

Supplemental Information
for
The Supply of Conspiracy Theories in State-controlled
Media

Gabriel Koehler-Derrick* Richard A. Nielsen[†] David Romney[‡]

December 5, 2022

For online publication only.

Abstract

This online appendix provides supplemental information about the analysis in “The Supply of Conspiracy Theories in State-controlled Media”

*Assistant Professor, New York University, Abu Dhabi, gkd225@nyu.edu

[†]Associate Professor, Massachusetts Institute of Technology, rnielsen@mit.edu.

[‡]Assistant Professor, Brigham Young University, david.romney@byu.edu

Contents

A Automated Conspiracy Theory Detection	3
B Regression Table for Main Models	6
C Robustness Checks	7
C.1 Model Specifications: Poisson Regression, OLS	7
C.2 Measurement: Aggregating to Article	7
C.3 Conspiracy Theory Framing: Only Endorsed	8
C.4 Classifying articles without key words	9
C.5 Including <i>Al-Ahram</i> from 1998	9
C.6 Individual ACLED Event Categories	10
C.7 Logged ACLED Events	10
C.8 Alternative Measure: ACLED death count	11
C.9 ACLED crisis measure, 75th and 90th percentile	11
C.10 Alternative Measure: START data	12
C.11 Alternative Measure: Clarke Anti-Regime Protest data, January 2012-June 2013 . .	13
C.12 Omitting Combination of control variables	13
C.13 Administration fixed effects	14
C.14 ACLED Same day and Day lag	14
C.15 Including <i>Al-Shuruq</i> and <i>Al-Gomhuria</i> Newspapers	15

A Automated Conspiracy Theory Detection

In this section, we discuss additional details about how we classify conspiracy theories.

Paragraph-level classification: The default unit in most text analysis settings is each natural text, which in our case, might reasonably be each news article as a whole. However, we found that conspiracy theory language is typically a relatively small part of any single article, so classifying entire articles as conspiracy theories or not resulted in substantial measurement error. We improved our classification accuracy by instead breaking articles into paragraphs (10.5 million).

Keyword-assisted classification approach: Conspiracy theories are relatively rare in our data. Based on our close reading, our best estimate is that to obtain 50 examples of conspiracy theory-related paragraphs using standard approaches, we would have to sample and code at least 15,000 paragraphs by hand. Training our model with such an imbalanced training set would have resulted in very low classification accuracy.

We turned to key words because our reading uncovered that find that Egyptian journalists reliably use these words when discussing conspiracy theories. We identified a list of 18 key words that are conspiracy-related including variants of: “conspiracy” (e.g. المؤامرة, التآمر, etc.); “trick” or “machination” (e.g. دسيسة, مكيدة, etc.) and “collusion” (e.g. المتواطئ, التواطؤ). Our review of a range of media materials confirms that these phrases are commonly used when conspiracy theories are discussed in the Arabic media, and are only sometimes associated with the discussion of other topics, because of other meanings or connotations. For example, we considered using other key words, such as variants of “plan” (خطط) and “interference” (التدخل), but an examination of the paragraphs in which these words were used showed that they were more often unrelated to conspiracy theories.

Hand-coding process: We sampled 1,500 articles that contain these key words and had two research assistants separately code these articles, paragraph by paragraph. We developed coding criteria for our coders by reading hundreds of conspiracy theories in both newspapers ourselves. We started from definitions of conspiracy theories from previous scholarship. We looked for text that alleged the role of “unseen and malevolent forces,” providing an interpretation of events using “Manichean” language, and discredited “mainstream” explanations (Oliver and Wood, 2014). However, a single paragraph need not describe each of these components in full to be coded a conspiracy theory, because authors often leave some aspects implicit. We instead identify what might be called conspiratorial language, based on our understanding that the politics of conspiracies may not necessarily require the author to explicitly provide a comprehensive “theory.”

We trained our coders to classify paragraphs in articles as conspiracy theory or non-conspiracy theory, based on this definition of a conspiracy theory. Where present, we asked them to identify the perpetrator and victims in each conspiracy theory, including several terms for vaguely specified entities. We also asked the coders to code the “frame” of each individual article taking into consideration whether the article appeared to endorse the conspiracy theory, whether the author presented the conspiracy theory in a neutral way, often through a direct quotation, or whether the author was critical of the conspiracy theory. We do not evaluate the truth of these theories. We instructed our coders to include any paragraph that fit our definition whether they considered the claim of conspiracy to be true or false.

While developing our coding rules, we observed some conspiracy-related content in contexts that were not immediately relevant to contemporary politics: movies, art, and books, sports, and

in discussions of historical events, particularly religious discussions about the early Muslim community. We had our research assistants generally exclude references to conspiracy theories in the context of the arts, sports, and historical events prior to the 1900s, unless they also related to contemporary Egyptian politics.

Our coders agreed in their top-level coding of “conspiracy theory” or “non-conspiracy theory” in 95.9% of paragraphs in these 1,500 articles. We reconciled the disagreements in this variable by having two of the authors read and adjudicate every paragraph for which our coders disagreed. We did not reconcile the other variables coded by the research assistants because they were not essential for our classification task.

Examples: A clear example of a conspiracy theory paragraph is the following:

...the restoration of the standing of the state is the most important of recent achievements, even if it [required] the use of force. We are facing a conspiracy and an enemy whose composition we do not yet know. Defense is a legitimate right in the face of a nebulous enemy...

...ان استعادة هيبة الدولة كان اهم منجزات الفترة الأخيرة حتى وان استخدمت القوة لأننا امام مؤامرة وعدو لم نعرف بعد كل اطرافه وهنا يصبح الدفاع حقا مشروعا امام عدو غامض...

A more difficult case, because of undetermined endorsement, was the following:

Many of my colleagues here (in Canada) explain this simply through conspiracy theory claims: For example, that the Brotherhood were (and still are) lackeys of foreign parties, and that there is an understanding between them and the US to resolve the Palestinian situation at the expense of Egypt through giving up a piece of the Sinai

بعض زملائي وزميلاتي هنا (في كندا) يشرحون هذا ببساطة بنظرية المؤامرة، مثلا أن الإخوان كانوا ولا يزالون عملاء الخارج وأن هناك تفاهما بينهم وبين واشنطن مثلا على حل القضية الفلسطينية على حساب مصر والتنازل عن جزء من سيناء.

Classifier details: We train our classifier on 22,190 paragraphs from the 1,500 articles that our research assistants coded by hand. Of these, 1,647 paragraphs were coded as conspiracy theories, and 20,543 were coded as not. We removed stop words and punctuation, and then stemmed the Arabic text before training the classifier. Using the Caret package (Kuhn et al., 2014), we partitioned the labeled data into an 80/20 split of training set (17,753 paragraphs) and test set (4,437). We used out-of-bag resampling with 10 resamples and 1,000 trees. After running the classifier on the identical training/test set over a range of possible parameter values (the number of trees, the number of out-of-bag resamples and the number of variables randomly sampled at each split), we choose the specification that performed best on overall accuracy, sensitivity, and specificity. Our final model yielded an accuracy of 0.977, with a sensitivity of 0.79 (accuracy at correctly identifying conspiracy theories) and a specificity of 0.99 (accuracy at correctly identifying non-conspiracy theories).

With this labeled set in hand, we then used a random forest classifier to classify the remaining

		Actual	
		conspiracy	not
Predicted	conspiracy	260	31
	not	69	4,077

Table 1: Confusion matrix showing classifier accuracy in the held-out portion of the training set.

449,297 paragraphs in the 31,096 articles that contained our key words but are not in the 1,500 article training set. Of these paragraphs, 52,412 (from 23,514 articles) are classified as conspiracy theories. Combining our hand-labeled set (1,647 paragraphs) with the set predicted by the classifier (52,412 paragraphs), results in 54,059 conspiracy theory paragraphs, from 24,806 articles. With 6,360,805 total paragraphs in *Al-Ahram* and 3,708,572 in *Al-Masry Al-Youm*, the percentage of conspiracy theory paragraphs in *Al-Ahram* is 0.52% and the percentage in *Al-Masry Al-Youm* is 0.51%.

B Regression Table for Main Models

Table 2 shows the full regression results corresponding to Figure 2 in the paper.

	Model 1 Outcome: Conspiracy theory paragraphs	Model 2 Outcome: Conspiracy theory paragraphs
Intercept	0.04	-0.49*
	0.067	0.14
ACLEDEvents	0.007*	-0.002
	0.0007	0.001
Al-Ahram	-0.18*	-0.14*
	0.027	0.027
ACLEDEvents × Al-Ahram	0.007*	0.006*
	0.0009	0.0009
Paragraphs	-0.0001	0.0006*
	0.00007	0.00007
Articles	0.008*	0.002*
	0.0007	0.0007
News Agency Paragraphs	-0.0009	-0.001*
	0.0005	0.0005
News Agency Articles	-0.004	0.012*
	0.004	0.004
Saturday	0.25*	0.23*
	0.036	0.033
Sunday	0.099*	0.17*
	0.037	0.034
Monday	0.15*	0.23*
	0.036	0.034
Tuesday	0.16*	0.24*
	0.036	0.034
Wednesday	0.13*	0.23*
	0.037	0.034
Thursday	0.13*	0.21*
	0.036	0.034
Year-Month Fixed Effects	No	Yes
N	8,805	8,805

Table 2: Negative Binomial regression models showing that Al-Ahram oversupplies conspiracy theories relative to Al-Masry Al-Youm in response to the same events. * indicates $p < 0.05$.

C Robustness Checks

This section demonstrates that the main finding presented in Table 2 is robust to a number of alternative modeling choices, measurement strategies, controls, and extends to the inclusion of two additional newspapers, one official and one independent. For brevity, we sometimes abbreviate “conspiracy theories” as “CTs.”

C.1 Model Specifications: Poisson Regression, OLS

Our main specifications use a generalized linear model with a negative binomial link that is appropriate for modeling (potentially over-dispersed) counts. To demonstrate that our results are not dependent on this modeling choice, Figure 1 shows the results of our same specification but utilizing a Poisson regression, typically used for count data, as well as a standard linear regression specification. In both cases we can see a clear difference in the supply of conspiracy theories between *Al-Ahram* and *Al-Masry Al-Youm* as the ACLED event count variable increases.

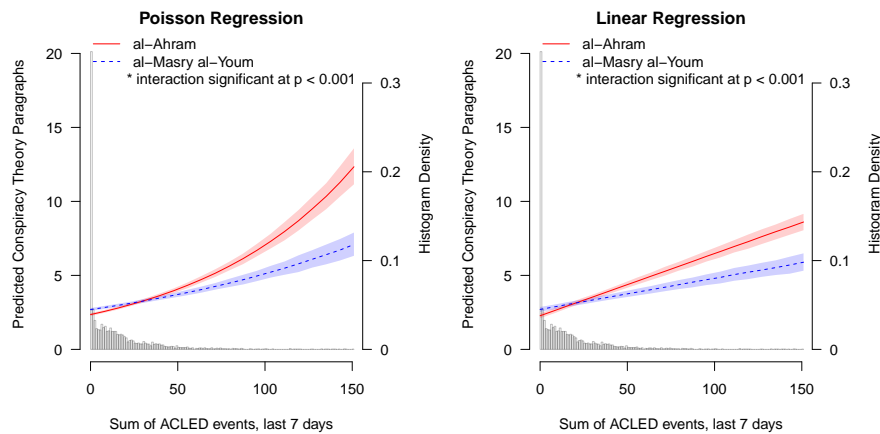


Figure 1: Main results robust to using Poisson or OLS regression.

C.2 Measurement: Aggregating to Article

An important measurement decision was to determine our unit of analysis—the level of text at which we think conspiracy theory language will be most detectable. The default unit in most text analysis settings is each natural text, in this case, each news article. However, based on manual examination of conspiracy theory-related articles, conspiracy theory language composed a relatively small portion of such articles. As noted in Section A, we attempt to improve classification accuracy by instead breaking articles into paragraphs, a decision we made before conducting any analysis. Figure 2 shows that our main result holds even if we use the article, instead of the paragraph as our unit of analysis.

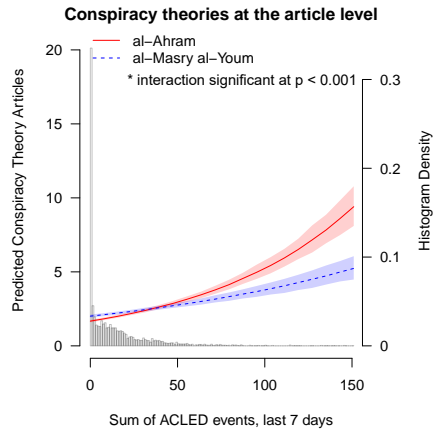


Figure 2: *Main results robust to collapsing to article*

C.3 Conspiracy Theory Framing: Only Endorsed

Our coders evaluated the framing of each article to distinguish between articles that endorsed a conspiracy theory, those that presented a conspiracy theory using neutral language, often in the context of a direct quote from a third party, or those that presented a conspiracy theory in a critical way, often through the use of sarcasm and humor. This is a difficult coding task and as we note in our paper there was less agreement between our coders on these three categories. Our purpose in coding the frame of each conspiracy theory was to avoid mistakenly counting criticism of conspiratorial thinking to be itself conspiracy theorizing. Our coders largely agreed when identifying conspiracy theory paragraphs (95.9% agreement), but struggled to agree when coding the framing (58.6% agreement). Figure 3 shows that our main result holds if we focus only on these “endorsing” articles.

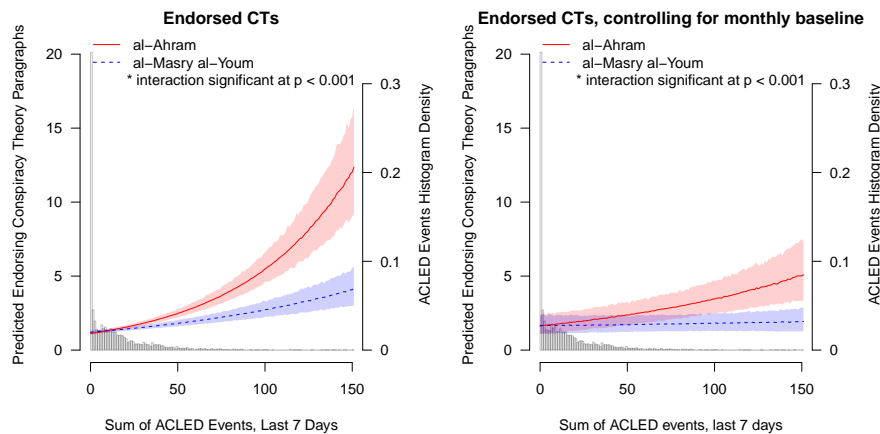


Figure 3: *Main results robust to using only articles that “endorse” conspiracy theories*

Our accuracy at predicting the framing of each conspiracy theory is lower, in large part because this is a more subtle task with greater fundamental uncertainty. We first attempted to predict all three categories — endorsing, neutral, and critical. The overall accuracy was 69.7 percent, but our

accuracy at identifying critical conspiracy theories was only 3.5 percent. Our theoretical reason for classifying conspiracy theory framing is to be sure that our results are not mistakenly driven by articles that do not promote conspiratorial thinking, so we collapsed our coding into two categories — endorsing versus critical/neutral — and achieved the same 69.7 percent accuracy (75 percent accuracy for the endorsement 63 percent accuracy for neutral/critical framing). We believe the difficulty of classifying whether conspiracy theories are endorsed may be intentional; the state may want to spread some theories while retaining plausible deniability.

C.4 Classifying articles without key words

We were concerned that perhaps our key word approach missed a significant number of conspiracy theories in articles that did not happen to use one of those key words. To allay our concerns, we apply both classifiers to the articles without key words in a second stage. The error rates above no longer apply for this set; the models will *overpredict* the prevalence of conspiracy theories in the non-keyword set because the model is trained on articles with a higher base rate of conspiracy theories. This second stage classification turned up only 315 additional conspiracy theory paragraphs in *Al-Ahram* and only 59 in *Al-Masry Al-Youm*. These numbers are small compared to 30,473 conspiracy theory paragraphs we identify in articles with our key words. Our hand inspection of these confirms that the model over-predicts conspiracy theories in this set, so we omit them from our main analysis. Including them has no substantive effect on the results reported in 2.

C.5 Including *Al-Ahram* from 1998

As we note in the main text *Al-Masry Al-Youm* began publication in 2004 and became available online in 2005. Because we are interested in a direct comparison between the two papers our main specifications include only the years in which we have data available for both newspapers (2005-2018). Figure 4 demonstrates that our main result is not dependent on excluding the years between 1998 and 2005 when we have data from *Al-Ahram* but not *Al-Masry Al-Youm*.

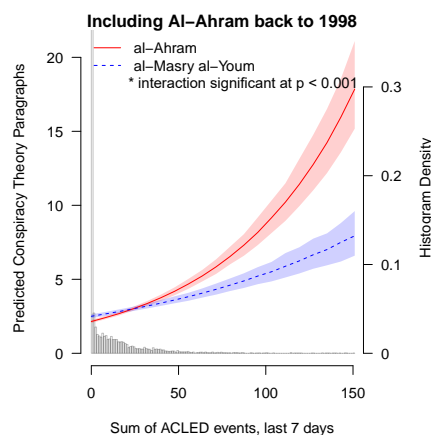


Figure 4: Main results robust to including full data from *Al-Ahram*

C.6 Individual ACLED Event Categories

The ACLED database reports its event data using six categories: battles, violence against civilians, explosions, protests, riots, and strategic events. Our main specification uses a total count of each of these categories. Figure 5 shows that our main result holds for each of these categories individually.

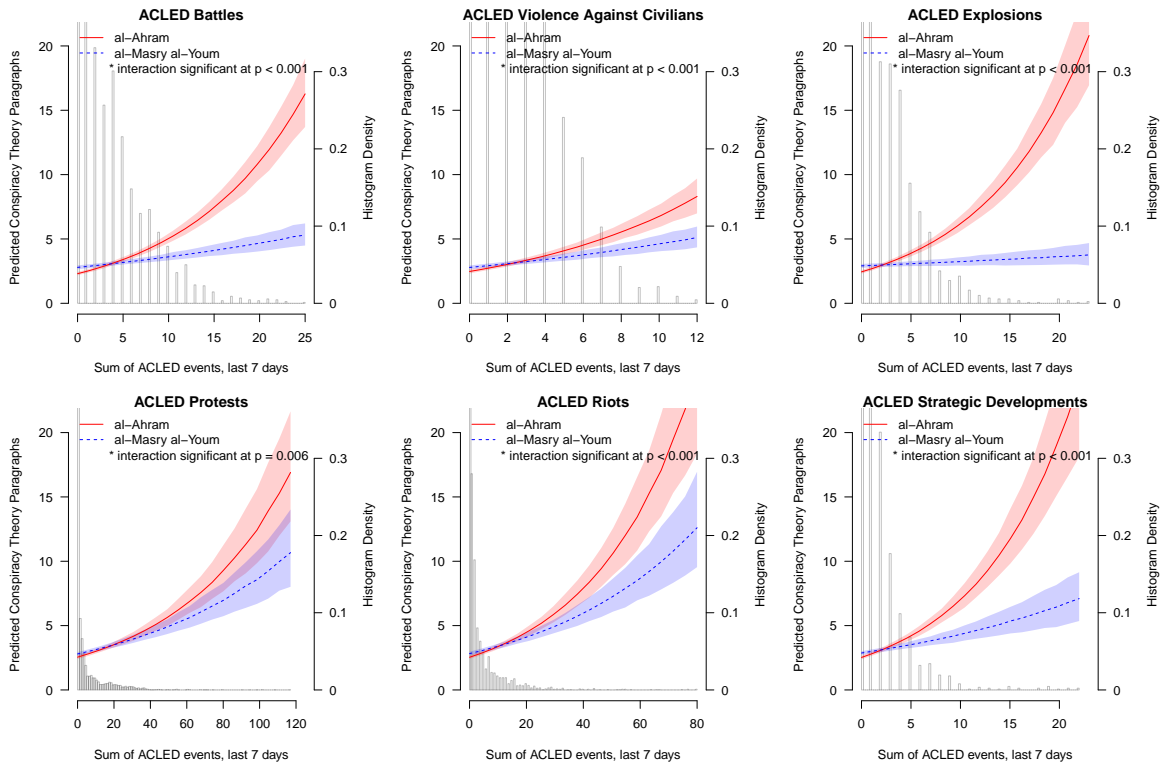


Figure 5: Main results robust different ACLED event types. The difference between Al-Ahram and Al-Masry Al-Youm is smaller for protests, but still statistically significant.

C.7 Logged ACLED Events

There is a wide variation in the count of ACLED events over the nearly twenty year period under examination. One concern is that skewness in the counts might be driving our result. To address this concern Figure 6 shows that our main result holds using a natural log +1 transformation of the ACLED event data.

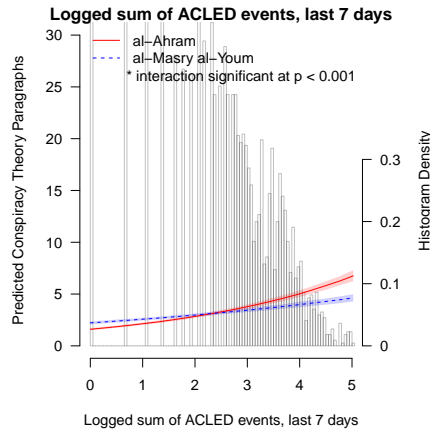


Figure 6: Main results robust to logging ACLED counts

C.8 Alternative Measure: ACLED death count

The ACLED data reports the total number of casualties for each event recorded in the dataset. It could be the case that government perception of threat differs in response to the number of casualties not the events themselves. Figure 7 shows that our main results hold if we use the ACLED data on casualties as an alternate measure of threat.

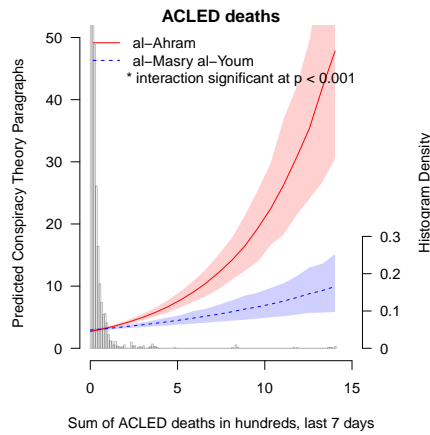


Figure 7: Main results robust to using ACLED death counts

C.9 ACLED crisis measure, 75th and 90th percentile

We consider the possibility that government perception of threat is perceived bluntly, as either low- or high-threat. Figure 8 shows that our main result is robust to an alternative measure of the ACLED event counts variable where we create a dummy variable to indicate if ACLED events are in the top 75th or top 90th percentiles.

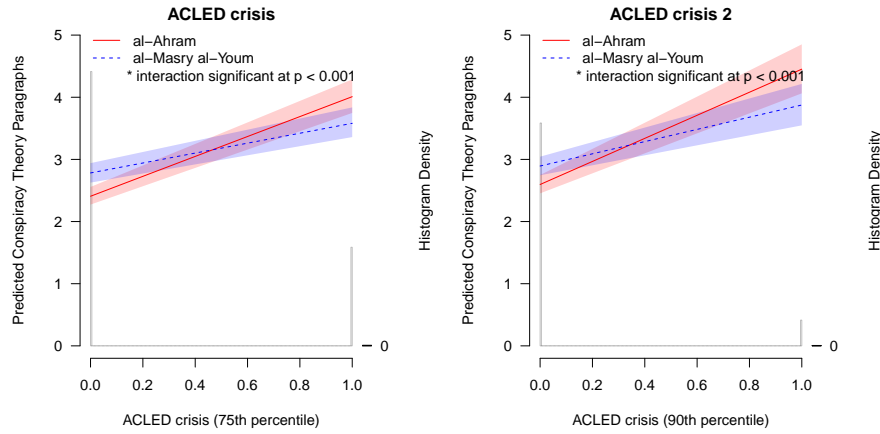


Figure 8: Main results robust to using ACLED “crisis” measure. The differences in slopes between *Al-Ahram* and *Al-Masry Al-Youm* are smaller using these blunt measure, but they are still statistically significant.

C.10 Alternative Measure: START data

We might be concerned that because the ACLED data relies on public reporting that it systematically undercounts or overcounts certain kinds of events. To address this concern we turn to the Global Terrorism Database, maintained by the START program at the University of Maryland (LaFree and Dugan, 2007). This provides an alternative measure of threat, albeit for a narrow category of events. Figure 10 shows that our main result holds using either the count of terrorist attacks or the casualty data recorded in the GTD.

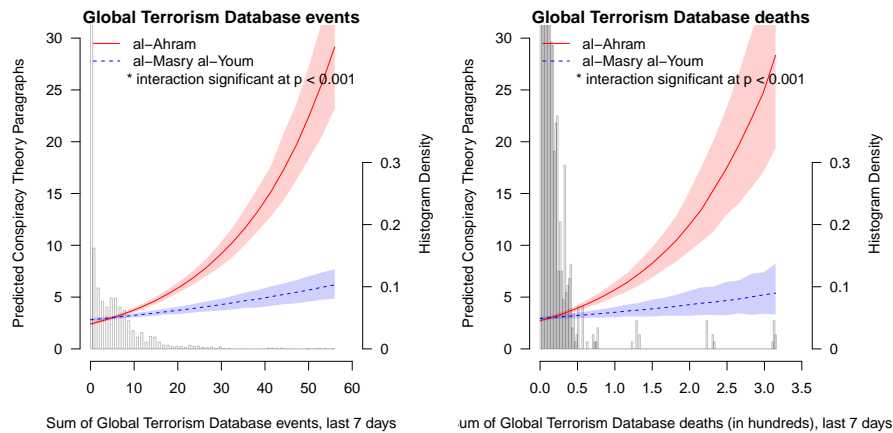


Figure 9: Main results robust to using *START’s Global Terrorism Dataset count data*

C.11 Alternative Measure: Clarke Anti-Regime Protest data, January 2012-June 2013

Clarke (2021) shows that ACLED undercounts peaceful, localized, and rural protest events in Egypt between 2012-2013. Clarke develops an improved data set of protest activity in Egypt between January 2012 and June 2013. He constructs this data set by coding articles from *Al-Masry Al-Youm* by hand, with a team of research assistants. The data are thus of very high quality, but limited to protests, not other threatening events, and limited in temporal coverage. To compare whether our results might change if we used Clarke's measure of protest as our proxy for threat, we fit a model substituting the sum of ongoing anti-regime protests in his data set for the ACLED event counts (which are correlated at 0.6). To make the results comparable, we re-estimate our main model with ACLED event counts using data for only the time in 2012-2013 covered by Clarke's data. This is a period in which our model predicts consistent undersupply by *Al-Ahram*, and the results reflect this. The undersupply is evident with either Clarke's protest data or the ACLED event counts. However, we cannot test whether our more interesting oversupply result holds from July 2013 onward because Clarke's data do not cover this time period.

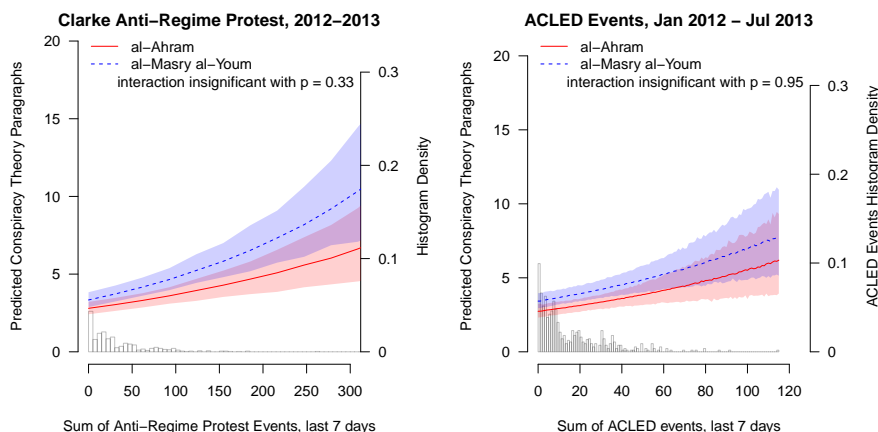


Figure 10: An improved measure of anti-regime protest from Clarke (2021) produces similar results to the ACLED event counts in the time period for which Clarke's data are available. *Note that the statistically insignificant interaction during this time-period is consistent with our theory and results reported in the paper.*

C.12 Omitting Combination of control variables

Our main regression employs two different kinds of controls to account for differences in the length of the paper across each day of the week. These include individual fixed effects for each day of the week, as well as count variables of the total number of articles and paragraphs for each day. Figure 11 shows that our main result holds if we include only these day of the week controls, or alternatively, the count variables.

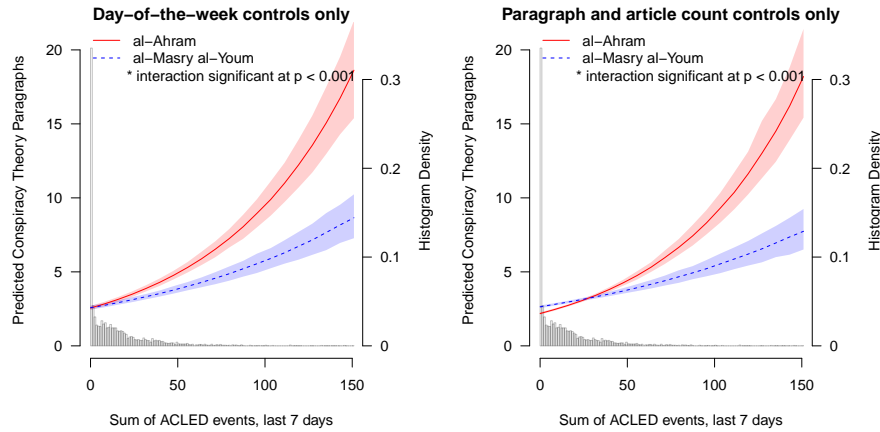


Figure 11: *Main results robust to including different temporal and corpora controls*

C.13 Administration fixed effects

We note in the paper that there we observe five different executives in our dataset: Hosni Mubarak, the SCAF, Mohamed Morsi, Adly Mansour, and Egypt’s current president Abdel Fatah al-Sisi. We might be concerned that our result is driven by the distinct political objectives of these executives and their administrations, especially given the power of the President in an authoritarian context like Egypt. Figure 12 shows that our main result holds when we include a fixed effect for each of these distinct administrations.

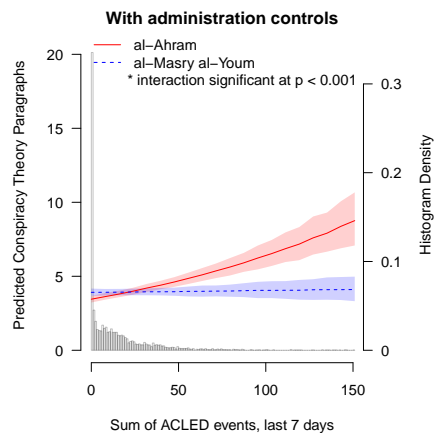


Figure 12: *Main results robust to including controls for each administration in the dataset*

C.14 ACLED Same day and Day lag

Our main specifications in the paper use a seven day moving average of ACLED events. We might be concerned that the government’s perception of threat is driven not by events over the entirety of the last week but rather that day or the previous day before publication. Figure 13 shows that our

main result holds if we use the ACLED event count the same day the newspaper was online or the day before.

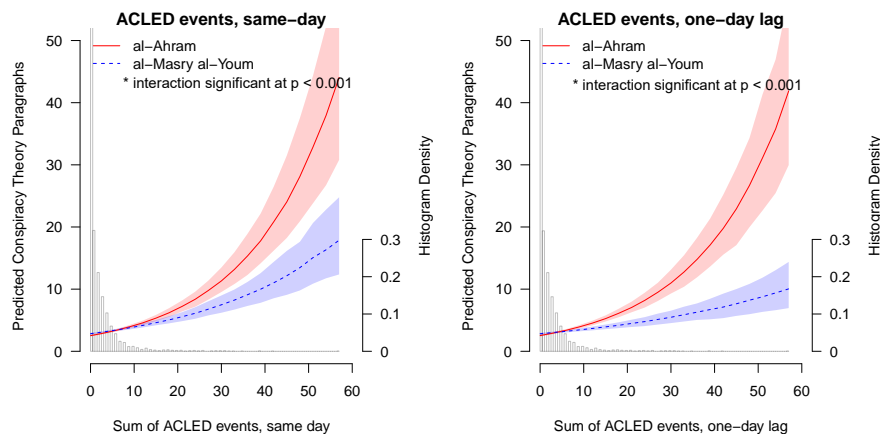


Figure 13: Main results robust to same day or one day lag of ACLED Events

C.15 Including *Al-Shuruq* and *Al-Gomhuria* Newspapers

Our analysis relies on two prominent papers one independent, *Al-Masry Al-Youm* and one government run, *Al-Ahram*. However, there are of course many other newspapers in Egypt and it is unclear without further evidence whether our argument would hold if we included additional media sources. This section briefly introduces two additional newspapers, *Al-Shuruq* (independent) and *Al-Gomhuria* (official), and presents our main specification including an additional 800,000 articles scraped and classified using the same procedure outlined in the main text

Founded in 2009 by Ibrahim al-Moallem, son of a prominent publisher, *Al-Shuruq* is by most measures the 2nd most influential independent daily newspaper in Egypt. Its reputation was forged in the aftermath of the 2011 revolution when its professionalism and relatively liberal politics, especially concerning domestic issues, led to a surge in web-traffic and circulation (Diab, 2011). Despite its independence, like all private newspapers in Egypt *Al-Shuruq* faces many of the same constraints as *Al-Masry Al-Youm* (Peterson, 2011).

Al-Gomhuria was created in the aftermath of the 1952 coup led by the Free Officers which finally forced the British from power in Egypt. Prior to 1952, most of Egypt's independent newspapers, like *Al-Ahram*, were foreign owned. *Al-Gomhuria* was explicitly created to counterbalance independent print media. *Al-Gomhuria's* first editor-in-chief was future president Anwar Sadat, and the paper rapidly developed a reputation as the mouthpiece of the Free Officers. As a national paper *Al-Gomhuria*, like *Al-Ahram*, is overseen by the National Press Authority. Unlike *Al-Ahram* whose regional prominence and importance endured even after it was nationalized by the Egyptian government, *Al-Gomhuria* has never enjoyed significant influence outside of Egypt.

Figure 14 shows that our main result is robust to including classified articles from the two additional newspapers: *Al-Shuruq* and *Al-Gomhuria*.

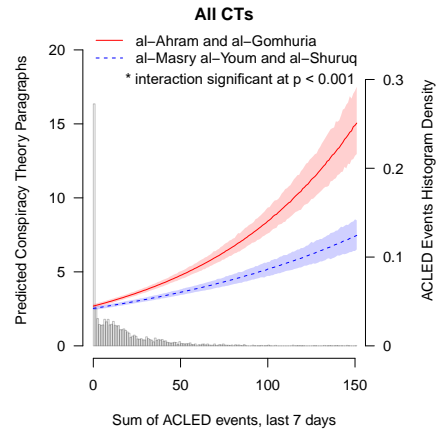


Figure 14: Main results robust to adding additional newspapers: *Al-Shuruq* (independent) and *Al-Gomhuria* (official)

References

- Clarke, Killian. 2021. “Which Protests Count? Coverage Bias in Middle East Event Datasets.” *Mediterranean Politics* pp. 1–27.
- Diab, Osama. 2011. “New Egypt, new media.”
URL: <https://www.theguardian.com/commentisfree/2011/mar/10/egypt-media-newspapers-mubarak-propaganda>
- Kuhn, M, Jed Wing, Stew Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, B Kenkel, RC Team et al. 2014. “caret: Classification and regression training. R package version 6.0–21.” *CRAN: Wien, Austria* .
- LaFree, Gary and Laura Dugan. 2007. “Introducing the global terrorism database.” *Terrorism and political violence* 19(2):181–204.
- Oliver, J. Eric and Thomas J. Wood. 2014. “Conspiracy theories and the paranoid style(s) of mass opinion.” *American Journal of Political Science* 58(4):952–966.
- Peterson, Mark Allen. 2011. “Egypt’s media ecology in a time of revolution.” *Arab Media & Society* 14:1–16.