

# UCLA

## UCLA Previously Published Works

### Title

How State and Protester Violence Affect Protest Dynamics.

### Permalink

<https://escholarship.org/uc/item/6d63p84b>

### Journal

The Journal of Politics, 84(2)

### ISSN

0022-3816

### Authors

Steinert-Threlkeld, Zachary

Chan, Alexander

Joo, Jungseock

### Publication Date

2022-04-01

### DOI

10.1086/715600

Peer reviewed



Published in final edited form as:

*J Polit.* 2022 April ; 84(2): 798–813. doi:10.1086/715600.

## How State and Protester Violence Affect Protest Dynamics

**Zachary C. Steinert-Threlkeld,**

University of California, Los Angeles

**Alexander M. Chan,**

University of California, Los Angeles

**Jungseock Joo**

University of California, Los Angeles

### Abstract

How do state and protester violence affect whether protests grow or shrink? Previous research finds conflicting results for how violence affects protest dynamics. This article argues that expectations and emotions should generate an n-shaped relationship between the severity of state repression and changes in protest size the next day. Protester violence should reduce the appeal of protesting and increase the expected cost of protesting, decreasing subsequent protest size. Since testing this argument requires precise measurements, a pipeline is built that applies convolutional neural networks to images shared in geolocated tweets. Continuously valued estimates of state and protester violence are generated per city-day for 24 cities across five countries, as are estimates of protest size and the age and gender of protesters. The results suggest a solution to the repression-dissent puzzle and join a growing body of research benefiting from the use of social media to understand subnational conflict.

---

Although formal and empirical research has established the importance of large protests (DeNardo 1985; Wouters and Walgrave 2017), their dynamics remain less understood. For example, Biggs (2003) argues for a positive feedback loop but does not specify when an initial protest is more likely to generate that process, and built environments or electoral fraud encourage protest participation (Tucker 2007; Zhao 1998). Empirical investigations have generated contradictory results for decades, making repression and protest dynamics an enduring puzzle (Davenport 2007). Studying these dynamics has been difficult because existing methodologies generate coarse (ordinal) estimates of violence and have difficulty measuring the size of protests, possibly contributing to the literature's conflicting results.

How violence affects protest dynamics depends on its source and severity. When it comes from the state, low amounts of violence mobilize more protesters while high amounts

---

Zachary C. Steinert-Threlkeld (zst@luskin.ucla.edu) is an assistant professor at University of California, Los Angeles, Los Angeles, CA 90095. Alexander M. Chan (alexander.chan@ucla.edu) was a graduate research assistant at University of California, Los Angeles, and is now an Infrastructure Data Scientist at Facebook. Jungseock Joo (jjoo@comm.ucla.edu) is an assistant professor at University of California, Los Angeles, Los Angeles, CA 90095.

Replication files are available in the *JOP* Dataverse (<https://dataverse.harvard.edu/dataverse/jop>). The empirical analysis has been successfully replicated by the *JOP* replication analyst. An online appendix with supplementary material is available at <https://doi.org/10.1086/715600>.

demobilize them, creating an n-shaped relationship between state violence and subsequent protest size. Lower than expected costs to protest and emotional backlash increase protest size, while higher than expected costs and dispiriting emotions shrink it. Protester violence always leads to smaller protests because it decreases the appeal of protesting and increases the probability of state violence.

In addition to explaining the importance of the source and severity of violence, this article improves the measurement of violence and protest size with a new methodology (computer vision) applied to large data (millions of geolocated tweets containing images). A convolutional neural network (CNN) is developed to recognize protest images. Of 42.6 million tweets from protest waves across five countries, 4.6 million contain an image. Approximately 115,000 of these images likely contain a protest. A second CNN generates binary and continuous measures of state and protester violence; these classifiers out-perform Google Vision, a third-party CNN. This scene classifier is complemented with a third CNN, a face classifier. This classifier estimates the gender and age of each face, allowing us to control for well-known correlations between these demographic features and protest participation (Nordås and Davenport 2013; Schaftenaar 2017). Summing faces in protest images generates estimates of protest size, and extensive corroboration shows that these estimates are consistent with actual protest size. This pipeline generates daily estimates of the size of protest, the severity of state and protester violence continuously valued from [0, 1], and potential confounds for 24 cities in five countries.

The next section explains how state and protester violence should affect protest size. After, detail on measuring and validating the concepts is provided. Then we explain how the concepts are measured and validate the classifier estimates. The penultimate section presents results and a battery of robustness checks. The conclusion suggests directions for future research.

## PROTEST DYNAMICS

Protesters want to join a large protest since they have larger net benefits because the probability of suffering repression is lower and the probability of policy change is higher. Costs and benefits are uncertain a priori, however, so individuals use realized state and protester violence to calculate their payoffs (Shadmehr and Bernhardt 2011). State violence causes different responses depending on its severity, generating an n-shaped relationship between it and subsequent protest size. Protester violence should always decrease protest size because it decreases the consumption benefit of protest while increasing the probability of repression.

### The importance of large protests

Three assumptions lead to the conclusion that large protests are more likely to change policy than small ones. If (1) the purpose of a protest is to convince political leaders to change a policy, (2) a leader cares about the median voter (Downs 1957) or his or her winning coalition exhibits some response to the median person (Bueno de Mesquita et al. 2003), and (3) a large protest's policy preference is closer to the median individual than a small one's, then a large protest is more likely to change policy than a small one.

Protest size matters regardless of a country's political institutions. In democracies, voting is the most common method of policy change but occurs infrequently. Protests, however, can occur at any time and usually have a clear policy goal (Battaglini 2017), and they provide an additional outlet for revealing pressing sentiments. While protest is unlikely to change an autocrat's policy, it nonetheless provides a key signal of discontent to which a government can respond (Bratton and Van de Walle 1992). This signal is especially pertinent if opinion polling is unreliable (Robertson 2007) or the media are not free (Qin, Strömberg, and Wu 2017).

This argument holds without assuming a leader aims for the median individual's policy preference. If a leader only desires to stay in power and a large protest means the probability of remaining in power is lower than the leader previously believed, a large protest is still more likely to lead to policy change than a small one. Although a regime should become less responsive to protests as its selectorate shrinks (Bueno de Mesquita et al. 2003), the importance of large protests increases inversely with the size of the selectorate since other mechanisms of policy change are foreclosed. This theory should therefore apply in a wide range of regimes such as rentier states, single-party regimes, and those with low Polity scores. While a large protest is not necessarily successful, almost all successful protests are large.

### State violence

State violence at protests should generate an n-shaped relationship via two mechanisms.<sup>1</sup> First, how protesters and bystanders respond to repression depends on its severity relative to expectations, not an absolute level. If individuals protest when the expected benefit outweighs the expected cost and the probability of suffering repression affects expected cost (Tullock 1971), then individuals decide to mobilize because of some expectation of harm. If the realized probability of harm is lower than the value an individual used in a personal calculation, then there will be bystanders who now expect to benefit from protesting. So long as the repression is not larger than expected for a larger number of those already protesting, repression can generate growth in protest size. If, however, repression is more severe than expected, more protesters will demobilize than bystanders mobilize, and protest size decreases. This relative expectation mechanism operates similarly to expectations about protester identity in information-signaling models (Lohmann 1994).

Second, repression triggers reflexive emotional responses that can cause more or fewer people to mobilize, depending on the emotions triggered (Jasper 2011). When repression generates emboldening emotions such as anger, joy, or pride, protesters are more likely to persist and bystanders are more likely to mobilize. Anger incites a desire to confront the agents committing repression and facilitates blame attribution against the state. Joy reflects pleasure as protesters sense progress toward their policy goal. Pride enhances feelings of self-worth and belonging, especially in contrast to "bad" state agents (Pearlman 2013).

Repression can also trigger dispiriting emotions such as fear, sadness, and shame, causing protest size to decrease. A fearful protester is more likely to cease protesting in order

---

<sup>1</sup> "State violence" refers to what others call "protest policing" (Earl et al. 2013).

to escape the threat of repression. A sad one has determined that the current policy cannot be changed by individual action. Others may feel shame, the belief that they have personally failed. In these cases, protesters demobilize and bystanders continue standing by: protests shrink in size. In addition, these emotions change individuals' risk aversion, perceived repression risk, and strategic consideration about other potential protesters, making individuals less likely to join a protest (Young 2019).

Emotions are especially useful for understanding how low amounts of repression can cause larger protests even when following no repression. If a protest at  $t_0$  has  $p_0$  participation and no repression, any repression at  $t_1$  should cause  $p_1 < p_0$ , ceteris paribus. But if repression induces feelings of anger, joy, or pride in more people than it creates fear, sadness, or shame,  $p_1 > p_0$ . This backlash effect is in addition to any increase in size due to the first mechanism, difference in expectations. As Siegel (2011, 1005) shows: "Less obvious is that, *if anger is strong enough, participation levels can be higher under repression than absent it. Individual anger at local repression endogenously enables aggregate backlash.* Further, for weak repression, comparatively little anger is needed to achieve backlash against the repressive entity." Emotions do not transform a calculating individual into an instinctual one. They do not replace calculation: they are inputs to that process.

Finally, we assume that a low level of repression triggers emboldening emotions while severe repression induces the dispiriting ones. If expectations and emotions influence protesters as described and the relationship between repression and triggered emotions follows the just-stated assumption, then:

**H1.** There exists an n-shaped relationship between the severity of state repression and the size of a protest the next day.

This relationship should hold in democracies and autocracies. For example, 2005 protests against chemical plants in Huashi, Zhejiang, drew much greater participation after an initial attempt to remove protesters' encampments led to hundreds of injuries; authorities eventually closed the plants (O'Brien and Deng 2015). The Occupy Wall Street movement in the United States experienced a surge in participants after New York City police arrested over 700 protesters marching on the Brooklyn Bridge, generating anger. In Egypt, the protests starting on January 25, 2011, were met with initial state resistance and some casualties; 18 days later, the Armed Forces forced President Hosni Mubarak to abdicate. Two years later, the Armed Forces launched a coup against the elected president, Mohamed Morsi. Large pro-Morsi protests erupted and continued for six weeks. The Armed Forces' initial attempts to demobilize them were counterproductive; finally, morning massacres on August 14 at the two main encampments killed at least 1,000 protesters, injured even more, and heralded the return of military rule (Shakir 2014).<sup>2</sup>

---

<sup>2</sup>The expectation of backlash conditional on repression severity is consistent with Francisco's (2004) argument for a strictly positive relationship between the two. Backlash in that theory encompasses nonprotest actions such as strikes, building occupations, or guerrilla action and allows for this substitution to occur much later (Moore 2000). The tests of that backlash theory also find that protesters initially demobilize in response to severe repression, in line with this article's argument.

## Protester violence

Protesters can also engage in violence, so a theory of protest dynamics should take their action into account. Protester violence should always decrease the size of protests for two reasons: it decreases the number of bystanders to whom protesting appeals and increases the cost of protesting to the remaining bystanders not deterred by protester violence.

One method by which bystanders determine whether to mobilize is to compare protesters' ideological distance from their own (Lohmann 1994). Since most individuals do not support violence or receive consumption value from it (Muñoz and Anduiza 2019; Simpson, Willer, and Feinberg 2018), protester violence signals that protesters, and therefore the policy changes for which they agitate, are likely not near the median policy preference. Being far from the mainstream, bystanders continue to stand by because the new policy the violent protesters seek is inferred to not reflect nonprotesters' preferred policy.

Protester violence decreases the likelihood of regime defections, further decreasing the pool of potential protesters. Peaceful protest convinces regime agents of their physical safety should they defect, increasing the probability that police, members of the armed forces, or legislators switch allegiances (Stephan and Chenoweth 2008). Violent protesters, however, induce fear in these agents that they will meet the same fate if they do not remain loyal. Violence therefore reduces the pool of those willing to protest, making the state stronger than if facing an otherwise equivalent peaceful protest.

Protests containing protester violence are seen as less legitimate than peaceful ones (Bashir et al. 2013), increasing the probability of repression and therefore the cost of protest. Peaceful protests enjoy high domestic and international legitimacy, so state violence against them risks generating a backlash that increases subsequent protests' size. But since the state can frame violent protesters as rioters, terrorists, or foreign agitators (Benford and Snow 2000), bystanders are more supportive of repressing violent protests than nonviolent ones (Murdie and Purser 2017). For the same reasons, the state is also less likely to receive international sanction when repressing violent protests.

Conversely, protester nonviolence increases the probability that a protest grows in size, especially when states repress. Because nonviolence increases the legitimacy of protests, it decreases the probability that a state represses, as the state will pay large reputation costs. The lower probability of repression induces more bystanders to mobilize, generating a positive feedback loop (Biggs 2003). In Morocco, for example, attempts to repress nonviolent protesters at the start of the Arab Spring led to larger crowd sizes (Lawrence 2017), and government violence in Tunisia did not prevent the spread of those protests.

Since protester violence alienates bystanders and potential regime defectors and increases the cost of protesting, it should be that:

**H2.** There exists a negative relationship between the severity of protester violence and the size of a protest the next day.

## METHODOLOGY

Research finds that state repression decreases protest (Moore 2000; Olzak, Beasley, and Olivier 2003), increases it (Davenport and Armstrong 2004; Gurr and Moore 1997; Hess and Martin 2006), or has no effect (Gupta, Singh, and Sprague 1993; Ritter and Conrad 2016). While this article is not the first to suggest repression severity generates an n-shaped relationship, previous studies have relied on coarse measures such as ordinal variables or annual number of political detainees at the country level. By generating a more precise measure of the severity of repression, this article resolves these contradictory findings.

Since the effect of state violence should vary on the basis of its severity, precise measures of it are required. Image analysis enables this precision in two ways. First, CNNs will generate continuously valued estimates from [0, 1]. Current leading data sets, by contrast, map violence onto different discrete categories (values of an ordinal variable). For example, the Social Conflict Analysis Database, Urban Social Disorder, and Armed Conflict Location and Event Data (ACLED) data sets record repression as a binary variable (Raleigh et al. 2010; Salehyan et al. 2012; Urdal and Hoelscher 2012). Repression elsewhere is coded as ordinal or nominal (Clark and Regan 2016; Goldstein 1992; Stephan and Chenoweth 2008), including in machine-coded event data (Boschee et al. 2015). Second, whether one text (e.g., a newspaper article) shows that a protest is violent will depend on the language used and the research teams' interpretation of that language. Researchers thus necessarily have to use coarse measures for large-n analysis. The closest continuously valued measure of repression is fatalities, which is not always recorded and is the most severe type of state violence. When focusing on violence in one setting, scholars have been more successful at disambiguating it to generate ordinal measures (Khawaja 1993; Olivier 1991).

This article develops three classifiers based on CNNs to automatically code the variables of interest: one identifies protest images, a scene classifier extracts data from them, and a face classifier generates size estimates and demographic controls.<sup>3</sup> The pipeline outlined in table 1 resembles the approach described in Zhang and Pan (2019). Step 1 uses keywords, including hard negatives like “concert” or “stadium,” and Google Image Search to acquire training images. Step 2 then trains a CNN on these images. Step 3 uses this classifier to identify protest images from just under 43 million geolocated tweets, resulting in 40,764 protest images. In step 4, workers from Amazon Mechanical Turk label these Twitter images. To measure state and protester violence, annotators are presented pairs of images and asked which is more violent, and the Bradley-Terry model generates continuous estimates from the resulting ordering (Bradley and Terry 1952). Each image is coded by at least two individuals; a third is used to break ties. These labels then train the second classifier (step 5), and this classifier provides the estimates for state and protester violence. Separately, in step 6, a third classifier assigns gender and age estimates to identified faces (Kärkkäinen and Joo 2019); the count of faces provides an estimate of protest size, and the demographic estimates generate control variables. These steps result in 40,764 tweets with

---

<sup>3</sup>The first two classifiers are in fact partially combined in implementation such that one integrated classifier can generate two sets of outputs, although they differ conceptually. This is called multitask learning (Girshick 2015). We discuss two classifiers separately because they are trained on different data and used in different steps.

new data about the demographics of individuals and severity and type of violence recorded in each image.

Figure 1 shows sample images of protest (fig. 1A), state (fig. 1B), and protester (fig. 1C) violence that the pipeline identifies. For an overview of CNNs, see appendix section S1. Section S2 provides additional detail on our pipeline and validation of the results. Section S3 shows manual validation from a team of research assistants specifically trained for this project.

## Data

We identify five protest periods from polities with diverse population, income, and institutional characteristics, mitigating the risk that subsequent findings arise from underlying similarities in the cases. These polities are Hong Kong, Pakistan, South Korea, Spain, and Venezuela. The primary criterion is to construct a sample from different types of regimes, which we measure with the Polity4 score. Hong Kong has a score of  $-4$ ; Venezuela, 4; Pakistan, 7; South Korea, 8; and Spain, 10. We also consulted the Varieties of Democracy data set to ensure these countries contain different media environments ( $e\_v2xme\_altinf\_5C$ ) and civil society freedom ( $e\_v2xcs\_ccsi\_5C$ ). With their varying amounts of freedom of assembly, the press, and civil society, studying these countries minimizes, although it does not eliminate, the possibility that results derive from case selection.

Table 2 details the cities included from these countries, the issues driving protest, and the frequency of protest images per city. For each period, we searched from one week before the first reported protest and one week after the last one. This process identifies 42,579,188 tweets containing 4,456,981 images. Keeping only tweets whose images generate a protest score of at least .849 results in 26,142 tweets with images.<sup>4</sup> These tweets are the data used for regressions.

We then aggregate tweets to their city of origin and the day they were created. Cities are kept for analysis when at least one-seventh of their days contain a protest image. Table 2 shows these 24 cities, which account for 6,303 protest images. (Most images do not have a location resolution more precise than the country.) These 6,303 protest images spread across 4,143 city-days. We treat missing dates as true zeroes, and a robustness check shows that this interpolation does not change results. Some of these images are duplicates, but later deduplication shows they do not affect inference.

Using these data introduces two ethical concerns. First, it is possible that minors are part of this study, as Twitter does not perform age verification for accounts and minors could appear in others' photos. Although many protests, such as Hong Kong's or the 2019 protests in Chile, feature prominent actors under age 18, we have only used faces from individuals estimated to be at least 20 years old. Second, protesters may not be as anonymous as they think. Although these data are observational and publicly available, individuals in photographs may not have consented to appear in those photographs. Authorities could

---

<sup>4</sup>This threshold is the value that maximizes recall with .85 precision.



monitor images shared on social media to identify people who protested; many already do (Purdy 2018).<sup>5</sup> To prevent the identification of individuals in our data, we have released only the aggregated city-day data.

### Operationalization

The dependent variable is  $\text{Log}_{10}(\text{Protest Size})_{i,t}$ , the logarithm of the sum of the number of faces in all protest photos from city  $i$  on day  $t$ . Because the resulting numbers are certainly lower than the true protest size (the largest protest in our data set contains 627 faces), five checks are performed to provide confidence that this operationalization actually measures protest dynamics. The estimates generated are consistent with others' and record actual events.

The first check manually validates the face counts per photo. To complement the manual validation from Amazon Mechanical Turk that figures A3–A5 (figs. A1–A22 are available online) show, we trained a team of three students to count faces in images and label whether they contain state or protester violence; we then compare their coding to our classifier's estimates. The number of faces the human coders identify closely matches the classifier's face count, and images humans label that contain state or protester violence receive much higher classifier estimates for those labels than those that do not. Appendix section S3 presents measures of intercoder reliability, and figure A10 shows this comparison.

Second, size estimates for large protests could be biased upward if the number of images and faces per image increases with protest size. Figure A11 shows no linear relationship between the size of a protest and the number of faces per photo. City-days with larger protests are therefore driven by the production of more protest images, not the sharing of crowded images. This result matches other work finding that counting faces in protest images generates accurate estimates of protest size variation (Sobolev et al. 2020).

Next, we consult other sources' estimates of protest size in Barcelona, Caracas, Seoul, and Hong Kong. For Caracas, we use crowd density estimates of protest images from Venezuelan newspapers (Rodríguez 2020); this methodology is used widely because it generates accurate estimates of large crowds without directly counting every participant (McPhail and McCarthy 2004).<sup>6</sup> For Barcelona and Hong Kong, we trained a team of undergraduates to follow the coding procedure and sources of Weidmann and Rod (2018). That approach did not generate enough size estimates for Seoul, so we use police and activist size reports provided by Wikipedia.<sup>7</sup> Table 3 shows this correlation for matched events and the results of two residual tests. Although higher correlations are preferable, the residual plots of figure A12 show that  $\text{Log}_{10}(\text{Protest Size})_{i,t}$  is not biased as a function of the protest size recorded in other sources. Therefore,  $\text{Log}_{10}(\text{Protest Size})_{i,t}$  is a noisy but consistent estimate of protest size available from other sources.<sup>8</sup>

<sup>5</sup>.At the same time, shared protest images can identify incriminating state behavior that would otherwise be denied (Lim 2013).

<sup>6</sup>.We also tried the Mass Mobilization in Autocracies Database (MMAD) data set (Weidmann and Rod 2018), but it recorded numeric estimates of protest size for only four events. Correlation with those is greater than .9.

<sup>7</sup>.See [https://ko.wikipedia.org/wiki/박근혜\\_대통령\\_퇴진\\_운동](https://ko.wikipedia.org/wiki/박근혜_대통령_퇴진_운동).

<sup>8</sup>.The other sources are not used for size estimates because they contain too many false negatives.

Fourth, this pipeline recovers a very large percentage of protests identified in other event data sets for these countries, as table 4 shows. The top row shows the number of city-days with protest observed using geolocated images for each city, region, or country. These records are compared to three leading event data sets; each cell includes the number of events that data set records and the percentage of those this methodology captures. These results are in line with what Zhang and Pan (2019) find for Sina Weibo in China: their protest detection pipeline finds 52% of the events the Integrated Crisis Early Warning System (ICEWS) does, 56% of the Global Database of Events, Language, and Tone's, and 88% of WiseNews's. Note as well that most events we record are not recorded in the other event data sets, in line with similar comparisons in Chile, China, and Venezuela (Steinert-Threlkeld and Joo 2020; Steinhardt and Goebel 2019).<sup>9</sup>

Finally, the temporal variation of protest size is consistent with actual events. Figure 2 shows this result: the correspondence is consistent and appears to be more than chance. We did not attempt to label every peak, but other work has found that social media records protests at least as well as other event data sets (Dowd et al. 2020).

The violence variables to test hypothesis 1 are Perceived State Violence $_{i,t-1}$  and its square. For hypothesis 2, we use Perceived Protester Violence $_{i,t-1}$ . These are the average of the classifier estimate for all protest images per city-day. Figure A13 shows a time series of the mean state violence recorded in protest images for the same four cities.

We describe the violence variables as “perceived” for three reasons. First, the true amount of violence is unknown because violence is a latent concept, not a physical entity, like temperature or pressure, directly measurable. Second, the images people share may be strategically chosen. This possible selection effect is true of any event data that rely on secondary sources, which is to say almost all event data. For discussion and analysis of bias that these measures may introduce, see appendix section S8. Third, the main analysis does not deduplicate images, meaning images that are shared often will have a greater impact on people's decision-making process than those only tweeted once. Deduplicating images to more closely approximate the “true” violence at events does not change results, as table 6 shows.

## Model

In addition to the operationalizations detailed in the previous section, we include six control variables. Two are demographic: the gender and age attributes of protesters. A society with greater gender equality is more likely to see nonviolent than violent action (Schaffenaar 2017), and the same is true at the movement level (Asal et al. 2013). The percentage of protesters who are male, Male Percent $_{i,t-1}$ , is therefore a variable for which we control. Since youth often spearhead mass protests and these effects are amplified when there are many of them, the percentage of participants ages 20–29, Young Adult Percent $_{i,t-1}$ , is a variable for which we control (González 2020; Urdal 2006).<sup>10</sup>

<sup>9</sup>Steinert-Threlkeld and Joo (2020) and Steinhardt and Goebel (2019) do provide exact comparisons but show that social media generates data on many more events than MMAD or ICEWS in Venezuela, about the same amount as ACLED in Chile, and much more than domestic and international organizations in China.

<sup>10</sup>The oldest age of legal adulthood (Japan) of which we are aware is 20, so we use it to be cautious.

We also generate two binary variables,  $Police_{i,t-1}$  and  $Fire_{i,t-1}$ . The police and fire variables are the sum of images containing a police officer or firefighter, respectively, based on the thresholds identified in table A2 (tables A1–A14 are available online). The variable  $Police_{i,t-1}$  provides another estimate of repression, while  $Fire_{i,t-1}$  approximates protester violence (fig. A14 shows that fire and protester violence have the highest city-day correlation of any two variables). Fifth,  $Tweets_{i,t-1}$  is the number of lagged protest images per country-day and controls for any protest information not directly measured, such as tactical knowledge about a protest (Little 2015). Sixth, we include a lagged dependent variable to account for autocorrelation as well as any regression to the mean. Table 5 provides descriptive statistics of these variables.

We build three models. The first uses only covariates that measure violence. The second focuses on the demographic control variables. The final model combine the two sets of variables. All independent variables are lagged one day. All models include city fixed effects and city-clustered standard errors. To guard against overfitting, we use fivefold cross-validation. Ordinary least squares is the estimator.

## RESULTS

Results match expectations: low amounts of state violence correlate with larger subsequent protests ( $p < .001$ ), although severe enough violence correlates with subsequently smaller protests ( $p < .001$ ). When protesters engage in violence, recorded subsequent protest size is smaller ( $p < .05$ ). In addition, protester violence has a much smaller coefficient than either state violence variable, with the largest magnitude occurring when states engage in high levels of violence. Table A5 shows these results.

Figure 3 shows marginal effects of state and protester violence from that model. From values of  $[0, .3)$ , state violence correlates with larger subsequent protest. At that amount of violence, protest size the next day is 137% higher than if there were no state violence. Moreover, state repression usually leads to larger protests: only 77 of 1,467 city-days of protest contain average state violence greater than .3. Increased protester violence always correlates with subsequently smaller protest. The change, however, is much smaller than for state violence: moving from no protester violence to its mean (.035) correlates with a 2% smaller protest, while the difference between state violence and its mean is an increase of just over 17%. A 1 standard deviation increase of state violence from 0 correlates with a 63% increase in protest size; a 1 standard deviation increase in protester violence from 0 correlates with a 12% smaller protest.

## ROBUSTNESS CHECKS

A series of robustness checks on state violence confirms its n-shaped relationship with subsequent protest size. First, images with more violence could contain fewer faces, causing the regression results to be driven by measurement problems. While lagging the independent variables mitigates this concern, figure 4 shows that no relationship appears to exist between the violence an image records and the number of faces contained therein. To the extent that one does, it is actually slightly positive, biasing against finding a negative relationship.

Second, the n-shaped relationship between state violence and the next day's protest size could be an artifact of fitting a parametric model with a square term. Figure 5 shows the results of tests demonstrating the persistence of this relationship. Whether fitting a local average of the relationship between Perceived State Violence $_{i,t-1}$  and  $\text{Log}_{10}(\text{Sum of Faces})_{i,t}$  or a spline with 50 knots, binning Perceived State Violence $_{i,t-1}$  into 10 evenly spaced groups, or regressing Perceived State Violence $_{i,t-1}$  on partial residuals, the n-shaped relationship between state violence and subsequent protest size holds.

Strategic behavior of protesters and state agents may endogenously cause the observed correlations. Protesters often seek international attention because it increases their legitimacy and raises the cost of repression (Bruns, Highfield, and Burgess 2013), and protesters are strategic about the language in which they make these appeals (Driscoll and Steinert-Threlkeld 2020; Metzger, Nagler, and Tucker 2015). These appeals may downplay protester violence and exaggerate the size of protests. To deter protester coordination, state actors will emphasize protester and state violence as well as downplay the size of crowds.

Table 6 shows that tests for this strategic behavior do not change inferences. Model 1 restricts analysis to those tweets in the country's lingua franca, as they are the most relevant for local actors. Model 2 drops tweets from bots (accounts controlled by computer code) since actors use them to strategically amplify messages.<sup>11</sup> Since strategic actors may emphasize particular protest features by repeatedly sharing photos, model 3 keeps only the first occurrence of an image.<sup>12</sup> Strategic actors are also more likely to have many followers (authority figures) or few (bots), so model 4 restricts analysis to only tweets from accounts within the 25th–75th percentile of their country's follower distribution. Model 5 focuses on attempts to manipulate estimates of protest size by making the dependent variable equal the logarithm of the number of unique accounts that share protest images. This new dependent variable is not affected by images chosen for the number of protesters they show or strategic actors tweeting frequently.

Across all models save the last, the results for state violence stay the same; model 5 also records an n-shaped relationship, but the coefficient on the squared term is too small to be statistically significant. (Table A8 shows that the raw number of users re-creates the n-shaped relationship.) Inferences about protester violence also do not change except for one model, model 3; its sign is still the same, and results weighted by the number of tweets per city-day (table A9) make the variable significant again.

Finally, table 7 shows that accounting for more complicated time dynamics does not change inferences about state violence, although it weakens the negative correlation of protester violence and subsequent protest size. A partial autocorrelation plot suggests 15 lags of the dependent variable, which model 1 includes. Model 2 includes weekend fixed effects, and model 3 includes weekday ones. Model 4 includes a control for the protest duration

<sup>11</sup>We submit every user to the Botometer service and remove tweets with a complete automation probability  $\geq .4$ , the threshold that produces the most accurate classification of bots (Varol et al. 2017). Table A13 shows that no more than 10.8% of tweets in any city are from bots.

<sup>12</sup>The data do not contain retweets because Twitter does not assign coordinates to retweets. They contain replies, and replies contain the image of the original tweet. Appendix sec. S11 explains the deduplication methodology, and table A14 shows the percentage of tweets per city that are duplicates.

and the number of consecutive days of state repression. In all specifications, the n-shaped relationship of state violence and subsequent protest remains statistically significant. Models 1 and 4 no longer find significant results for protester violence, suggesting that protesters may become more violent, the longer a protest lasts. Model 4 also shows that protests decrease in size over time, although persistent state violence correlates with larger protests.

Appendix section S8 presents two checks of the data-generating process to address concerns about selection bias. Users who share protest images may differ from those who share nonprotest images; figure A15 shows, and  $t$ -tests confirm, these users do not differ. Users may strategically choose the level of geographic specificity to assign to a tweet, depending on the tweet content. Figure A16 shows the distribution of estimates for number of faces and perceived state and protester violence by country and four levels of geographic specificity. Most tweets are with neighborhood or city specificity, and the classifier results do not systematically vary by geographic level within a country.

Appendix section S9 presents nine additional robustness checks. Section S9.1 performs the same nonparametric examinations of Perceived Protester Violence $_{i,t-1}$  that were performed on Perceived State Violence $_{i,t-1}$ . Section S9.2 presents the results from a vector autoregression, the results of which are consistent with table A5. Section S9.3 aggregates tweets to their state or country to see whether individuals selecting the tweet's geographic level biases results. Section S9.4 confirms that modeling days without protest images as days of no protest does not affect results. Section S9.5 shows that different transformations of the dependent variable do not change results. Section S9.6 weights city-days by number of tweets and their inverse; results do not change, suggesting these findings are not an artifact of Twitter prevalence in certain locales. Section S9.7 analyzes tweets most likely to originate at a protest: those from mobile phones or during protest hours. Section S9.8 shows results controlling for tweets about protest but without images. Section S9.9 disaggregates results by country. In all models, the protester and state violence variables exhibit the same relationships found in the other results, although protester violence's coefficient occasionally becomes close enough to zero to be indistinguishable from it.<sup>13</sup>

## DISCUSSION

The results presented here suggest that state violence generates an n-shaped correlation with subsequent protest size, while protester violence is consistently negative. While emphasizing the severity of repression during protest policing is not new (Khawaja 1993; Muller 1985), measuring nonlethal repression as a continuous variable is. This insight is not measurable without applying recent advances in computer vision to large data sets of individual behavior, in this case geolocated images shared on Twitter.

Synthesizing this article's results with others' suggests protest dynamics work as follows. Preventative repression, such as arresting a group's leaders or seizing their offices, makes it harder for protests to start (Danneman and Ritter 2013; Sullivan 2016; Sutton, Butcher, and Svensson 2014). Once started, protester violence decreases support for protest and therefore

<sup>13</sup>We also reran the main model but with  $\text{Log}(\text{Tweets})_{i,t-1}$ . The results do not change.

its subsequent size (Wasow 2020). Repression produces differential effects depending on its severity: light state violence generates backlash, while severe state violence “works.”

Future research should focus on several areas. It is possible that the time during a protest cycle when repression occurs correlates with its severity; future work should interact state violence with when in a protest cycle that violence occurs. Mechanisms connecting state and protester violence to changes in protest size also require investigation; emotions can be measured from social media text and images (Steinert-Threlkeld and Joo 2020), and expectations can be measured with longitudinal surveys (Cantoni et al. 2019). In addition, global event data sets could be constructed using this methodology. An advantage of using images is that they are closer to a universal language than text (Graber 1996). Classifiers generated from images are therefore less context dependent than those from text, so they can be applied across settings and time periods. For more detail on the advantages of images for generating protest event data, see Steinert-Threlkeld (2019).

Social media make valuable contributions to understanding subnational conflict. Driscoll and Steinert-Threlkeld (2020) and Metzger et al. (2015) use social media as a quasi poll to understand mass protests in Ukraine. Steinert-Threlkeld (2017) shows that coordinating activity correlates with more protests when that activity comes from less influential members of society. Larson et al. (2019) study 130,000,000 protesters in France, finding that protesters are embedded in denser social networks than bystanders. Muchlinski et al. (2021) measure electoral violence using CNNs and tweets. Social media, moreover, often generates more extensive records of protest than wire reports or newspapers, the primary source for existing event data sets (Steinert-Threlkeld and Joo 2020; Steinhardt and Goebel 2019). These data are not a fad or detached from reality: they offer unprecedented insights into human behavior and should continue to grow in importance to political science research.

If a picture is worth 1,000 words, then it would require approximately 2 kilobytes of storage (Jagenstedt 2008). Images from consumer cell phones and digital cameras, however, require at least 3 megabytes of storage, usually more. Even images shared on social media platforms, which are compressed from their original size, require hundreds of kilobytes of space. A picture, in other words, is worth anywhere from 50,000 (100 kilobytes) to 1,500,000 words (3 megabytes).<sup>14</sup> A picture is actually worth a book, and there are vast libraries waiting to be explored.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

We would like to thank the editor and three reviewers for their excellent and punctual feedback. Justin Jerro, Xiaofeng Lin, Jun Luo, Jack Schafer, Donghyeon Won, and Cindy Yuan provided essential research assistance. Alexei Abrahams, Konstantin Ash, Michael Chwe, Emilio Ferrara, Tim Groeling, Susan Hyde, Jason Jones, Andrew Little, Susanne Lohmann, Brandon Merrell, Tamar Mitts, Kevin O'Brien, Jennifer Pan, Margaret Roberts,

<sup>14</sup>This estimate is poetic. Another way to think of images is that they have high entropy, meaning they cannot be compressed as much as text. The greater size of images reflects this greater difficulty of compressing them, not necessarily a true quantum of information.

Jacob Shapiro, Joshua Tucker, Nils Weidmann, and Thomas Zeitoff have provided valuable feedback throughout. Feedback at the American Political Science Association, European Political Science Association, International Conference on Computational Social Science, International Studies Association, Peace Science, Politics and Computational Social Science, Political Methodology, and Southern Political Science Association also refined the article. Research costs money, and support from the National Science Foundation (SBE-SMA 1831848), California Center for Population Research, and internal University of California, Los Angeles, grants made this work possible. All remaining errors are unintentional.

## REFERENCES

- Asal Victor, Legault Richard, Szekely Ora, and Wilkenfeld Jonathan. 2013. "Gender Ideologies and Forms of Contentious Mobilization in the Middle East." *Journal of Peace Research* 50 (3): 305–18.
- Bashir Nadia Y., Lockwood Penelope, Chasteen Alison L., Nadolny Daniel, and Noyes Indra. 2013. "The Ironic Impact of Activists: Negative Stereo-types Reduce Social Change Influence." *European Journal of Social Psychology* 43 (7): 614–26.
- Battaglini Marco. 2017. "Public Protests and Policy Making." *Quarterly Journal of Economics* 132 (1): 485–549.
- Benford Robert D., and Snow David A. 2000. "Framing Processes and Social Movements: An Overview and Assessment." *Annual Review of Sociology* 26:611–39.
- Biggs Michael. 2003. "Positive Feedback in Collective Mobilization: The American Strike Wave of 1886." *Theory and Society* 32:217–54.
- Boschee Elizabeth, Lautenschlager Jennifer, Sean O'Brien Steve Shellman, Starz James, and Ward Michael. 2015. "ICEWS Coded Event Data." Harvard Dataverse. 10.7910/DVN/28075.
- Bradley Ralph Allan, and Terry Milton E. 1952. "Rank Analysis of Incomplete Block Designs." Pt. 1, "The Method of Paired Comparisons." *Biometrika* 39 (3/4): 324–45.
- Bratton Michael, and Van de Walle Nicolas. 1992. "Popular Protest and Political Reform in Africa." *Comparative Politics* 24 (4): 419–42.
- Bruns Axel., Highfield Tim, and Burgess Jean. 2013. "The Arab Spring and Social Media Audiences: English and Arabic Twitter Users and Their Networks." *American Behavioral Scientist* 57 (7): 871–98.
- Bueno de Mesquita Bruce, Smith Alastair, Siverson Randolph M., and Morrow James D. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.
- Cantoni Davide, Yang David Y., Yuchtman Noam, and Zhang Y. Jane. 2019. "Protests as Strategic Games: Experimental Evidence from Hong Kong's Anti-authoritarian Movement." *Quarterly Journal of Economics* 134 (2): 1021–77.
- Clark David H., and Regan Patrick M. 2016. "Mass Mobilization." Unpublished manuscript, Binghamton University. <https://massmobilization.github.io/about.html>.
- Danneman Nathan, and Ritter Emily H. 2013. "Contagious Rebellion and Preemptive Repression." *Journal of Conflict Resolution* 58 (2): 254–79.
- Davenport Christian. 2007. "State Repression and Political Order." *Annual Review of Political Science* 10 (1): 1–23.
- Davenport Christian, and Armstrong David A. II. 2004. "Democracy and the Violation of Human Rights: A Statistical Analysis from 1976 to 1996." *American Journal of Political Science* 48 (3): 538–54.
- DeNardo James. 1985. *Power in Numbers: The Political Strategy of Protest and Rebellion*. Princeton, NJ: Princeton University Press.
- Dowd Caitriona, Justino Patricia, Kishi Roudabeh, and Marchais Gauthier. 2020. "Comparing 'New' and 'Old' Media for Violence Monitoring and Crisis Response in Kenya." *Research and Politics* 7 (3): 1–9.
- Downs Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Driscoll Jesse, and Steinert-Threlkeld Zachary C. 2020. "Social Media and Russian Territorial Irredentism: Some Facts and a Conjecture." *Post-Soviet Affairs* 36 (2): 101–21.
- Earl Jennifer, Heather McKee Hurwitz Analicia Mejia Mesinas, Tolan Margaret, and Arlotti Ashley. 2013. "This Protest Will Be Tweeted: Twitter and Protest Policing during the Pittsburgh G20." *Information, Communication and Society* 16 (4): 459–78.

- Francisco Ronald A. 2004. "After the Massacre: Mobilization in the Wake of Harsh Repression." *Mobilization: An International Journal* 9 (2): 107–26.
- Girshick Ross. 2015. "Fast r-cnn." In *Proceedings of the IEEE International Conference on Computer Vision*. Los Alamitos, CA: IEEE Computer Society, 1440–48.
- Goldstein Joshua S. 1992. "A Conflict-Cooperation Scale for WEIS Events Data." *Journal of Conflict Resolution* 36 (2): 369–85.
- González Felipe. 2020. "Collective Action in Networks: Evidence from the Chilean Student Movement." *Journal of Public Economics* 188:104220.
- Graber Doris A. 1996. "Say It with Pictures." *Annals of the American Academy of Political and Social Science* 546 (1): 85–96.
- Gupta Dipak K., Singh Harinder, and Sprague Tom. 1993. "Government Coercion of Dissidents: Deterrence or Provocation?" *Journal of Conflict Resolution* 37 (2): 301–39.
- Gurr Ted Robert, and Moore Will H. 1997. "Ethnopolitical Rebellion: A Cross-Sectional Analysis of the 1980s with Risk Assessments for the 1990s." *American Journal of Political Science* 41 (4): 1079–103.
- Hess David, and Martin Brian. 2006. "Repression, Backfire, and the Theory of Transformative Events." *Mobilization: An International Journal* 11 (2): 249–67.
- Jagenstedt Philip. 2008. "How Much a Thousand Words Are Worth." *Preentious Nonsense*. <https://blog.foolip.org/2008/05/17/how-much-a-thousand-words-are-worth/>.
- Jasper James M. 2011. "Emotions and Social Movements: Twenty Years of Theory and Research." *Annual Review of Sociology* 37 (1): 285–303.
- Kärkkäinen Kimmo, and Joo Jungseock. 2019. "FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age." arXiv. <https://arxiv.org/pdf/1908.04913.pdf>.
- Khawaja Marwan. 1993. "Repression and Popular Collective Action: Evidence from the West Bank." *Sociological Forum* 8 (1): 47–71.
- Larson Jennifer M., Nagler Jonathan, Ronen Jonathan, and Tucker Joshua A. 2019. "Social Networks and Protest Participation: Evidence from 130 Million Twitter Users." *American Journal of Political Science* 63 (3): 690–705.
- Lawrence Adria K. 2017. "Repression and Activism among the Arab Spring's First Movers: Evidence from Morocco's February 20th Movement." *British Journal of Political Science* 47 (3): 699–718.
- Lim Merlyna. 2013. "Framing Bouazizi: 'White Lies,' Hybrid Network, and Collective/Connective Action in the 2010–11 Tunisian Uprising." *Journalism* 14 (7): 921–41.
- Little Andrew T. 2015. "Communication Technology and Protest." *Journal of Politics* 78 (1): 152–66.
- Lohmann Susanne. 1994. "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–91." *World Politics* 47 (1): 42–101.
- McPhail Clark, and McCarthy John. 2004. "Who Counts and How: Estimating the Size of Protests." *Contexts* 3 (3): 12–18.
- Metzger Megan, Nagler Jonathan, and Tucker Joshua A. 2015. "Tweeting Identity? Ukrainian, Russian, and #Euromaidan." *Journal of Comparative Economics* 44 (1): 16–40.
- Moore Will H. 2000. "The Repression of Dissent: A Substitution Model of Government Coercion." *Journal of Conflict Resolution* 44 (1): 107–27.
- Muchlinski David, Yang Xiao, Birch Sarah, MacDonald Craig, and Ounis Iadh. 2021. "We Need to Go Deeper: Measuring Electoral Violence using Convolutional Neural Networks and Social Media." *Political Science Research and Methods* 9 (1): 122–39.
- Muller Edward N. 1985. "Income Inequality, Regime Repressiveness, and Political Violence." *American Sociological Review* 50 (1): 47–61.
- Muñoz Jordi, and Anduiza Eva. 2019. "'If a Fight Starts, Watch the Crowd': The Effect of Violence on Popular Support for Social Movements." *Journal of Peace Research* 56 (4): 485–98.
- Murdie Amanda, and Purser Carolin. 2017. "How Protest Affects Opinions of Peaceful Demonstration and Expression Rights." *Journal of Human Rights* 16 (3): 351–69.
- Nordås Ragnhild, and Davenport Christian. 2013. "Fight the Youth: Youth Bulges and State Repression." *American Journal of Political Science* 57 (4): 926–40.



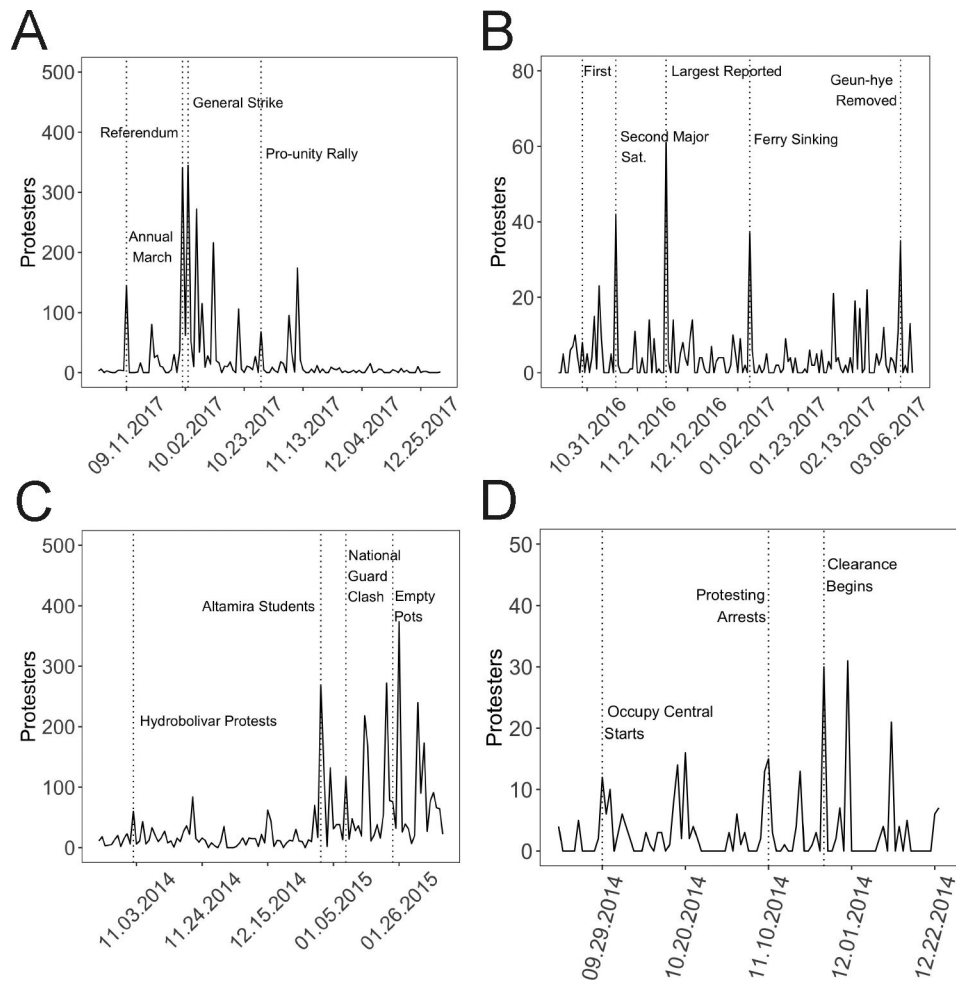
- O'Brien Kevin J., and Deng Yanhua. 2015. "Repression Backfires: Tactical Radicalization and Protest Spectacle in Rural China." *Journal of Contemporary China* 24 (93): 457–70.
- Olivier Johan L. 1991. "State Repression and Collective Action in South Africa, 1970–84." *South African Journal of Sociology* 22 (4): 109–17.
- Olzak Susan, Beasley Maya, and Olivier Johan. 2003. "The Impact of State Reforms on Protest against Apartheid in South Africa." *Mobilization* 8 (1): 27–50.
- Pearlman Wendy. 2013. "Emotions and the Microfoundations of the Arab Uprisings." *Perspectives on Politics* 11 (2): 387–409.
- Purdy Chase. 2018. "China Is Launching a Dystopian Program to Monitor Citizens in Beijing." *Quartz* <https://qz.com/1473966/china-is-starting-a-big-brother-monitoring-program-in-beijing/>.
- Qin Bei, Strömberg David, and Wu Yanhui. 2017. "Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda." *Journal of Economic Perspectives* 31 (1): 117–40.
- Raleigh Clionadh, Linke Andrew, Hegre Havard, and Karlsen Joakim. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset; Special Data Feature." *Journal of Peace Research* 47 (5): 651–60.
- Ritter Emily Hencken, and Conrad Courtenay R. 2016. "Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression." *American Political Science Review* 110 (1): 85–99.
- Robertson Graeme B. 2007. "Strikes and Labor Organization in Hybrid Regimes." *American Political Science Review* 101 (4): 781–98.
- Rodríguez Francisco. 2020. "Things Fall Apart: Nicolás Maduro and the Unraveling of Venezuela's Populist Experiment, 2012–2020." Unpublished manuscript.
- Salehyan Idean, Hendrix Cullen, Hammer Jesse, Case Christina, Linebarger Christopher, Stull Emily, and Williams Jennifer. 2012. "Social Conflict in Africa: A New Database." *International Interactions* 38 (4): 503–11.
- Schaftenaar Susanne. 2017. "How (Wo)men Rebel: Exploring the Effect of Gender Equality on Nonviolent and Armed Conflict Onset." *Journal of Peace Research* 54 (6): 762–76.
- Shadmehr Mehdi, and Bernhardt Dan. 2011. "Collective Action with Uncertain Payoffs: Coordination, Public Signals, and Punishment Dilemmas." *American Political Science Review* 105 (4): 829–51.
- Shakir Omar. 2014. "All According to Plan: The Rab'a Massacre and Mass Killings of Protesters in Egypt." Technical report, Human Rights Watch. <https://www.hrw.org/report/2014/08/12/all-according-plan/raba-massacre-and-mass-killings-protesters-egypt>.
- Siegel David A. 2011. "When Does Repression Work? Collective Action in Social Networks." *Journal of Politics* 73 (4): 993–1010.
- Simpson Brent, Willer Robb, and Feinberg Matthew. 2018. "Does Violent Protest Backfire? Testing a Theory of Public Reactions to Activist Violence." *Socius: Sociological Research for a Dynamic World* 4:23780231188 0318.
- Sobolev Anton, Joo Jungseock, Chen Keith, and Steinert-Threlkeld Zachary C. 2020. "News and Geolocated Social Media Accurately Measure Protest Size Variation." *American Political Science Review* 114 (4): 1343–51.
- Steinert-Threlkeld Zachary C. 2017. "Spontaneous Collective Action: Peripheral Mobilization during the Arab Spring." *American Political Science Review* 111 (2): 379–403.
- Steinert-Threlkeld Zachary C. 2019. "Comment: The Future of Event Data Is Images." *Sociological Methodology* 49 (1): 68–75.
- Steinert-Threlkeld Zachary C., and Joo Jungseock. 2020. "Protest Event Data from Geolocated Social Media Content." APSA Preprints. 10.33774/apsa-2020-mjz4s.
- Steinhardt H. Christoph, and Goebel Christian. 2019. "Better Coverage, Less Bias: Using Social Media to Measure Protest in Authoritarian Regimes." Presented at the Conference on New Forms of Evidence for the Study of Contention in China, Department of East Asian Studies, University of Vienna, May 8.
- Stephan Maria J., and Chenoweth Erica. 2008. "Why Civil Resistance Works." *International Security* 33 (1): 7–44.

- Sullivan Christopher M. 2016. "Political Repression and the Destruction of Dissident Organizations." *World Politics* 68 (4): 645–76.
- Sutton Jonathan, Butcher Charles R., and Svensson Isak. 2014. "Explaining Political Jiu-Jitsu: Institution-Building and the Outcomes of Regime Violence against Unarmed Protests." *Journal of Peace Research* 51 (5): 559–73.
- Tucker Joshua A. 2007. "Enough! Electoral Fraud, Collective Action Problems, and Post-Communist Colored Revolutions." *Perspectives on Politics* 5 (3): 535–51.
- Tullock Gordon. 1971. "The Paradox of Revolution." *Public Choice* 11:89–99.
- Urdal Henrik. 2006. "A Clash of Generations? Youth Bulges and Political Violence." *International Studies Quarterly* 50:607–29.
- Urdal Henrik, and Hoelscher Kristian. 2012. "Explaining Urban Social Disorder and Violence: An Empirical Study of Event Data from Asian and Sub-Saharan African Cities." *International Interactions* 38 (4): 512–28.
- Varol Onur, Ferrara Emilio, Davis Clayton A., Menczer Filippo, and Flammini Alessandro. 2017. "Online Human-Bot Interactions: Detection, Estimation, and Characterization." arXiv. <https://www.indosmedia.com/wp-content/uploads/2017/06/1703.03107.pdf>.
- Wasow Omar. 2020. "Agenda Seeding: How 1960s Black Protests Moved Elites, Public Opinion and Voting." *American Political Science Review* 114 (3): 638–59.
- Weidmann Nils B., and Rod Espen Geelmuyden. 2018. "Coding Protest Events in Autocracies." In *The Internet and Political Protest in Autocracies*. Oxford: Oxford University Press, chap. 4.
- Wouters Ruud, and Walgrave Stefaan. 2017. "Demonstrating Power: How Protest Persuades Political Representatives." *American Sociological Review* 82 (2): 361–83.
- Young Lauren E. 2019. "The Psychology of State Repression: Fear and Dissent Decisions in Zimbabwe." *American Political Science Review* 113 (1): 140–55.
- Zhang Han, and Pan Jennifer. 2019. "CASM: A Deep-Learning Approach for Identifying Collective Action Events with Text and Image Data from Social Media." *Sociological Methodology* 49:1–48.
- Zhao Dingxin. 1998. "Ecologies of Social Movements: Student Mobilization during the 1989 Prodemocracy Movement in Beijing." *American Journal of Sociology* 103 (6): 1493–529.

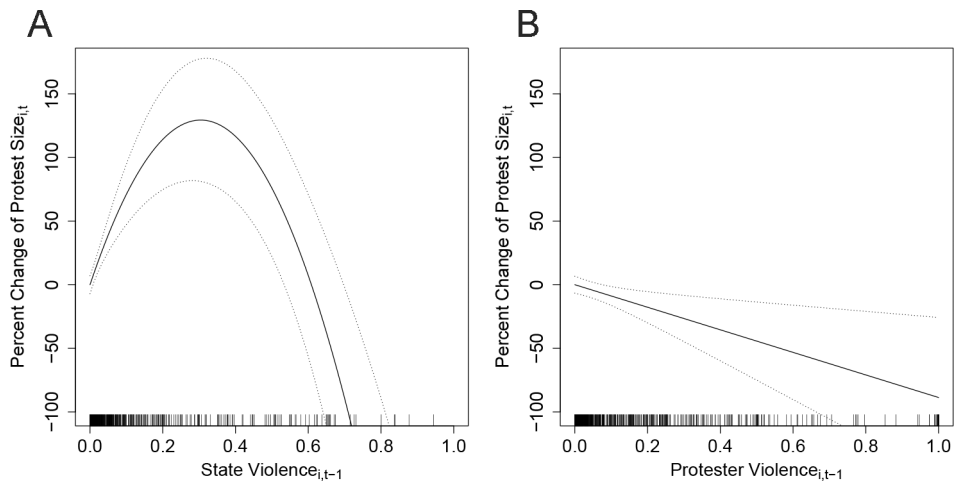


**Figure 1.**

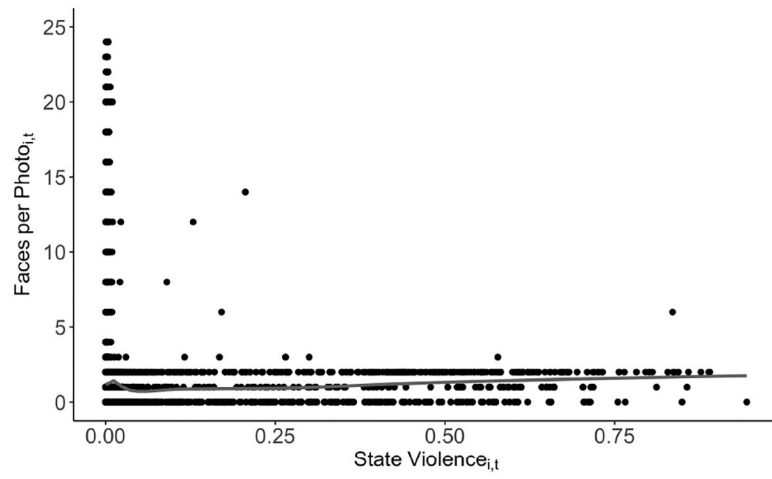
*A*, Sample images and the protest classifier's rating of them. *B*, Protest images with their state violence rating. *C*, Protest images' protester violence rating. Labels specify each image's city and label probability. The use of hard negatives in the training set ensures that scenes that contain crowds (*A*, *left*), individuals walking on streets (*C*, *left*), or a nonprotest sign (*A*, *second*) are not included in analysis. Color version available as an online enhancement.



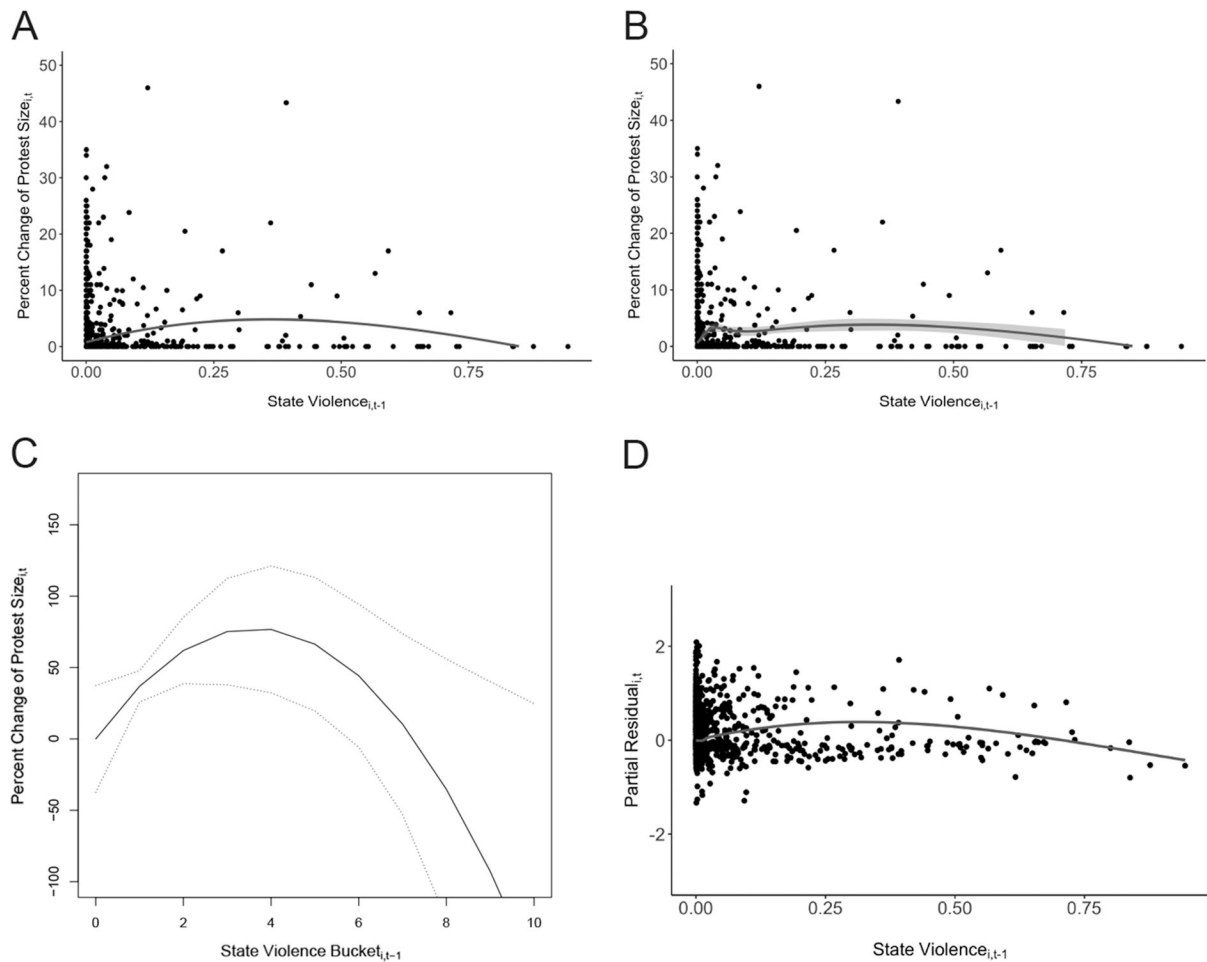
**Figure 2.** Verifying protest size, time series: *A*, Barcelona; *B*, Seoul; *C*, Caracas; *D*, Hong Kong. Measuring protest size using faces in images records changes in protest size that match reported dynamics.



**Figure 3.** Marginal effects of Perceived State Violence<sub>*i,t-1*</sub> (A) and Perceived Protester Violence<sub>*i,t-1*</sub> (B) from model 3 of table A5. State violence exhibits an n-shaped relationship with subsequent protest size, while protester violence always correlates with smaller subsequent protests.



**Figure 4.** No correlation exists between state violence in an image and the number of faces. A linear fit suggests a slight positive relationship, and restricting the relationship to images with fewer than 10 faces does not change results.



**Figure 5.** State violence results remain in flexible operationalizations: *A*, LOESS (span = .2); *B*, spline (50 knots); *C*, binned marginal effects; *D*, LOESS on partial residuals. The n-shaped relationship holds in a nonparametric relationship (*A*, *B*). Generating fixed effects for state violence in bins of width .1 finds the same relationship (*C*). *D*, Regressing state violence on the partial residuals of model 3 from table A5.

## Protest Data Pipeline

Table 1.

Step	Input	Source	Output
Collecting images for training set:			
1. Image search	Keywords	Google	100,000 images
2. Train a protest image classifier	Images from step 1	Self	Initial CNN
3. Protest images from Twitter corpus	Model from step 2	Twitter	40,764 images
Developing protest and scene classifier:			
4. Manual annotation	Images from step 3	Amazon Mechanical Turk	13 ground-truth labels
5. Train a CNN	Training data from step 4	Self	Protest and scene classifier
Face attribute classification:			
6. Face classification	...	Karkkainen and Joo (2019)	Gender, age, and size estimates

Note. The protest data pipeline encompasses six primary steps. Appendix sec. S2 provides more detail. CNN = convolutional neural network.



Table 2.

## Protest Periods

City	Country	Start	End	Issue	Protest Images/Day	Protest Images/Day if >0
Central	Hong Kong	2014.09.18	2014.12.23	China reforms	1.96	5.00
Kowloon	Hong Kong	2014.09.18	2014.12.23	China reforms	1.29	2.92
Lahore	Pakistan	2017.11.07	2017.11.23	Blasphemy	.18	1
Kimhae	South Korea	2016.10.20	2017.03.14	Anti-incumbency	.47	1.92
Seoul	South Korea	2016.10.20	2017.03.14	Anti-incumbency	2.40	3.76
Citutat Vella	Spain	2017.09.01	2017.12.31	Secession	.94	4.95
Barcelona	Spain	2017.09.01	2017.12.31	Secession	3.07	11.60
Girona	Spain	2017.09.01	2017.12.31	Secession	1.10	3.26
Granera	Spain	2017.09.01	2017.12.31	Secession	.62	2.33
Granollers	Spain	2017.09.01	2017.12.31	Secession	.23	1.25
Lleida	Spain	2017.09.01	2017.12.31	Secession	.42	1.88
Mataro	Spain	2017.09.01	2017.12.31	Secession	.51	2.33
Reus	Spain	2017.09.01	2017.12.31	Secession	.35	1.68
Sabadell	Spain	2017.09.01	2017.12.31	Secession	.96	2.66
St. Cugat d. Valles	Spain	2017.09.01	2017.12.31	Secession	.31	2.06
St. Felin d. Pallerols	Spain	2017.09.01	2017.12.31	Secession	.61	2.19
St. Salvador d. Guardiola	Spain	2017.09.01	2017.12.31	Secession	.48	2.15
Tarragona	Spain	2017.09.01	2017.12.31	Secession	.57	1.94
Terrassa	Spain	2017.09.01	2017.12.31	Secession	.57	2.22
Boca del Río	Venezuela	2014.03.27	2017.12.17	Anti-Maduro	.26	1.34
Caracas	Venezuela	2014.03.27	2017.12.17	Anti-Maduro	4.82	7.63
Caucagua	Venezuela	2014.03.27	2017.12.17	Anti-Maduro	.53	1.72
Maracaibo	Venezuela	2014.03.27	2017.12.17	Anti-Maduro	.39	1.49
Valencia	Venezuela	2014.03.27	2017.12.17	Anti-Maduro	.41	1.62

Note. The last column is the average number of photos for days containing a protest photo.

**Table 3.**

## Verifying Protest Size Estimate Using Other Sources

City	Source	Matched Events	Correlation	S-W	K-S
Barcelona	AFP, BBC, AP	5	.7980	.0397	.4714
Caracas	Rodríguez (2020)	18	.4101	.9598	.5975
Hong Kong	AFP, BBC, AP	11	.3477	.8135	.7967
Seoul	Wikipedia: Police	12	.3686	.7294	.8785
Seoul	Wikipedia: Activists	21	.4349	.0491	.7121

Note. Taking the log of the sum of faces in protest photos correlates with logged estimates from newspapers (Barcelona, Hong Kong), Wikipedia (Seoul), and crowd density estimates (Caracas). The S-W column shows the  $p$ -value from the Shapiro-Wilks test, and the K-S column is for the Kolmogorov-Smirnov test; both are conducted on the residuals from regressing  $\text{Log}10(\text{Protest Size})_{i,t}$  on the log of the reported protest size. Lahore is not shown because we found no newspaper estimates of protest size from there. AFP = Agence France-Presse; BBC = British Broadcasting Corporation; AP = Associated Press.

**Table 4.**

## Verifying Protest Coverage Using Other Event Data

Source	Data	Catalonia	Hong Kong	Lahore, Pakistan	South Korea	Venezuela
Twitter	Images	1,421	137	17	232	2,336
ACLEd	Local and international news; reports	...	...	9, 33%	...	...
ICEWS	Local and international news	54, 94%	49, 59%	2, 0%	99, 82%	365, 49%
MMAD	AFP, BBC, AP	...	69, 91%	...	...	4, 100%

Note. The top row is the number of events in the Twitter data used for regressions. Each subsequent row shows the number of events in data set and the percentage match with Twitter; e.g., ICEWS (Integrated Crisis Early Warning System) records 54 events in Catalonia, 94% of which the Twitter data contain. All ICEWS events with a city of "nan" are dropped. MMAD contains more events from Hong Kong here than in table 3 because this table does not restrict events to those with size estimates. ACLEd = Armed Conflict Location and Event Data; MMAD = Mass Mobilization in Autocracies Database; AFP = Agence France-Presse; BBC = British Broadcasting Corporation; AP = Associated Press.

Table 5.

## Summary Statistics for Regression Variables

Statistic	N	Mean	SD	Min	Max
Protest Size <sub><i>it</i></sub>	4,143	2.57	15.36	.00	627.00
Perceived Protester Violence <sub><i>it,t-1</i></sub>	4,121	.03	.12	.00	1.00
Perceived State Violence <sub><i>it,t-1</i></sub>	4,121	.02	.08	.00	.94
Police <sub><i>it,t-1</i></sub>	4,121	.001	.04	.00	1.00
Fire <sub><i>it,t-1</i></sub>	4,121	.08	.42	.00	7.00
Male Percent <sub><i>it,t-1</i></sub>	4,121	.03	.11	.00	1.00
Young Adult Percent <sub><i>it,t-1</i></sub>	4,121	.02	.08	.00	1.00
Tweets <sub><i>it,t-1</i></sub>	4,121	1.52	7.02	.00	238.00

Table 6.

## Results Robust to Strategic Behaviors

	Country Language (1)	No Bots (2)	Duplicated Images (3)	IQR Users (4)	DV: $\text{Log}_{10}(\text{Number of Users})_{i,t}$ (5)
Perceived Protester Violence $_{i,t-1}$	-.1982*** (.0387)	-.1438** (.0704)	-.1301 (.0853)	-.1530*** (.0446)	-.1784*** (.0495)
Perceived State Violence $_{i,t-1}$	1.5665*** (.4553)	1.2236*** (.3349)	1.2138** (.5284)	1.1894*** (.3333)	.3938* (.2340)
Perceived State Violence $^2_{i,t-1}$	-2.5718*** (.8864)	-2.0184*** (.5787)	-2.0015** (.9100)	-2.0125*** (.5509)	-.6597 (.4044)
Police $_{i,t-1}$	.5070 (.3090)	.6357* (.3712)	.9419* (.5190)	.8552** (.3763)	.2250 (.2720)
Fire $_{i,t-1}$	.0755*** (.0249)	.0912*** (.0202)	.0594 (.0507)	.0576* (.0316)	.0667*** (.0152)
Male Percent $_{i,t-1}$	-.0617 (.1264)	-.1848* (.1078)	-.0833 (.1201)	-.0195 (.0458)	-.1068** (.0543)
Young Adult Percent $_{i,t-1}$	.0173 (.0845)	.2166** (.0905)	.2848** (.1370)	.2584*** (.0953)	.1162** (.0466)
Tweets $_{i,t-1}$	.0249*** (.0047)	.0119** (.0047)	.0162** (.0069)	.0148** (.0060)	.0037* (.0021)
DV $_{i,t-1}$	.1145 (.0772)	.1766** (.0782)	.1490* (.0844)	.1047*** (.0363)	.4062*** (.1004)
Intercept	.0694*** (.0138)	.1186*** (.0157)	.1319*** (.0172)	.1270*** (.0160)	.1027*** (.0105)
N	3,481	4,121	2,533	3,462	4,121
Adjusted $R^2$	.3144	.2732	.2223	.1897	.4125
Cluster SE	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes

Note. Model 1 keeps only tweets in each country's lingua franca. Model 2 drops all tweets from accounts identified as bots. Model 3 removes duplicate images. Model 4 keeps only tweets from users within the interquartile range (IQR) of their country's follower distribution. Model 5's dependent variable (DV) is the logarithm transformation of the number of users tweeting protest images. Standard errors (in parentheses) are clustered by city. FE = fixed effects.

\*  $p < .1$ .

\*\*  $p < .05$ .

$p < .10$   
\*\*\*

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 7.**

Time Effects Do Not Change Results

	15 Lags (1)	Weekend FE (2)	Weekday FE (3)	Duration Controls (4)
Perceived Protester Violence <sub><i>i,t-1</i></sub>	-.0366 (.0831)	-.1543* (.0830)	-.1571* (.0831)	-.0489 (.0775)
Perceived State Violence <sub><i>i,t-1</i></sub>	.7348** (.3226)	1.2716*** (.3722)	1.2864*** (.3727)	.8831*** (.3283)
Perceived State Violence <sup>2</sup> <sub><i>i,t-1</i></sub>	-1.2775** (.5291)	-2.0853*** (.6515)	-2.1055*** (.6528)	-1.5302*** (.5435)
Police <sub><i>i,t-1</i></sub>	.6381* (.3864)	.7512 (.4584)	.7553* (.4570)	.5867* (.3391)
Fire <sub><i>i,t-1</i></sub>	.0229 (.0334)	.0995*** (.0360)	.0998*** (.0360)	.0127 (.0313)
Male Percent <sub><i>i,t-1</i></sub>	-1.1952*** (.0735)	-1.1805*** (.0734)	-1.1789*** (.0731)	-1.1800*** (.0674)
Young Adult Percent <sub><i>i,t-1</i></sub>	.2127* (.1093)	.1930* (.1030)	.1927* (.1032)	.1575 (.1009)
Tweets <sub><i>i,t-1</i></sub>	.0059*** (.0021)	.0088*** (.0030)	.0087*** (.0031)	.0055*** (.0020)
DV <sub><i>i,t-1</i></sub>	.0864*** (.0308)	.1998*** (.0352)	.2002*** (.0352)	.1141*** (.0294)
Protest Days <sub><i>i,t</i></sub>				-.0006*** (.0001)
Consec. State Violence <sub><i>i,t</i></sub>				.0638*** (.0064)
Intercept	.0278** (.0130)	.1161*** (.0184)	.1260*** (.0218)	.2655*** (.0271)
<i>N</i>	3,777	4,121	4,121	4,121
Adjusted <i>R</i> <sup>2</sup>	.3489	.2684	.2686	.3505
Cluster SE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes

Note. Dependent variable (DV): Log [0(Sum of Faces)]<sub>*i,t*</sub>. The n-shaped relationship of state violence and subsequent protest size holds when controlling for time dynamics. Protest Days<sub>*i,t*</sub> refers to the length of a protest, and Consec. State Violence<sub>*i,t*</sub> the number of consecutive days of state violence. FE = fixed effects.

$p < .01$   
\*\*\*  
 $p < .05$   
\*\*  
 $p < .1$   
\*

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript