**Online Appendix:**

**Full list of rebel group Twitter accounts**

Primary analysis in this study is based on an original corpus consisting of around 270,000 tweets from Twitter accounts identified as the official, centrally-controlled accounts for the groups of interest. As described in the main text of the article, we began by extracting a list of conflict episodes between states and nonstate challengers that were active between January 2006 and July 2018 from the UCDP/PRIO Armed Conflict Dataset, version 18.1. This list consisted of 477 conflict episodes involving 155 unique rebel organizations. We then worked through this set of 155 groups, identifying official group Twitter accounts where they existed, yielding a final list of 76 Twitter accounts belonging to 55 rebel groups. Table A1 below lists each of these groups, their locations, and associated accounts in full, and includes account-level data on tweet activity, followers, and friends as well as the original date of account activation.

[Table A1 about here]

**Notes on topic classification**

A central task in this analysis was to classify the tweets we collected according to the topics they contain. Broadly speaking, there are three approaches to this task. In the first, hand coding, every tweet is read and categorized by a human coder. The accuracy of this approach benefits from close human supervision, as might be expected, but it is a slow, resource-intensive process that quickly becomes unfeasible as the number of documents in a corpus grows. In the second approach, unsupervised learning, the data are preprocessed and then passed through some variety of clustering algorithm that identifies similarities between documents and sorts them into categories accordingly. With this approach, classification is fully-automated, and the researcher’s task is to interpret the clusters returned by the algorithm. In this sense, unsupervised approaches are truly “letting the data speak,” but they allow little room for the researcher to adhere to a theoretically-informed typology if one exists. A middle ground between these approaches is supervised learning, where human coders prepare a training set of hand coded observations which are then used to train a classifier to identify the features that characterize each category. This way, the human researcher is kept “in the loop” while still benefiting from the advantages of automation at scale.

In this study, since we started with a theoretically-informed topic classification scheme, we adopted a supervised learning approach. We created a training set by hand-coding a random sample of 10,000 tweets out of the corpus. Then, we trained classifiers on these data, working to maximize accuracy in predicting the human-generated labels of the training data over 10-fold cross validation. We tried three classes of algorithm for this task—linear support vector machines (SVM), nonlinear SVM with an RBF kernel, and random forest classifiers. Although it was not the most accurate model we tried, we ultimately settled on the linear SVM because of two useful characteristics. First, linear SVMs produce coefficients that can be directly interpreted as feature weights. This provided us with the ability to see what terms the algorithm considered highly predictive of one category over another, allowing for validation of the model before applying it to unlabeled data. Second, the errors we observed in the linear SVM tended to be in favor of overpredicting the “other” category, as shown in Figure A1 on the bottom panel. This is a desirable characteristic, since it is essentially biasing the model toward the null rather than in favor of any of our categories of interest.

[Figure A1 about here]

**Notes on follower location identification**

In this study, we used the user-reported location field of each rebel account’s followers to create measures of the proportion of their audience who were domestically and internationally located. This measure allowed us to make some interesting observations about the way social media affords rebel groups new opportunities to frame and present their message for a broad international audience. Unfortunately, these measures are inherently noisy, as outlined in the main text of this article, primarily because Twitter does not require users to provide location information. As a result of this policy, we observe a significant nonresponse rate. With an ideal measure, the proportion of domestic and international followers would be a perfect inverse of each other, negatively correlated at r = -1. In the data we collected, this is not quite the case, as shown in Figure A2 on the right-hand panel. In the measures we were able to observe, the proportion of domestic and international followers correlates at r = -0.60. This highlights two facts. First, the lack of perfect correlation confirms that these measures are noisy. Second, though, the fact that they do correlate and in the expected direction demonstrates that they at least providing a successful approximation of the quantities of interest.

[Figure A2 about here]

**Additional Figures**

Figure A3 below shows the breakdown of primary and secondary account languages in the Rebel Twitter Dataset. We considered a language to be “primary” if more than 50% of the tweets sent from a given account used that language, “secondary” otherwise.

[Figure A3 about here]

Figures A4 – A6 show examples of tweets classified respectively as *mobilization*, *operations*, and *self-promotion*, as they appear on twitter.com.

[Figure A4 about here]

[Figure A5 about here]

[Figure A6 about here]