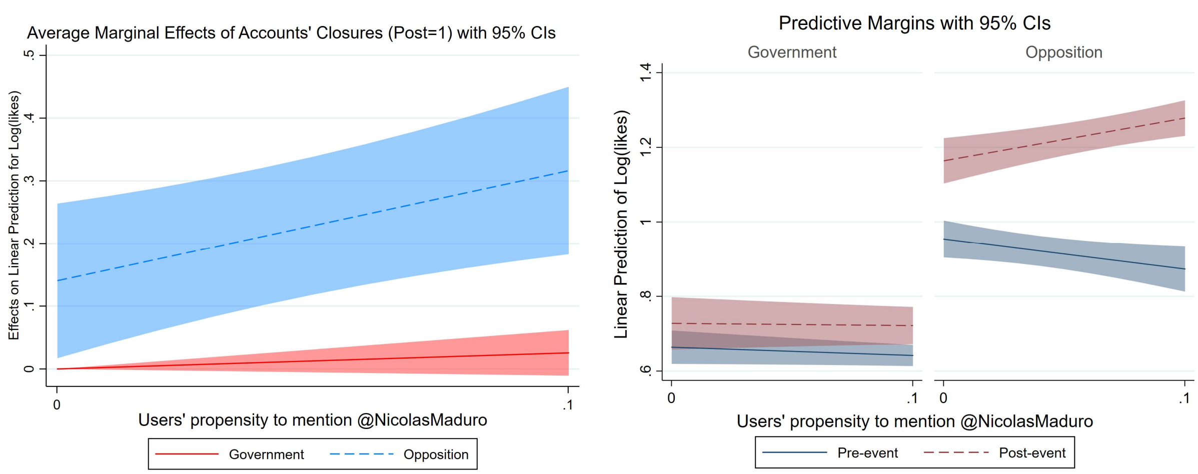

Figure A9: Heterogeneity by users' propensity to mention @NicolasMaduro

Model specification: Fixed effects (left) and Random effects (right)