Proceedings of the Sixteenth International AAAI Conference on Web and Social Media (ICWSM 2022)

MMCHIVED: Multimodal Chile and Venezuela Protest Event Data

Zachary Steinert-Threlkeld* **and Jungseock Joo**

University of California, Los Angeles

Abstract

This paper introduces the Multimodal Chile & Venezuela Protest Event Dataset (MMCHIVED). MMCHIVED contains city-day event data using a new source of data, text and images shared on social media. These data enables the improved measurement of theoretically important variables such as protest size, protester and state violence, protester demographics, and emotions. In Venezuela, MMCHIVED records many more protests than existing datasets. In Chile, it records slightly more events than the Armed Conflict Location and Events Dataset (ACLED). These extra events are from small cities far from Caracas and Santiago, an improvement of coverage over datasets that rely on newspapers, and the paper confirms they are true positives. While MMCHIVED covers protest events in Chile and Venezuela, the approach used in the paper is generalizable and could generate protest event data in 107 countries containing 97.14% of global GDP and 82.7% of the world's population.

Introduction

At the end of 2013, Nicolas Maduro, Hugh Chávez's Vice-President, narrowly won Venezuela's presidential election. Protests over security started in 2014 in response to the murder of former Miss Venezuela Mónica Spear, and they quickly expanded in scope and intensity; one non-profit organization recorded 9,286 that year (Patilla 2015). Protests continued for the next several years, and recent contention between Juan Guaidó, the majority leader of the National Assembly, and President Maduro stem from these events. In Chile on October 6, 2019, the Panel of Public Transport Experts announced a moderate fare increase for buses and subways. The next day, secondary and tertiary students led fare evasion protests across the capital, and this defiance triggered nationwide protests against inequality. One year later, 78% of Chileans voted to replace the Pinochet Constitution with a new one to be written by a 155 member assembly.

This paper introduces the Multimodal Chile and Venezuela Event Dataset, MMCHIVED. The key contribution of this dataset is to take advantage of geolocated text and images shared on Twitter. Social media data are particularly useful for the study of protests for two reasons. First, since almost anyone can publish content, they often record more protests than other sources of raw data (Zhang and Pan 2019). Second, their documentary use facilitates the measurement of several theoretically important variables in ways with which other approaches struggle. These variables include protest size, continuous measures of protester and state violence, protester demographics, and emotions. These advances allow social scientists studying contentious politics to test theories whose concepts have eluded precise measurement at scale, such as sex or violence. Results from these methods suggest, for example, that violence and protester demographics are more powerful predictors of protest dynamics than emotions.

The dataset builds on existing event data in several ways. First, its methodology still relies on general knowledge generated from mass media to know where and when to search for protests. This approach is easier than the *sui generis* detection facing datasets such as Mass Mobilization in Autocracies Database (MMAD) and Armed Conflict Location and Event Data (ACLED) which parse media to extract protest events (Croicu and Weidmann 2015). Second, other researchers have used images to generate protest data on emotions and violence (Won, Steinert-Threlkeld, and Joo 2017; Steinert-Threlkeld, Chan, and Joo 2022). Images can complement information which may be missing in text and improve the reliability and precision of event detection and description (Steinert-Threlkeld, Chan, and Joo 2022). MMCHIVED builds on these approaches by generating Chile protest data, generating emotion data from text, recording protester demographic data, and combining all of these variables in one dataset. Third, this paper evaluates MMCHIVED by comparing it to MMAD, ACLED, the Integrated Conflict Early Warning System (ICEWS), and the Temporally Extended, Regular, Reproducible International Event Records (TERRIER) (Boschee et al. 2015; Liang et al. 2018) MMCHIVED therefore contributes to a growing literature comparing event data from social media to event data from tradtional media (Steinhardt and Gobel 2019; Dowd et al. 2020). While MMCHIVED covers protest events in Chile and Venezuela, the approach used in the paper is generalizable and could generate protest event data in 107 countries containing 97.14% of global GDP and 82.7% of the world's population (see Appendix, Fig. 13.)

*zst@luskin.ucla.edu. The dataset can be accessed via Dataverse (Steinert-Threlkeld and Joo 2021).

Why Multimodal Protest Event Data?

"Multimodal event data" refers to the generation of datasets of politically relevant actions using more than one type of data, where type means method of communication. For example, using text from a newspaper and text from a website is unimodal because text is the one type of data. Combining audio records with their speakers' covariates to understand Supreme Court decisions is an example of multimodal data (Dietrich, Enos, and Sen 2019). MMCHIVED is multimodal because it uses text and images to generate its protest records. Multimodal data represent a promising avenue for the generation of protest data because different types of data can generate new or improved measures of theoretically relevant variables.

Social media present particularly rich sets of multimodal data. While newspapers contain text and images, their incentives and space constraints make it very difficult to generate multimodal protest data. Though it is much easier to create a social media account than become a journalist, each social media account resembles a journalist because it can document events. Since there are many more social media accounts than journalists, space is unlimited, and many accounts "publish" frequently, social media generate orders of magnitude more "stories" than newspapers. There is therefore the potential to observe more events using social media than with other sources (Steinhardt and Gobel 2019). Moreover, these stories — tweets — always contain text and often contain images, making them natural candidates for multimodal event data. We focus on Twitter because it is a global platform that many researchers have productively used to study contentious politics, and its data are easier to obtain than competitors such as Facebook or Instagram.

Images can provide improved estimates of protest size, state and protester violence, and protester demographics. As more people protest, policy change becomes more likely (Wouters and Walgrave 2017). Crowd size estimation is its own field of study with a well-developed methodology (McPhail and McCarthy 2004), but this type of estimation rarely is used for event datasets because it is expensive and requires foreknowledge of protests. Most event datasets therefore rely on media reports of size; these estimates are usually words such as "hundreds", "thousands", or as round numbers, making them difficult to use in statistical analysis. When an image contains faces, the sum of face counts from protest images generates accurate estimates of protest size; cell phone location data validates this measurement strategy (Sobolev et al. 2020).[1]

Images can also generate estimates of protester and state violence in continuous values, and these precise measures can make important contributions to the repression-dissent literature. Many studies find no correlation between repression and subsequent dissent (Ritter and Conrad 2016). Others find a positive correlation (Francisco 2004), a negative one (Ferrara 2003), or argue that repression's effects depend on the time period analyzed (Rasler 1996). Part of the mixed results may be due to the difficulty of measuring violence from text sources, which causes existing datasets to report violence as an ordinal, usually binary, variable. Computer vision models, on the other hand, can generate continuous value estimates of violence (Chen et al. 2016; Won, Steinert-Threlkeld, and Joo 2017), and these more precise estimates may untangle the contradictory results of previous research.

Protester demographics may affect protest dynamics, but measures of protester demographics do not exist in panel data (Oliver, Marwell, and Teixeira 1985). Youth have their own grievances and have greater proclivity to protest (Nordås and Davenport 2013). Gender within movement leadership affects outcomes (Robnett 1996), and gender is a frequent question in surveys of protesters (Kern 2011). Race is a common mobilizing feature across countries (Scarritt, McMillan, and Mozaffar 2001). Some event datasets indicate whether a social movement includes individuals from various backgrounds as nominal variables at the movement, not protest, level (Chenoweth and Stephan 2011). The Social Conflict Analysis Database (SCAD) codes whether an event is focused on female or LGBTQ issues but not the gender or sexual orientation of event participants (Salehyan et al. 2012). The Mass Mobilization in Autocracies Database records actor attributes as a string variable *at the report level*, but the event level version of the dataset omits that variable (Weidmann and Rød 2019). Projects that code protester attributes from survey data rarely include longitudinal observations and are not part of larger event dataset efforts (Fisher, Dow, and Ray 2017). Generating age, gender, and race estimates of protesters using faces in protest images therefore can advance several lines of scholarship around demographics and protest.

How emotions affect individuals' decision to protest is a growing area of research, but they have eluded incorporation into event data (Jasper 2011). They mobilize bystanders and mediate the effect of repression, depending on which are triggered (Pearlman 2013). Because of the difficulty of measuring emotional states in real time, research in this area has been theoretical (Pearlman 2013) and experimental (Young 2019) but not observational. Though some work has estimated the emotional content of images (Won, Steinert-Threlkeld, and Joo 2017), MMCHIVED uses text because many theoretically relevant emotions, such as pride or shame, are less apparent in images. In addition to text itself, emojis (expressive digital faces) are transmitted as Unicode text, so they can be treated as text and used to measure emotions.

Constructing the Dataset

MMCHIVED extends the methodology developed in (Won, Steinert-Threlkeld, and Joo 2017) and (Steinert-Threlkeld, Chan, and Joo 2022). It builds on the former by generating separate estimates of protester and state violence and creating a protest size variable. It builds on the latter by measuring emotions from text. It builds on both by measuring protester demographics, emotions, and analyzing protests in Chile. Since those two papers provide technical details and evaluation of the pipeline that generates much of MMCHIVED's data, this section focuses on issues unique to MMCHIVED.

Figure 1 shows sample images with their classifier estimate. The images for protest and protester violence are from

[1]That paper also shows that counting tweets with protest keywords generates accurate estimates of protest size.

Chile; those for state violence, from Venezuela. Each image is labeled with its city and classifier estimate.

MMCHIVED identifies protest images and assigns protester and state violence scores using the methodology of (Won, Steinert-Threlkeld, and Joo 2017) and (Steinert-Threlkeld, Chan, and Joo 2022). Briefly, we first searched Google Images using keywords related to protest such as “protest” or “BLM” and negative one such as “concert” to collect a large number of negative images that may look similar to a protest scene. A CNN is trained on these data and applied to tweets, generating 40,764 images for workers on Amazon Mechanical Turk (AMT) to code. Their labels are then given to a ResNet-50 CNN (He et al. 2016). The resulting trained model is the protest image classifier MMCHIVED uses to identify protest images. To estimate protester and state violence, different AMT workers generate pairwise annotations, and the Bradley-Terry model uses those annotations to generate continuous estimates of violence (Bradley and Terry 1952).

Analyzing faces in protest images enables the estimation of protester demographics; as far as we are aware, MMCHIVED is the first dataset to estimate protester demographics from images. A publicly available model, FairFace, is applied to the identified faces in order to generate age, gender, and race estimates (Karkkainen and Joo 2021). FairFace is the only model trained from a racially balanced dataset comprised of amateur images, and it substantially improves classification accuracy, especially for the non-White populations.

Text and off-the-shelf software libraries estimate the emotional content of tweet text. After standard preprocessing steps, emojis are converted to the words they evoke using R’s `emo` package. With this additional emoji information, R’s `syuzhet` package then identifies whether a tweet contains any of eight emotions.

The end result of the pipeline is a dataset of tweets containing protest images, with the metadata in Table 1 added to each tweet. If a tweet only contains text, it is not kept. Tweets are not filtered for the presence of bots since previous research has found little evidence of bots in geolocated tweets (Samper-Escalante et al. 2021).

If a duplicated image exists, only the first appearance is kept before aggregation. Deduplication prevents false positive events when the same image is shared across multiple days or cities. Deduplication can also be thought of as removing a potential violation of SUTVA. 43.54% of images from Venezuela are duplicates, as are 18.79% from Chile. Figure 9 shows the distribution of duplication rate by city, and manual investigation of a sample of images confirms the removed images are duplicates.

The final step is aggregation by city-day using the operation in Table 1’s aggregation column. MMCHIVED assumes that the date of a protest is the date the image was tweeted, and it uses the city that Twitter assigns.[2] Protest size is the sum of face counts in protest images per city-day, which accurately measures protest size variation (Sobolev et al. 2020). Aggregation allows MMCHIVED to use multiple observations of heterogeneous quality to estimate an event variable’s unknowable true value, similar to aggregating individual news articles to generate event data (Cook and Weidmann 2019).[3]

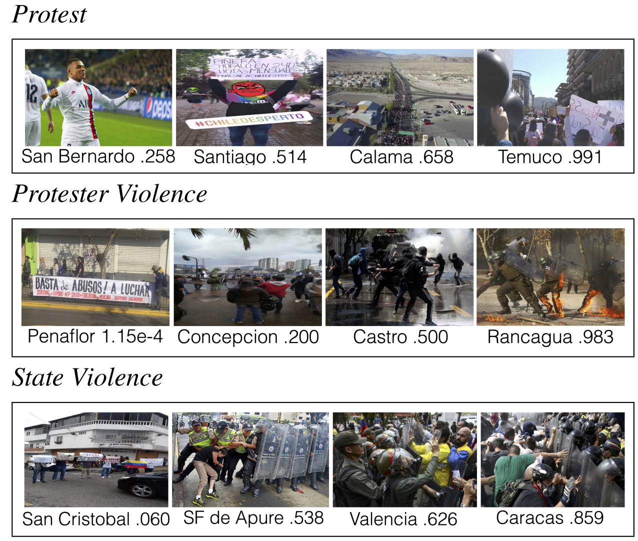


Figure 1: Example Images and Classifier Annotation [0-1]

Since the pipeline is applied to specific countries and periods during which protest is known to have occurred, identifying protest images for MMCHIVED is conceptually simpler than the task facing event datasets without *a priori* knowledge of protests. For example, MMAD and ACLED parse news reports without knowing if a country experienced protest on a given day. The equivalent for MMCHIVED would be if all tweet images were downloaded and analyzed regardless of country or day. This equivalent requires more computation infrastructure than available to us, though it is technically feasible.

The Dataset

Since Twitter, and social media more broadly, is popular in both countries, the generation of multimodal event data is feasible (Munger et al. 2019; Kemp 2021). MMCHIVED records protests in Venezuela from November 1, 2014 - February 10, 2015 and in Chile from October 1, 2019 - December 31, 2019.

Comparison to Similar Datasets

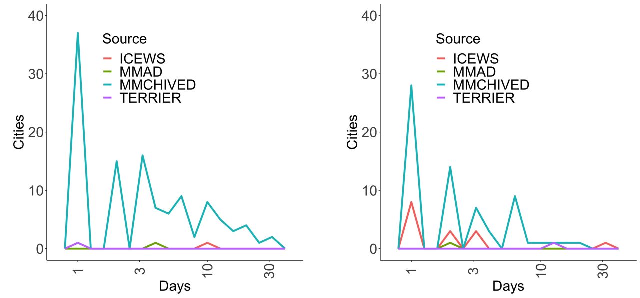
To evaluate MMCHIVED, we compare its records to four datasets. ICEWS has a similar geographic scope but has recorded protests in Latin America much longer (Boschee et al. 2015). MMAD records protest at the event and report level in autocracies, which includes Venezuela. The Temporally Extended, Regular, Reproducible International Event Records (TERRIER) is an extension of ICEWS, though its data stops at the end of 2015 (Liang et al. 2018). ACLED is a global dataset that started recording protests in Latin America in 2019 (Raleigh et al. 2010). ICEWS and TERRIER are

[2]Twitter strips images of their metadata, so it is not possible to directly test this assumption. For more detail on how Twitter determines the location of a tweet, see (Steinert-Threlkeld 2018).

[3]MMCHIVED cannot, however, provide its “articles", tweets, because of privacy constraints.

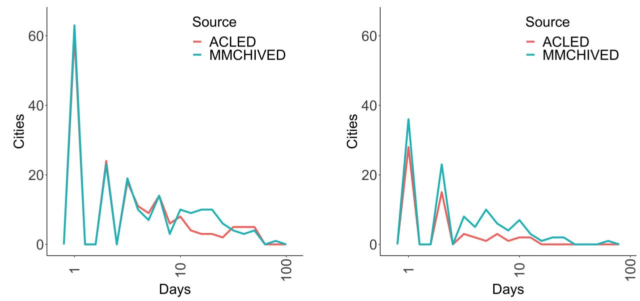
Variable	*Values*	*Aggregation*	*Content*
Protester Violence	[0,1]	Mean	Image
State Violence	[0,1]	Mean	Image
Face, Age 0-2	[0, ∞]	Sum	Image
Face, Age 3-9	[0, ∞]	Sum	Image
Face, Age 10-19	[0, ∞]	Sum	Image
Face, Age 20-29	[0, ∞]	Sum	Image
Face, Age 30-39	[0, ∞]	Sum	Image
Face, Age 40-49	[0, ∞]	Sum	Image
Face, Age 50-59	[0, ∞]	Sum	Image
Face, Age 60-69	[0, ∞]	Sum	Image
Face, Age 70+	[0, ∞]	Sum	Image
Face, Male	[0, ∞]	Sum	Image
Face, Female	[0, ∞]	Sum	Image
Face, White	[0, ∞]	Sum	Image
Face, Latin	[0, ∞]	Sum	Image
Face, Black	[0, ∞]	Sum	Image
Face, East Asian	[0, ∞]	Sum	Image
Face, Southeast Asian	[0, ∞]	Sum	Image
Face, South Asian	[0, ∞]	Sum	Image
Face, Middle Eastern	[0, ∞]	Sum	Image
Anger	{0,1}	Sum	Text
Anticipation	{0,1}	Sum	Text
Disgust	{0,1}	Sum	Text
Fear	{0,1}	Sum	Text
Joy	{0,1}	Sum	Text
Sadness	{0,1}	Sum	Text
Surprise	{0,1}	Sum	Text
Trust	{0,1}	Sum	Text

Table 1: Metadata Added per Tweet



(a) City-days with Protest. (b) City-days with Repression

Figure 2: Distribution of City-Days, Venezuela



(a) City-days with Protest. (b) City-days with Repression

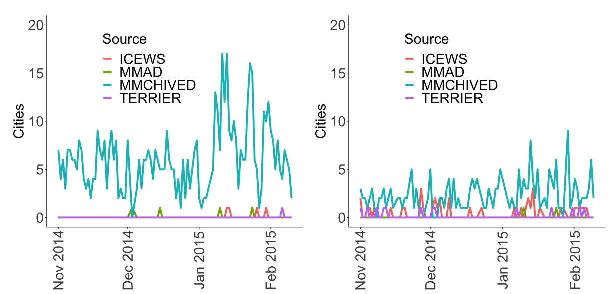
Figure 3: Distribution of City-Days, Chile

fully automated datasets while MMAD and ACLED use human coders. The first three are used exclusively for Venezuela data while ACLED is only for Chile. MMAD and TERRIER do not include Chile at the end of 2019; ACLED did not cover Venezuela during MMCHIVED's time period. Though ICEWS includes Chile, its Chile records are not analyzed here because it recorded as few there as in Venezuela. For detail on how events were identified in the other datasets, see Section "Selecting Events in Other Datasets".

The first result shows the distribution of city-days with protest or repression in Venezuela (Figure 2) and Chile (Figure 3). MMCHIVED detects many protest and repression events that occur only one day, e.g. 35 cities were recorded with one day of protest in Venezuela versus 62 in Chile. In Venezuela, MMCHIVED records more protest and repression events in more cities than MMAD, TERRIER, or ICEWS. TERRIER only identifies one day of protest in this time period, Caracas on February 6, 2015. ICEWS identifies 9 protests in Caracas and 1 in San Cristobal on January 13, 2015. MMCHIVED identifies 24 protests in San Cristobal across as many days. MMAD identifies 6 protests on 6 dates, all from Caracas. In Chile, MMCHIVED and ACLED record very similar distributions, much closer across the two datasets than for Venezuela.

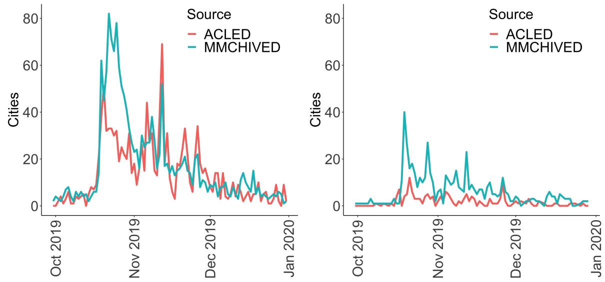
The second result compares time trends. MMCHIVED also records more sustained protest in Venezuela than the other datasets, as the time series in Figure 4 show. Figure 4(a) shows that MMCHIVED provides a more continual measure of protests than the text datasets. Restricting analysis to cities included in both datasets shows they record similar numbers of repressive events, though "Use tactics of violent repression" (event code 175 and the most natural comparison to our measure) is recorded less frequently than in the image dataset. In Chile, shown in Figure 5, the time series records largely match, though MMCHIVED records much more widespread protest and repression in mid-October.

These four figures suggest that MMCHIVED ameliorates selection and fatigue bias. Matching previous research, ICEWS, MMAD, and TERRIER record more repression events than protests (Myers and Caniglia 2004). MMCHIVED, however, records more protests than repression events. While the True number of protest and repres-



(a) Cities with Protest (b) Cities with Repression

Figure 4: Trends in Protest and Repression in Venezuela



(a) Cities with Protest (b) Cities with Repression

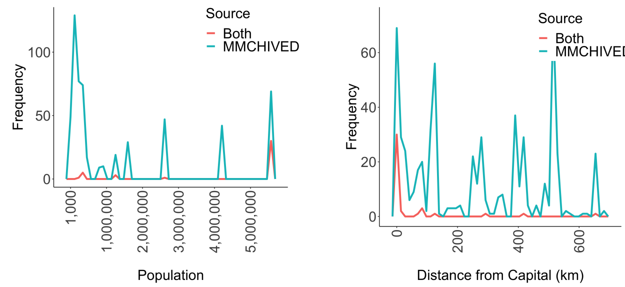
Figure 5: Trends in Protest and Repression in Chile

sion events will never be known, the increased prevalence of protests in MMCHIVED compared to repression suggests that it exhibits less selection bias towards state violence than datasets that rely on traditional media. MMCHIVED also does not appear to exhibit fatigue bias, a common concern for event data (Beieler et al. 2016). The persistence of its records is in contrast to many datasets that rely on major international newspapers and suggests a major benefit of incorporating local media and sources when possible (Nam 2006). In Venezuela, only ICEWS records repression events consistently. Selection and fatigue bias do not appear to affect ACLED in Chile, probably due to its expansive use of sources. That ACLED and MMCHIVED match so well in Chile also suggests that the lack of selection and fatigue bias in Venezuela is not an artifact of the data creation process.

Next, Figures 6 and 7 show that MMCHIVED records protests from small cities far from each country's capital. Each panel shows the distribution of events recorded in MMCHIVED or one of the other datasets as a function of city population (a) or distance from Caracas or Santiago (b). In Venezuela, MMCHIVED's extra events are recorded in small cities far from Caracas, as Figure 6 shows. ICEWS, MMAD, and TERRIER exhibit a clear bias to large cities near the capital, a well-known behavior of media-based event data (Myers and Caniglia 2004).

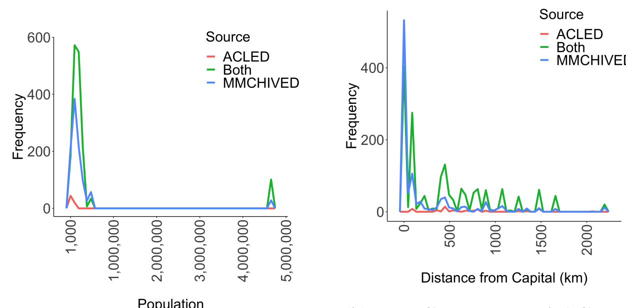
In Chile, on the other hand, ACLED and MMCHIVED perform similarly, as Figure 7 shows. While there are events that only MMCHIVED records, they do not differ in city population or distance from Santiago; the same is true of the much smaller number of events that only ACLED records. Most events are recorded in each dataset, and they also match the ACLED and MMCHIVED distributions. 103 events are only in ACLED versus 1,104 only in MMCHIVED; 1,668 are in both.

Lastly, three additional checks validate MMCHIVED's data. First, Figure 10 in Appendix shows that MMCHIVED's protest size estimates are not biased. Second, Section "MMCHIVED Records True Positives" shows that the extra events MMCHIVED records are true positives. Third, Section "Complements Existing Datasets" shows that even when MMCHIVED and other datasets record the same events, they record different features and so complement each other.



(a) MMCHIVED records small cities. (b) MMCHIVED records cities far from the capital

Figure 6: Comparing Missingness in Venezuela (Note: "Both" refers to MMCHIVED and at least one of ICEWS, MMAD, or TERRIER. No Venezuela event was observed in a dataset that was not also recorded in MMCHIVED.)



(a) MMCHIVED and ACLED record small cities. (b) MMCHIVED and ACLED record cities far from the capital.

Figure 7: Comparing Missingness in Chile

Application to Protest Dynamics

Modeling protest dynamics demonstrates the utility of MMCHIVED for the study of protests. The outcome of interest is protest size, operationalized by taking the logarithm of the sum of the number of faces in protest photos. [4] To measure violence, the average of the classifier output for protester violence and state violence is taken for all images from a city-day. To capture non-linear dynamics, we include squared terms of each. To measure emotion, we record the percent of tweets that contain keywords for anger, joy, sadness, or fear; categories are not mutually exclusive. We measure protester demographics several ways. For age, we record the percent of protester faces estimated be 19 or younger (youth), 20-29 (young adult), or 30 and older (adult). For gender, we record the percent of faces that are male. For race, we record the percent of faces FairFace classifies as White, Latinx, or Black. Because dominance of protest attendance by one demographic group should be associated with lower protest size and this complexity will not be captured with only linear terms, square terms are included. All independent variables are lagged one day. Country fixed effects are included and

[4]This does not estimate the exact number of people in the scene due to missed and redundant detection and selective posting. However this captures the relative size of protester groups.

standard errors are clustered by city. Table 2 presents these results; i indexes cities and t days.

While this paper does not develop theoretical expectations around violence, demographics, or emotions, the results intrigue. Regarding state violence, the regression finds the same curvilinear relationship between state violence and subsequent protest others have found (Zhukov 2018). Protester violence only has the expected negative relationship at large values, though this result is not statistically significant (Simpson, Willer, and Feinberg 2018). According to (Pearlman 2013), anger and joy should positively correlate with increased protest, while fear and sadness should negatively correlate. Only sadness is in the direction expected, and it is the only emotion variable to achieve statistical significance. Regarding demographics, that $Young\ Adult\ \%_{i,t-1}$ and $Male\ \%_{i,t-1}$ are positive and significant align with the biographic availability framework (Beyerlein and Hipp 2006). For race, $White\ \%_{i,t-1}$ and $Black\ \%_{i,t-1}$ exhibit n-shaped, statistically significant relationships; the $Latinx$ variables, the same but not significant. Moreover, that the demographic variables all exhibit this shape provides more confidence in the validity of MMCHIVED variables.

Discussion

One concern is that MMCHIVED will only record major events since it relies on general knowledge generated from mass media to know where and when to search for protests. While this concern is true of all event datasets because media prefer to report on large protests (Myers and Caniglia 2004), MMCHIVED records protests from smaller cities further from capital cities than existing datasets. The average number of tweets per city-day is 13.92, the median 3, and maximum 1,606: MMCHIVED records major and non-major protests.

It may also be the case that characteristics of who tweets affects inferences about protest dynamics. For example, manual inspection of accounts can identify journalists (Lotan et al. 2011), and it is easy to subset the data for verified accounts. Tweets from these accounts may be less emotional than others, and this difference could explain the relatively poor performance of the emotions model. Others have found, however, that excluding them does not change results and that protesters and non-protesters have similar mobility patterns (Larson et al. 2019).

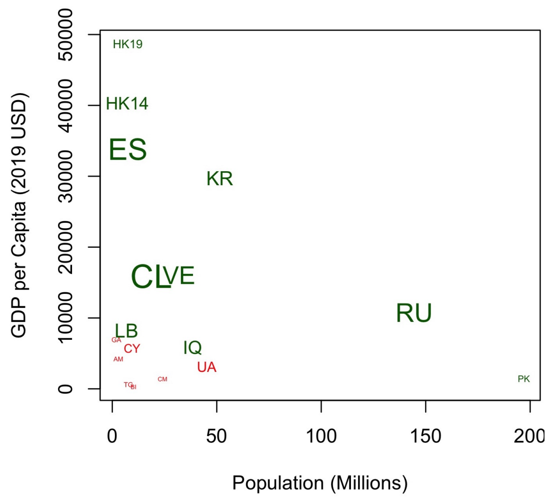
To estimate the geographic coverage that multimodal event data from Twitter can provide, we analyze data from the seventeen countries and nineteen periods listed in Table 3. The bold rows are those for which at least one protest image per week was identified at the city level. Figure 8 plots these nineteen periods against country population and gross domestic product per capita, revealing that those two variables demarcate the boundary of MMCHIVED's methodology.

More geolocated tweets can be collected, and more protests therefore detected, with four modifications of the pipeline. First, MMCHIVED relies on the one percent streaming endpoint, which Twitter provides for free. With Version 2 of the Twitter API, any tweet ever published from any account still public is available, potentially making many more pieces of content available. Second, a researcher can connect to the filtered stream endpoint and download tweets

DV: Concept:	Faces Violence (1)	Tweets Emotions (2)	Faces Demographics (3)
$Protester\ Violence_{i,t-1}$	.1517		
	(.1231)		
$Protester\ Violence^2_{i,t-1}$	−.1256		
	(.1278)		
$State\ Violence_{i,t-1}$	.4840		
	(.3556)		
$State\ Violence^2_{i,t-1}$	−.9798*		
	(.5826)		
$Anger\ \%_{i,t-1}$		−.0084	
		(.0151)	
$Joy\ \%_{i,t-1}$		−.0050	
		(.0107)	
$Fear\ \%_{i,t-1}$		.0231	
		(.0147)	
$Sadness\ \%_{i,t-1}$		−.0182**	
		(.0088)	
$Youth\ \%_{i,t-1}$			.1922
			(.3049)
$Youth\ \%^2_{i,t-1}$			−.1656
			(.3342)
$Young\ Adult\ \%_{i,t-1}$			.3073**
			(.1479)
$Young\ Adult\ \%^2_{i,t-1}$			−.0757
			(.2562)
$Adult\ \%_{i,t-1}$			.1969
			(.2165)
$Adult\ \%^2_{i,t-1}$			.0046
			(.1514)
$Male\ \%_{i,t-1}$			.3332**
			(.1643)
$Male\ \%^2_{i,t-1}$			−.4230*
			(.2163)
$White\ \%_{i,t-1}$			.7341**
			(.3272)
$White\ \%^2_{i,t-1}$			−.9520*
			(.5185)
$Black\ \%_{i,t-1}$			1.2250**
			(.4890)
$Black\ \%^2_{i,t-1}$			−2.3428**
			(1.0517)
$Latinx\ \%_{i,t-1}$			.0326
			(.1485)
$Latinx\ \%^2_{i,t-1}$			−.0265
			(.1439)
$DV_{i,t-1}$	.5185***	.5748***	
	(.1149)	(.1102)	
Intercept	.3954***	.2726***	.5590***
	(.1047)	(.0528)	(.1069)
N	5,962	7,612	5,962
Adjusted R^2	.5075	.1459	.3668
FE	Country	Country	Country
Cluster SE	City	City	City

*p < .1; **p < .05; ***p < .01

Table 2: How Violence, Emotions, and Demographics Affect Protest Dynamics. The Emotions model uses tweets as the dependent variable because more tweets contain emotion information than images; the number of tweets is a reliable estimate of protest size (Sobolev et al. 2020). The Demographics model excludes a lagged dependent variable to provide enough variation for the other variables. i indexes cities and t days. DV: Faces = Log(Faces + 1)$_{i,t}$, Tweets = Log(Tweets + 1)$_{i,t}$.



Note: Red countries have fewer than one protest image per seven days, at a city level of resolution. Size corresponds to the number of protest images per day.

Figure 8: Protest is Detected in Populous or Medium-Income Countries

from specific accounts in real time. Twitter now limits academic accounts to downloading 10 million tweets per month, so caution is required. If this quota binds, the third option is to purchase tweets. Fourth, post-processing can improve geolocation. In Venezuela, for example, MMCHIVED discards about half of the protest images because Twitter only provided their location at the country level. One can infer an account's location, and assign it to all tweets from that account, using the user biography, self-reported location, or locations mentioned in the tweet itself (Ryoo and Moon 2014).

Future work should focus on expanding the event types MMCHIVED records and incorporate researcher oversight (Oostdijk et al. 2020). It may be feasible, however, to generate events around speeches or military activity, as governments increasingly communicate on social media with text and image, and text from government accounts could be interpreted as actions falling into one of CAMEO's categories. A protest image classifier can screen all images for those likely to contain a protest, and teams of trained human coders can manually annotate them. This approach will work best for the demographic and size variables. Depending on resources available, this hybrid approach may limit the breadth of panel data.

Appendices

Selecting Events in Other Datasets

ICEWS and TERRIER use the same event ontology, Conflict and Mediation Event Observations (CAMEO), so extracting the same type of events is straightforward. Any event with the root code of `14` is a protest. Repression can occur in various guises under root code `15` ("Exhibit Force Posture"), `17`, ("Coerce"), or `18` ("Assault"). Event code `175` ("Use tactics of violent repression") is the closest to protest policing, but the news datasets record few enough of those that the following results include all events we identify as repression.[5] For MMAD, each row is a protest, and we use the `max_secengagement` variable, which is ordinal, to identify repression. Once protest and repression events are identified in these text datasets, they are merged with the image event data based on matching city days.

Duplicate Investigation

Figure 9 shows the image duplication rate by city and country. Bar height is the number of tweets per city. MMCHIVED deduplicates images before aggregating by city and day.

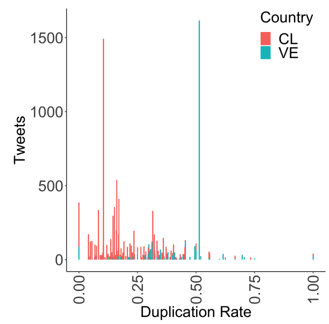
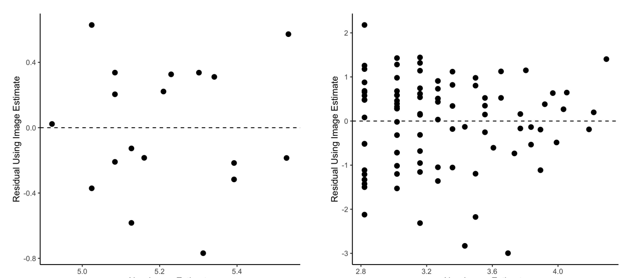


Figure 9: Duplication Rate by City

Residual Investigation

Figure 10 shows that the size estimates in MMCHIVED are not biased when compared to newspaper estimates. Shown are the residuals from regressing MMCHIVED's protest size estimate on newspaper estimates reported from Venezuela and Chile. The x-axis of each figure is the estimate from newspapers. These results match similar findings in the United States (Sobolev et al. 2020).



(a) Venezuela (b) Chile

Figure 10: No Bias Found in MMCHIVED Size Estimates. Note: The x-axis is the logged reported protest size from the Associated Press, Agence-France Press, or BBC Monitoring. The y-axis is the residual from regressing MMCHIVED's protest size on the reported protest size. Shapiro-Wils and Kolmogorov-Smirnov tests confirm the residuals are normally distributed: MMCHIVED's protest size estimates appear unbiased.

[5]Event codes `151, 153, 170, 171, 1711, 1712, 1723, 1724, 173, 180, 182, 1823` are the other CAMEO events we include as repression.

MMCHIVED Records True Positives; Other Datasets, False Negatives

Importantly, the extra events MMCHIVED records are true positives. Wikipedia provides a comprehensive record of protests in Venezuela. Though we do not compare MMCHIVED to Wikipedia because the latter is much stronger for 2015 than 2014, many protests from those pages appear in MMCHIVED. For example, residents protested in Caracas on January 25th, 2015, and our raw data contain 12 tweets with protest images from there then; the other datasets record no protest. MMCHIVED records 16 protests on January 23rd, compared to 0 in the other datasets; Wikipedia documents protests on both of these days. Figures 11(a) and (b) show images from protests recorded in Merida and Caracas. Anti-Maduro protests occurred on February 4, 2015 in Merida (Patillia 2015); ICEWS records an arrest of citizens in Caracas on that day, and the other datasets record nothing. MMCHIVED also records a teacher protest in Caracas on November 6, 2014 that does not appear in the other datasets (Patilla 2014).

(a) 02.04.2015: Peaceful Protest, Merida

(b) 11.06.2014: Peaceful Protest, Caracas

Figure 11: Protest Events Recorded in Social Media but not Other Datasets

MMCHIVED Complements Existing Datasets

Next, MMCHIVED complements existing datasets when they record the same events. MMCHIVED records different types of repression information for events that the text datasets also record, as Figures 12(a) and (b) show. ICEWS records a protest in Caracas on January 24, 2015, but it does not record protester violence (Figure 12a) or the presence of police (Figure 12b); TERRIER records no protest in Venezuela. MMAD records this protest, no protester violence, and the presence of police. The image, however, permits differentiation between normal and riot police as well as measuring the size of the police force.

(a) 01.24.2015: Protester Violence, Caracas

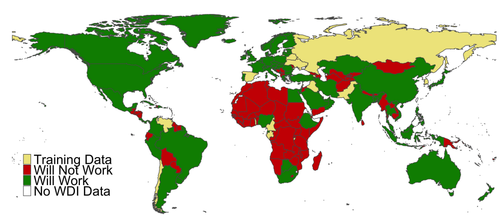
(b) 01.24.2015: State Response, Caracas

Figure 12: Images Provide Additional Detail About Events in Other Datasets

In Chile, MMCHIVED and ACLED complement each other on what they record. ACLED records 66 types of actors, 51 conducting action against 15. These actors, however, are broad labels, like "Labour Group" or "Civilians". The attribute characteristics that images can generate can then provide demographi information about the these actors. Regarding size, 63.65% of ACLED's events have sized recorded as "no report", 2.5% are "thousands", 2.24% are "hundreds", and the vast majority of the rest include phrases such as "around", "nearly", or "at least". Only 2.48% contain an exact number, reflecting the difficulty of measuring protest size when relying on text. In contrast, 46.93% of events recognized from images contain countable faces and therefore a size estimate. Regarding violence, ACLED can infer protester violence from the "violent demonstration" or "mob violence" sub-event types and state violence from the "protest with intervention" and "excessive force against protesters" sub-event types or the `fatalities` variable. Except for `fatalities`, these variables are nominal, in contrast to the continuously valued estimates of protester and state violence images produce.

Global Potential of MMCHIVED's Methodology

To determine the applicability of this approach, we run a logistic regression using the 19 periods in Table 3 to model the relationship between a country's population, gross domestic product, and recording of protest. The outcome of this model is a 1 if at least one protest is observed in one city once per week and a 0 otherwise. A clear boundary at predicted values of .38 emerges, so the same model is applied to data on every country. Figure 13 shows the result of this model; green countries are those in which MMCHIVED is predicted to record at least one protest per week in one city. These countries account for 82.7% of the world's population and 97.14% of its GDP.



Note: Using geolocated images shared on Twitter will produce protest event data in 107 countries. This number will increase if one aggregates by week or augments the image collection strategy.

Figure 13: Predicting the Geographic Coverage of Protest Event Data from Twitter Geolocated Images

Determinants of Method Appropriateness

Table 3 shows the nineteen countries to which we applied the image methodology.

Country	Start	End	Issue
Armenia	03.01.18	05.31.18	Anti-incumbency
Belarus	02.18.17	05.02.17	Unemployment tax
Burundi	04.01.15	12.01.15	Elections
Chile	**10.01.19**	**12.31.19**	**Inequality**
Cameroon	11.01.16	12.01.17	Bilingualism
Egypt	06.01.17	06.31.17	Islands to Saudi Arabia
Gabon	08.20.16	09.27.16	Elections
Hong Kong	**09.18.14**	**12.23.14**	**China reforms**
Hong Kong	**03.01.19**	**12.31.19**	**China reforms**
Iraq	**10.01.19**	**12.31.19**	**Incompetence**
Lebanon	**10.01.19**	**12.31.19**	**Anti-tax**
Pakistan	**11.01.17**	**11.30.17**	**Blasphemy protests**
Russia	**03.12.17**	**04.26.17**	**Corruption**
Catalonia, Spain	**09.01.17**	**12.31.17**	**Secession**
South Korea	**10.20.16**	**03.14.17**	**Anti-incumbency**
Togo	08.01.17	12.01.17	Anti-incumbency
Ukraine	11.21.13	03.21.14	European Integration
Venezuela	**11.01.14**	**02.10.15**	**Grievances**
Venezuela	**12.29.16**	**12.17.17**	**Anti-Maduro**

Table 3: Protest Periods. Bold rows are those countries where at least one city contains a protest photo for at least $\frac{1}{7}$ of its period.

Ethical Statement

Participation in a protest event as well as posting an image about it on social media can lead to harmful outcomes to individuals because state authorities or political opponents may attempt to track them down. To protest the identities of protest participants and Twitter users, we do not share any images or individual tweets. The dataset was completely made from public content in Twitter. All metadata are aggregated at the city-day level. We believe that our dataset will contribute to fairer and less-filtered reports about protest events and facilitate relevant research.

Acknowledgments

This work was supported by NSF SBE-SMA 1831848.

References

Beieler, J.; Brandt, P. T.; Halterman, A.; Schrodt, P. A.; and Simpson, E. M. 2016. Generationg Political Event Data in Near Real Time: Opportunities and Challenges. In *Computational Social Science: Discovery and Prediction*, 98–120. Cambridge: Cambridge University Press. ISBN 9781316257340.

Beyerlein, K.; and Hipp, J. R. 2006. A Two-stage model for a two-stage process: How biographical availability matters for social movement mobilization. *Mobilization: An International Quarterly*, 11(3): 219–240.

Boschee, E.; Lautenschlager, J.; O'Brien, S.; Shellman, S.; Starz, J.; and Ward, M. 2015. ICEWS Coded Event Data. https://doi.org/10.7910/DVN/28075. Accessed: 2019-04-10.

Bradley, R. A.; and Terry, M. E. 1952. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4): 324–345.

Chen, B.; Escalera, S.; Guyon, I.; Ponce-López, V.; Shah, N.; and Simón, M. O. 2016. Overcoming calibration problems in pattern labeling with pairwise ratings: application to personality traits. In *European Conference on Computer Vision*, 419–432. Springer.

Chenoweth, E.; and Stephan, M. J. 2011. *Why Civil Resistance Works*. New York City: Columbia University Press.

Cook, S. J.; and Weidmann, N. B. 2019. Lost in Aggregation: Improving Event Analysis with Report-Level Data. *American Journal of Political Science*, 63(1): 250–264.

Croicu, M.; and Weidmann, N. B. 2015. Improving the selection of news reports for event coding using ensemble classification. *Research and Politics*, 2(4): 1–8.

Dietrich, B. J.; Enos, R. D.; and Sen, M. 2019. Emotional Arousal Predicts Voting on the U.S. Supreme Court. *Political Analysis*, 27(2): 237–243.

Dowd, C.; Justino, P.; Kishi, R.; and Marchais, G. 2020. Comparing 'New' and 'Old' Media for Violence Monitoring and Crisis Response in Kenya. *Research and Politics*, 7(3): 1–9.

Ferrara, F. 2003. Why Regimes Create Disorder: Hobbes's Dilemma During a Rangoon Summer. *Journal of Conflict Resolution*, 47(3): 302–325.

Fisher, D. R.; Dow, D. M.; and Ray, R. 2017. Intersectionality takes it to the streets: Mobilizing across diverse interests for the Women' s March. *Science Advances*, 3: 1–8.

Francisco, R. A. 2004. After the Massacre: Mobilization in the Wake of Harsh Repression. *Mobilization: An International Journal*, 9(2): 107–126.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Jasper, J. M. 2011. Emotions and Social Movements: Twenty Years of Theory and Research. *Annual Review of Sociology*, 37(1): 285–303.

Karkkainen, K.; and Joo, J. 2021. FairFace: Face Attribute Dataset for Balanced Race. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 1548–1558.

Kemp, S. 2021. Digital 2021: Chile. Technical report, : DataReportal.

Kern, H. L. 2011. Foreign Media and Protest Diffusion in Authoritarian Regimes: The Case of the 1989 East German Revolution. *Comparative Political Studies*, 44(9): 1179–1205.

Larson, J. M.; Nagler, J.; Ronen, J.; and Tucker, J. A. 2019. Social Networks and Protest Participation: Evidence from 130 Million Twitter Users. *American Journal of Political Science*, 63(3): 690–705.

Liang, Y.; Jabr, K.; Grant, C.; Irvine, J. A.; and Halterman, A. 2018. New Techniques for Coding Political Events Across Languages. *IEEE 18th International Conference on Reuse and Integration*, 89–93.

Lotan, G.; Ananny, M.; Gaffney, D.; Boyd, D.; Pearce, I.; and Graeff, E. 2011. The Revolutions Were Tweeted: Information Flows During the 2011 Tunisian and Egyptian Revolutions Web. *International Journal of Communications*, 5: 1375–1406.

McPhail, C.; and McCarthy, J. 2004. Who Counts and How: Estimating the Size of Protests. *Contexts*, 3(3): 12–18.

Munger, K.; Bonneau, R.; Nagler, J.; and Tucker, J. A. 2019. Elites Tweet to Get Feet Off the Streets: Measuring Regime Social Media Strategies During Protest. *Political Science Research and Methods*, 7(4): 815–834.

Myers, D. J.; and Caniglia, B. S. 2004. All the Rioting That's Fit to Print: Selection Effects in National Newspaper Coverage of Civil Disorders, 1968-1969. *American Sociological Review*, 69(4): 519–543.

Nam, T. 2006. What You Use Matters: Coding Protest Data. *PS: Political Science & Politics*, 39(2): 281–287.

Nordås, R.; and Davenport, C. 2013. Fight the Youth: Youth Bulges and State Repression. *American Journal of Political Science*, 57(4): 926–940.

Oliver, P.; Marwell, G.; and Teixeira, R. 1985. A Theory of the Critical Mass. I. Interdependence, Group Heterogeneity, and the Production of Collective Action. *American Journal of Sociology*, 91(3): 522–556.

Oostdijk, N.; van Halteren, H.; Başar, E.; and Larson, M. 2020. The connection between the text and images of news articles: New insights for multimedia analysis. In *12th International Conference on Language Resources and Evaluation, Conference Proceedings*, 4343–4351. Marseille: European Language Resources Association. ISBN 9791095546344.

Patilla, L. 2014. Profesores protestaron cerquita de Miraflores. https://www.lapatilla.com/2014/11/06/profesores-protestaron-cerquita-de-miraflores-fotos/. Accessed: 2020-05-14.

Patilla, L. 2015. En 2014 se registraron 9.286 protestas, cifra inédita en Venezuela. https://www.lapatilla.com/2015/01/19/en-2014-se-registraron-9-286-protestas-cifra-inedita-en-venezuela/. Accessed: 2020-05-14.

Patillia, L. 2015. #MaduroRenuncia: La consigna de la protesta merideña de este miércoles. https://www.lapatilla.com/2015/02/04/madurorenuncia-la-consigna-de-la-protesta-meridena-de-este-miercoles-fotos/. Accessed: 2020-05-14.

Pearlman, W. 2013. Emotions and the Microfoundations of the Arab Uprisings. *Perspectives on Politics*, 11(02): 387–409.

Raleigh, C.; Linke, A.; Hegre, H.; and Karlsen, J. 2010. Introducing ACLED: An Armed Conflict Location and Event Dataset: Special Data Feature. *Journal of Peace Research*, 47(5): 651–660.

Rasler, K. 1996. Concessions, Repression, and Political Protest in the Iranian Revolution. *American Sociological Review*, 61(1): 132–152.

Ritter, E. H.; and Conrad, C. R. 2016. Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression. *American Political Science Review*, 110(1): 85–99.

Robnett, B. 1996. African-American Women in the Civil Rights Movement, 1954-1965: Gender, Leadership, and Micromobilization. *American Journal of Sociology*, 101(6): 1661–1693.

Ryoo, K.; and Moon, S. 2014. Inferring Twitter user locations with 10 km accuracy. *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion*, 643–648.

Salehyan, I.; Hendrix, C.; Hammer, J.; Case, C.; Linebarger, C.; Stull, E.; and Williams, J. 2012. Social Conflict in Africa: A New Database. *International Interactions*, 38(4): 503–511.

Samper-Escalante, L. D.; Loyola-González, O.; Monroy, R.; and Medina-Pérez, M. A. 2021. Bot Datasets on Twitter: Analysis and Challenges. *Applied Sciences*, 11(9): 4105.

Scarritt, J. R.; McMillan, S. M.; and Mozaffar, S. 2001. The Interaction Between Democracy and Ethnopolitical Protest and Rebellion in Africa. *Comparative Political Studies*, 34(7): 800–827.

Simpson, B.; Willer, R.; and Feinberg, M. 2018. Does Violent Protest Backfire? Testing a Theory of Public Reactions to Activist Violence. *Socius: Sociological Research for a Dynamic World*, 4: 237802311880318.

Sobolev, A.; Joo, J.; Chen, K.; and Steinert-Threlkeld, Z. C. 2020. News and Geolocated Social Media Accurately Measure Protest Size. *American Political Science Review*, 1–9.

Steinert-Threlkeld, Z.; Chan, A.; and Joo, J. 2022. How State and Protester Violence Affect Protest Dynamics. *The Journal of Politics*.

Steinert-Threlkeld, Z.; and Joo, J. 2021. MMCHIVED: Multimodal Chile and Venezuela Event Data. https://doi.org/10.7910/DVN/3BTUXJ. Accessed: 2021-02-07.

Steinert-Threlkeld, Z. C. 2018. *Twitter as Data*. Cambridge University Press.

Steinhardt, H. C.; and Gobel, C. 2019. How should we measure protests in authoritarian regimes? A comparison of traditional, dissident, and social media event data from China. *Working paper*.

Weidmann, N. B.; and Rød, E. G. 2019. *The Internet and Political Protest in Autocracies*. Oxford University Press.

Won, D.; Steinert-Threlkeld, Z. C.; and Joo, J. 2017. Protest Activity Detection and Perceived Violence Estimation from Social Media Images. In *Proceedings of the 25th ACM International Conference on Multimedia*, 786–794. ISBN 9781450349062.

Wouters, R.; and Walgrave, S. 2017. Demonstrating Power: How Protest Persuades Political Representatives. *American Sociological Review*, 82(2): 361–383.

Young, L. E. 2019. The Psychology of State Repression: Fear and Dissent Decisions in Zimbabwe. *American Political Science Review*, 113(1): 140–155.

Zhang, H.; and Pan, J. 2019. CASM: A Deep-Learning Approach for Identifying Collective Action Events with Text and Image Data from Social Media. *Sociological Methodology*, 49: 1–48.

Zhukov, Y. M. 2018. Repression 'Works' (just not in moderation). *Working paper*.