Supplemental Information

for

The Supply of Conspiracy Theories in State-controlled Media

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

May 13, 2021

For online publication only.

Abstract

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

[*]Postdoctoral Fellow, Brown University, gabriel_koehler-derrick@brown.edu
[†]Associate Professor, Massachusetts Institute of Technology, rnielsen@mit.edu
[‡]Postdoctoral Fellow, Harvard University, dromney@fas.harvard.edu

Contents

A Automated Conspiracy Theory Detection

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

The default unit in most text analysis settings is each natural text, in our case, each news article. However, we found that classifying entire articles as conspiracy theories or not resulted in substantial measurement error. Conspiracy theory language is typically a relatively small part of any single article, so we improved our classification accuracy by instead breaking articles into paragraphs. This results in roughly 10.5 million paragraphs that we classify.

Conspiracy theories are relatively rare in our data. Training our model with a simple random sample of paragraphs would have been intractable. Approximately one out of every 300 paragraphs in these newspapers is conspiracy theory-related; to obtain even 50 examples of conspiracy theory-related paragraphs using standard approaches, we would have had to sample and code 15,000 paragraphs by hand.

Instead, we exploited the fact that Egyptian journalists use formulaic and consistent vocabulary when discussing conspiracy theories. We identified a list of 18 keywords that are conspiracy-related including variants of: "conspiracy" (e.g. المؤامرة, التآمر, etc.); "trick" or "machination" (e.g. مكيدة, دسيسة, etc.) and "collusion" (e.g. التواطؤ, متواطئ, etc.). 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.[1] Next we took a sample of 1,500 paragraphs that contain these keywords, had two research assistants separately classify these paragraphs as conspiracy or non-conspiracy, as well as identify the perpetrator and victims in each conspiracy theory. After a reconciliation process we used these paragraphs as our training sets for the classifier that ran on paragraphs containing the conspiracy keywords—to further limit our analysis to political conspiracies of interest. The human coding by two RAs was based on (1) the presence of the keywords identified above, (2) Sufficient length for the coders to make a determination about the nature of the paragraph, usually at least 25 words—this is because it is difficult (though not impossible) for very short paragraphs to meet our definition of a conspiracy theory—and (3) we also excluded references which used our keywords but in the context of the arts, sports, and historical events prior to the 1900s.

Our goal is to accurately measure the incidence of references to conspiracy theories in news articles by *Al-Ahram* and *Al-Masry Al-Youm*. In principle, this measurement problem is relatively straightforward; we identify content that satisfies our definition of a conspiracy theory: paragraphs that contain our key words referencing "conspiracy" or "machinations," are of sufficient length for our human coders to determine whether the author is presenting a conspiracy in the text, and not in our excluded categories of art, sports or historical events before 1900. This is a reasonable task for a human coder, although it does take time and and invariably there are some differences of opinion between coders, even among paragraphs that meet all of the criteria above. For example, a clear example of a conspiracy paragraph was 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…

On the other hand, 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 ولايزالون عملاء الخارج وأن هناك تفاهما بينهم وبين واشنطن مثلا على حل القضية الفلسطينية على حساب مصر والتنازل عن جزء من سيناء.

It is also somewhat difficult to code off-hand references to conspiracy theories. A single word might evoke a network of meanings for readers "in the know." These problems all relate to conceptual issues of measurement: How would we know conspiracy theory language when we see it?

A.1 Instructions to coders

We developed a set of guidelines to guide our two coders as they hand-coded our labeled set. Existing definitions used by previous scholars characterize conspiracy theories as highlighting the role of "unseen and malevolent forces," providing an interpretation of events using "Manichean" language, and finally discrediting "mainstream" explanations (Oliver and Wood, 2014). However, our approach is more precisely characterized as identifying conspiratorial language, based on our understanding that the politics of conspiracies may not necessarily require the author to explicitly provide a comprehensive "theory." We see this approach as a necessary adaption of accepted definitions of conspiracy theories to our particular context. Thus, our strategy focused on determining a set of guidelines to weed out clear non-conspiracies from the set of paragraphs containing our keywords, rather than on identifying Manichean language.

We developed coding criteria for our coders by reading hundreds of conspiracy theories in both newspapers outside the training set ourselves. As part of this process we observed that conspiracy related content appeared in some surprising parts of the newspaper. We regularly encountered our keywords in discussions of movies, art, and books, as well as in the sports section, and often in discussions of historic events, particularly religious discussions about the early Muslim community. While all of these articles employed our keywords we instructed the coders to exclude these categories as they were less relevant to our focus on contemporary conspiracy theories. Another insight from this process was the crucial decision to instruct the coders not to evaluate the conspiracy theories based on their veracity, as this would inevitably have introduced subjectivity and bias into the coding process.

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. This task was fundamentally more nuanced and therefore we find less agreement between the coders. We report a robustness check with this variable below.

[1] For example, we considered using other keywords, such as iterations of the words for "plan" (خطط) and "interference" (التدخل), but an examination of the paragraphs in which these words were used showed that they were more often than not about subjects totally unrelated to our focus, such as urban planning.

A.2 Classifier details

We train our classifier on 22,190 paragraphs from the 1,500 articles that our research assistants coded by hand. 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 chose 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 and a specificity of 0.99.

With this labeled set in hand, we then used a random forest classifier to classify the remaining 449,297 paragraphs in the 18,087 articles that contained our keywords. Combining our labeled set with the set predicted by the classifier, 27,039 (82.2%) of our original 33,720 paragraphs remained marked as conspiracy theories.

B Regression Results

Table 1 shows the full regression results corresponding to Figure 3 in the paper.

	(1) All CTs	(2) All CTs	(3) Endorsed CTs	(4) Endorsed CTs
Intercept	0.04	-0.49*	-0.67*	-1.01*
	0.067	0.14	0.11	0.21
ACLED Events	0.007*	-0.002	0.008*	0.001
	0.0007	0.001	0.001	0.002
Al-Ahram	-0.18*	-0.14*	-0.069	-0.002
	0.027	0.027	0.044	0.043
ACLED Events × Al-Ahram	0.007*	0.006*	0.008*	0.006*
	0.0009	0.0009	0.001	0.001
Paragraphs	-0.0001	0.0006*	-0.0003*	0.0004*
	0.00007	0.00007	0.0001	0.0001
Articles	0.008*	0.002*	0.01*	0.002*
	0.0007	0.0007	0.001	0.001
News Agency Paragraphs	-0.0009	-0.001*	0.001	0.002*
	0.0005	0.0005	0.0008	0.0008
News Agency Articles	-0.004	0.012*	-0.024*	-0.008
	0.004	0.004	0.006	0.006
Saturday	0.25*	0.23*	0.21*	0.21*
	0.036	0.033	0.058	0.054
Sunday	0.099*	0.17*	0.097	0.18*
	0.037	0.034	0.059	0.056
Monday	0.15*	0.23*	0.13*	0.24*
	0.036	0.034	0.059	0.056
Tuesday	0.16*	0.24*	0.15*	0.26*
	0.036	0.034	0.059	0.056
Wednesday	0.13*	0.23*	0.13*	0.27*
	0.037	0.034	0.059	0.056
Thursday	0.13*	0.21*	0.042	0.13*
	0.036	0.034	0.059	0.056
Year-Month Fixed Effects	No	Yes	No	Yes
N	8,805	8,805	8,805	8,805

Table 1: *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 1 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.

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.

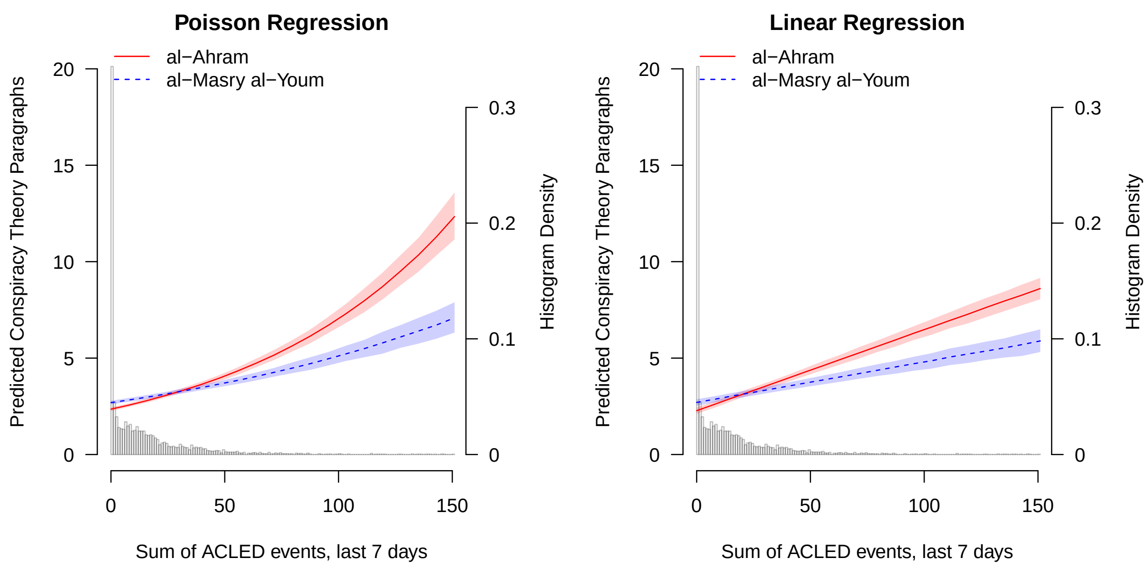

Figure 1: *Main results robust to using Poisson or OLS regression*

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.

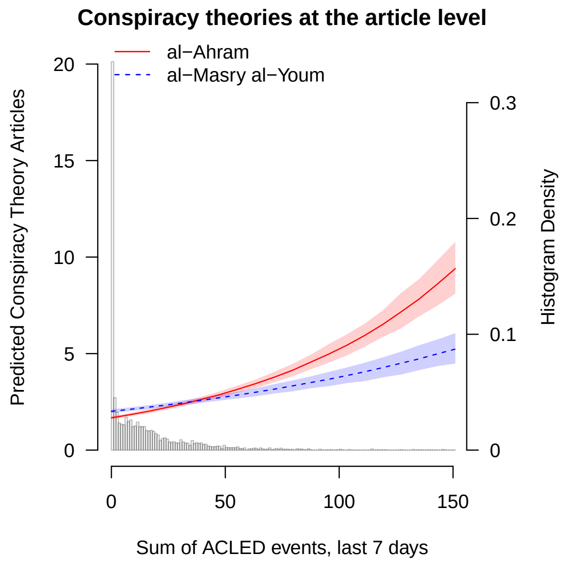

Figure 2: *Main results robust to collapsing to article*

C.3 Conspiracy Theory Framing: Only Endorse

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.

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). This accuracy remains lower than we hoped, so we report our main models below with two measures of conspiracy theories: the count of all conspiracy theory paragraphs per day and the count of only those we classify as endorsed. 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.

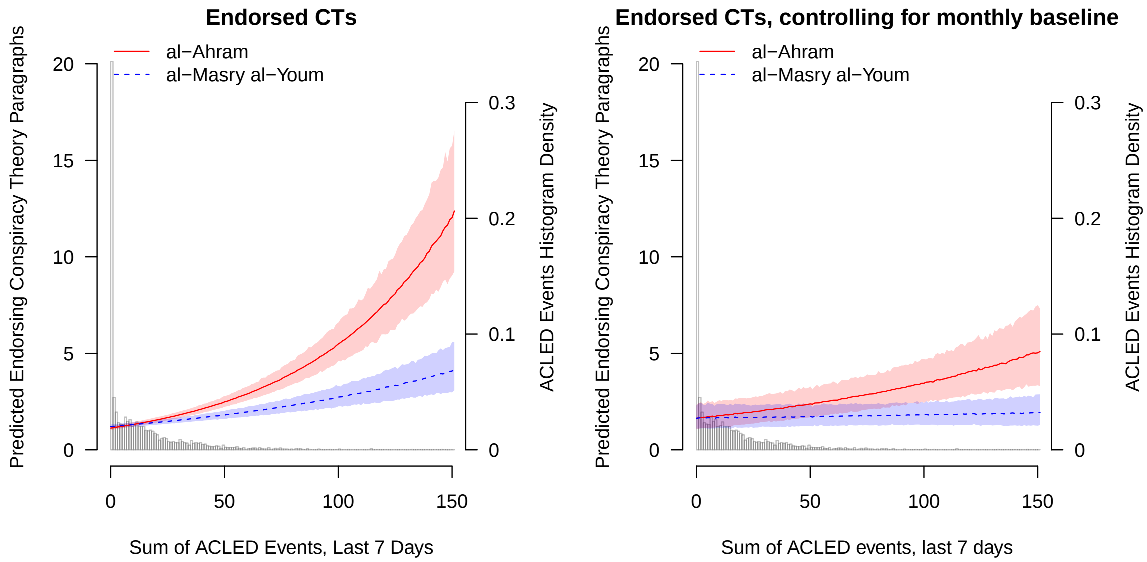

Figure 3: *Main results robust to using only articles that "endorse" conspiracy theories*

C.4 Texts without keywords

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. To our relief, 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 1.

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*.

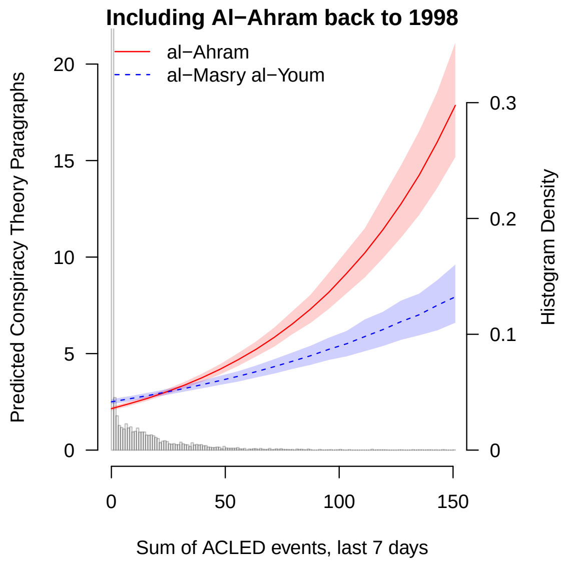

Figure 4: *Main results robust to including full data from Al-Ahram*

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.

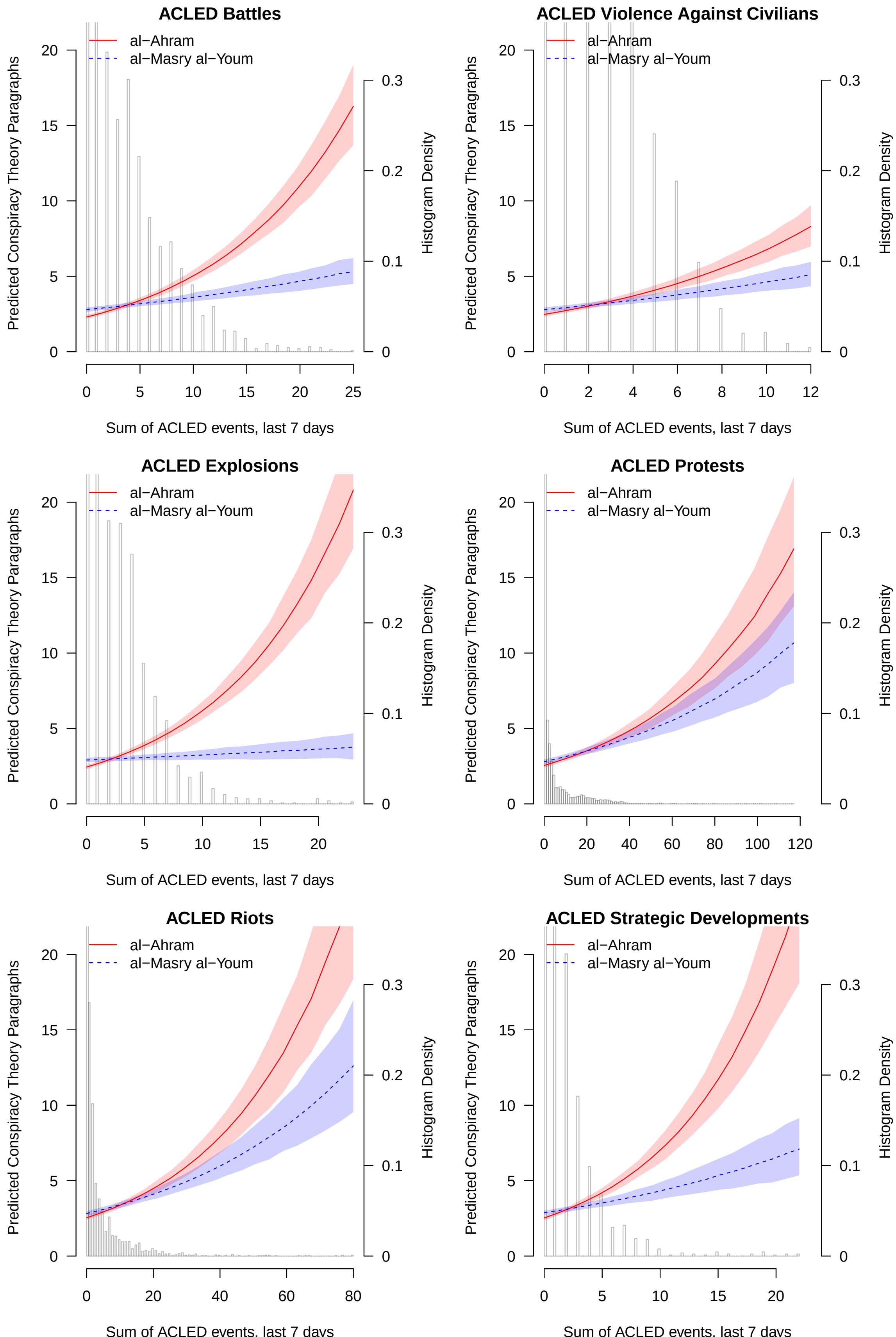

Figure 5: *Main results robust different ACLED event types*

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.

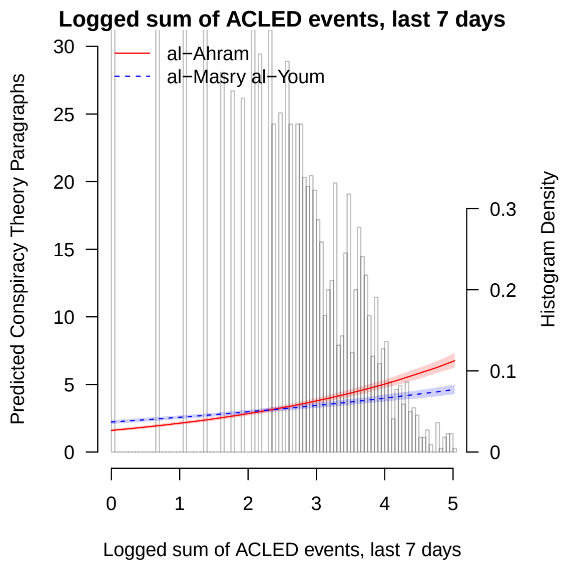

Figure 6: *Main results robust to logging ACLED counts*

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.

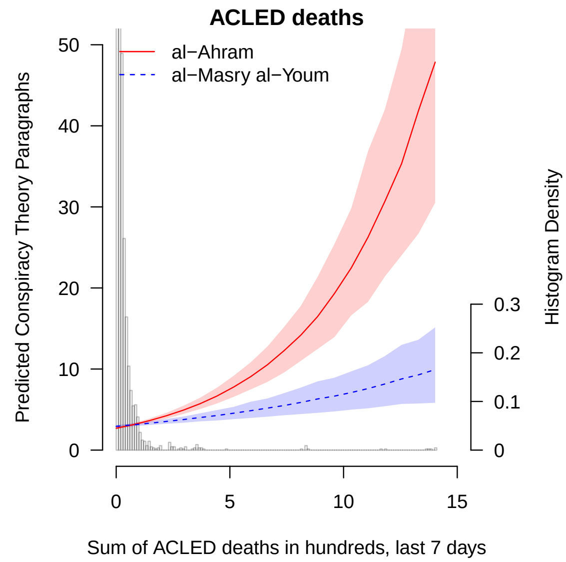

Figure 7: *Main results robust to using ACLED death counts*

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.

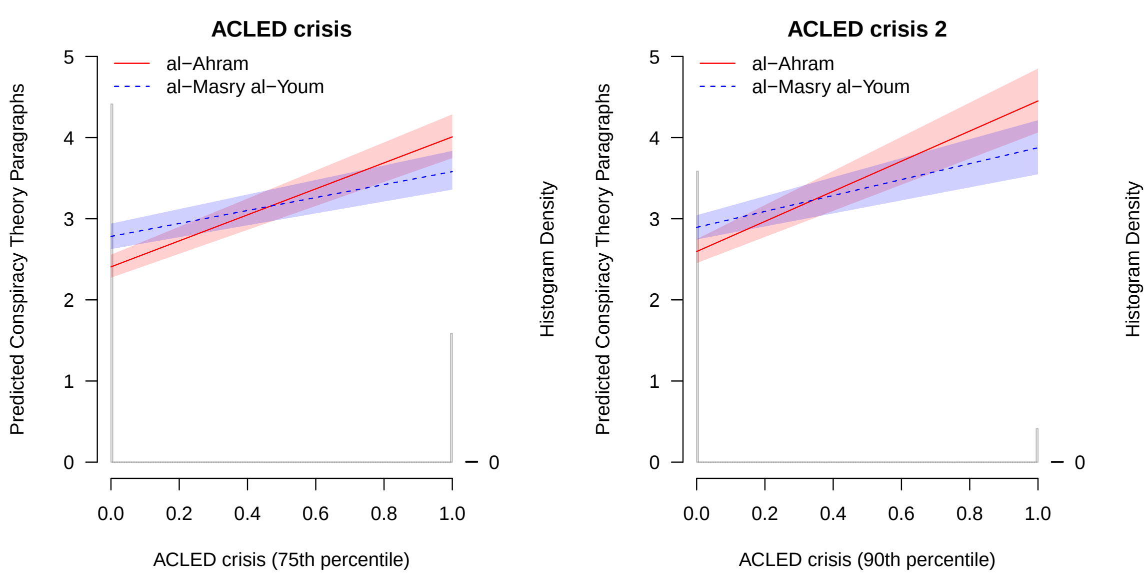

Figure 8: *Main results robust to using ACLED "crisis" measure*

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 9 shows that our main result holds using either the count of terrorist attacks or the casualty data recorded in the GTD.

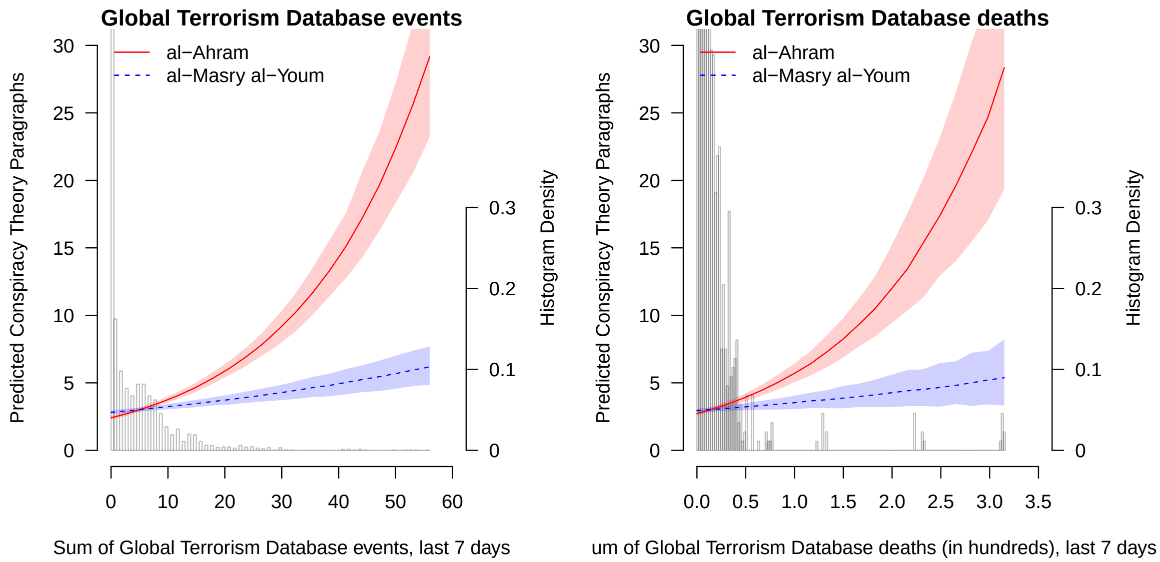

Figure 9: *Main results robust to using START's Global Terrorism Dataset count data*

C.11 Alternative Measure: 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 10 shows that our main result holds if we include only these day of the week controls, or alternatively, the count variables.

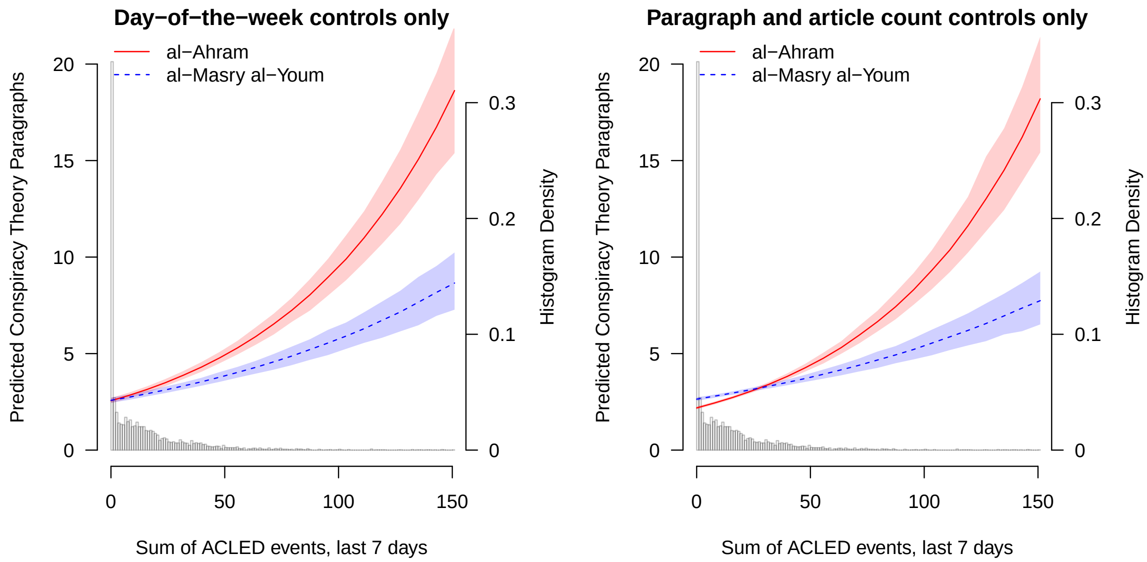

Figure 10: *Main results robust to including different temporal and corpora controls*

C.12 Alternative Measure: 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 11 shows that our main result holds when we include a fixed effect for each of these distinct administrations.

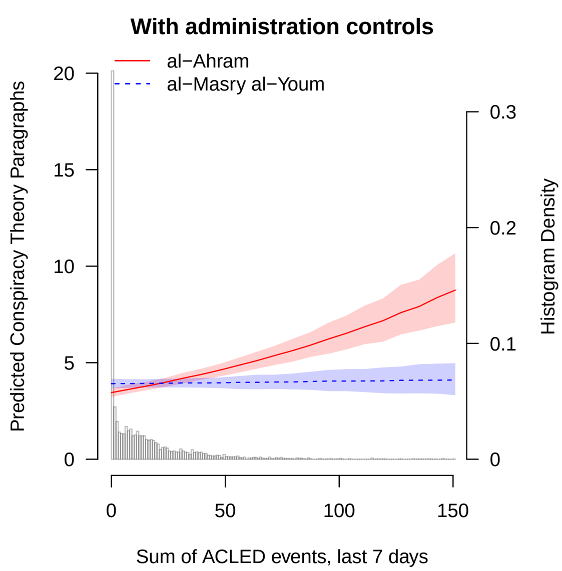

Figure 11: *Main results robust to including controls for each administration in the dataset*

C.13 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 12 shows that our main result holds if we use the ACLED event count the same day the newspaper was online or the day before.

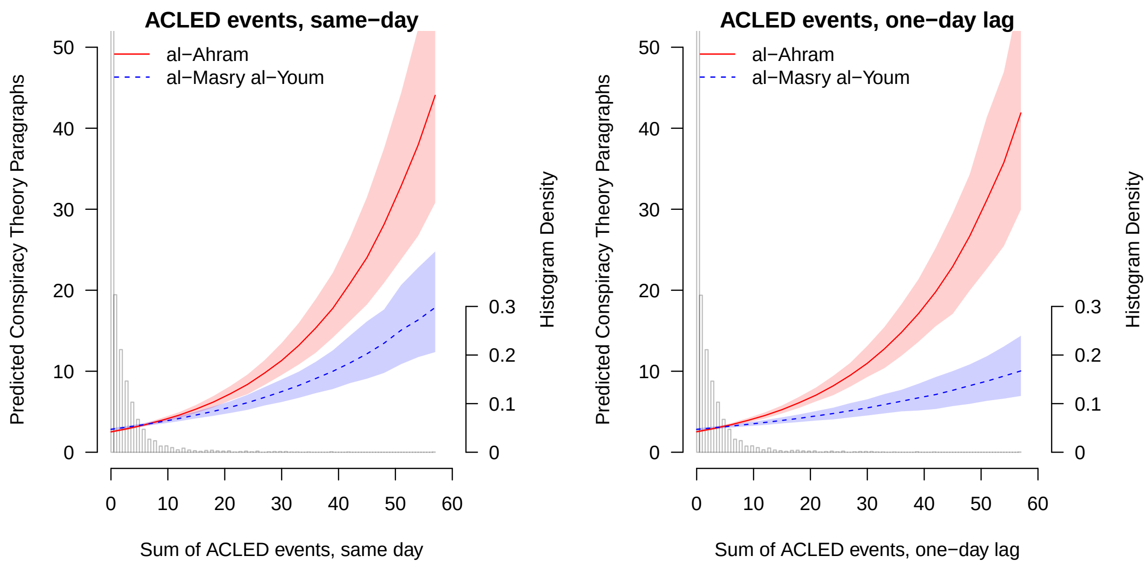

Figure 12: *Main results robust to same day or one day lag of ACLED Events*

D Generalizability: Background on *Al-Shuruq* and *Al-Gomhuria*

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.

D.1 *Al-Shuruq* and *Al-Gomhuria* Results

Figure 13 shows that our main result is robust to including classified articles from two additional newspapers: *Al-Shuruq* and *Al-Gomhuria*.

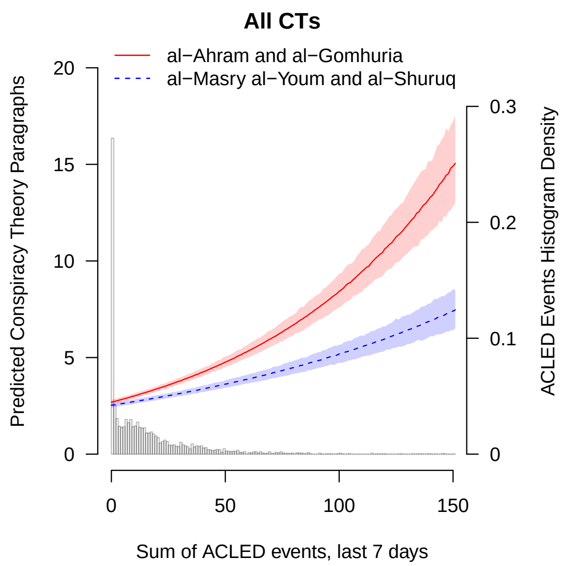

Figure 13: *Main results robust to adding additional newspapers: Al-Shuruq (independent) and Al-Gomhuria (official)*

E Which authors supply conspiracy theories?

Our first observation concerns how conspiracy theories tend to end up in Al-Ahram. We observed that a relatively small number of authors write conspiracy theories quite frequently, while the rest write them rarely. We calculate the percent of each author's articles that include a conspiracy theory and identify 175 authors who use conspiracy theories in at least 10 percent of their articles. We find that these authors are more likely to write when the number of ACLED events was high the previous day. By itself, this does not reveal the intent of the government, but it does suggest one of its means for communicating conspiracy theories: a particular set of authors prone to conspiracy theorizing are mobilized to write in *Al-Ahram* as the number of violent events increases.

F Additional Details about the Oversupply Qualitative Analysis

Table 2 shows that references to perpetrators, victims, and incidents from the prior two weeks of ACLED were present only 13% of the time for the perpetrator, 6% of the time for the victim,

In ACLED Prior Two Weeks:	Articles (141 total)	Days (30 total)
Perpetrators	18 (13%)	8 (27%)
Victims	9 (6%)	6 (20%)
Incidents	13 (9%)	9 (30%)

Table 2: *Oversupplied Conspiracy Theories in Al-Ahram rarely mention incidents or actors involved in recent events recorded in the ACLED data.*

and 9% of the time for the incident. We interpret the fact that the vast majority of *Al-Ahram's* conspiracy theories on these days have little connection to recent events, to be consistent with the idea that the state seeks to deflect the public's attention rather than craft a (counter)narrative about events as they occur.

References

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.