Chinese Investment and Elite Sentiment in Southeast Asia: An Event Study of Influence Along the Belt and Road *

Yining Sun, Ethan B. Kapstein, and Jacob N. Shapiro

Princeton University

September 5, 2023

Contents

1 The Dataset of Chinese Investment and Politician Sentiment **7**

1.1 Events: South China Sea Disputes 7

1.2 Treatment: Chinese Projects 7

1.2.1 Validation of Chinese projects 8

1.2.2 Types of Chinese projects 9

1.2.3 Total value of Chinese projects 11

1.2.4 Time range of Chinese projects 14

1.2.5 Geolocations of Chinese projects 17

1.3 Outcome: Elite Sentiment 19

1.3.1 Text processing pipeline 19

1.3.2 Sample social media posts and machine translation 23

1.3.3 Sentiment analysis algorithms: Google NLP API and SpaCy 28

1.3.4 Histogram of politician sentiment 28

1.3.5 Time trend of politician sentiment toward China 30

1.3.6 Spatial distribution of identified politicians that engage in China-related conversation 37

*The authors thank Cedrik Forbes, Fikry A Rahman, and Mutti Anggitta for invaluable research assistance. Alicia R. Chen contributed to the data development. This project received generous support from the Department of Defense. All errors are our own.

1.4 Unit of Analysis 42
1.4.1 Constituency-date level 42
1.4.2 Politician-date level 43

2 Descriptive Analysis **44**
2.1 Summary Statistics 44
2.2 Content Classification 44
2.3 Number of Posts by Chinese Investment Around Shocks 48
2.4 Elite Sentiment by Chinese Investment Around Shocks 59

3 Regression Results **71**
3.1 Google NLP API Sentiment 71
3.2 SpaCy Sentiment 79

4 Robustness Checks **87**
4.1 Unit Fixed Effects Only 87
4.2 Event Windows 95
4.3 Placebo Tests 103
4.3.1 Shock = One day after the event date 103
4.3.2 Shock = One day before the shock 111
4.4 Alternative Measurement of Chinese Investment 119

5 Heterogeneous Effects **127**
5.1 Incumbent vs. Challengers 127
5.2 Public Transportation vs. Powerplants 129

6 Power Analysis **131**

List of Tables

A-1 List of South China Sea Disputes 7
A-2 Comparison of Datasets on Chinese Investment in Southeast Asia 8
A-3 Politicians with Identified Social Media Handles 20
A-4 List of Filtering Keywords 20
A-5 Summary Statistics of China-related Social Media Posts 21
A-6 Summary Statistics of Politicians that Switch Constituencies 37
A-7 Summary Statistics of Outcome Variables in the Stacked Event Study Design 45

List of Figures

A-1 Types of Chinese Projects 10

A-2 Total Value of BRI Projects 11

A-3 Time Distribution of BRI Projects 14

A-4 Geolocations of Chinese Projects 17

A-5 Text Processing Pipeline 19

A-6 Descriptive Analysis of Politicians that Mentioned China on Social Media 22

A-7 Sample tweets in the Philippines 24

A-8 Machine translation examples 26

A-8 Machine translation examples (continued) 27

A-9 Histogram of Politician Sentiment 29

A-10 Time Trend of Politician Sentiment toward China 31

A-11 Politician Sentiment Toward China Before and After COVID Shock 33

A-12 Politician Sentiment Toward China Before and After COVID Shock: Condition on BRI Treatment 35

A-13 Map of Politicians Engaged in China-related Topics 39

A-14 Top 12 Categories of Politician Posts about China in General 46

A-15 Time Trend of Top 12 Categories of Politician Posts about China 47

A-16 Categories of Politician Posts about China Around Shocks 48

A-17 Number of Posts about China Grouped by Project Treatment Status: Unbalanced Panel 50

A-18 Number of Posts about China Grouped by Project Treatment Status: Balanced Panel 55

A-19 Elite Sentiment Toward China Grouped by Project Treatment Status: Unbalanced Panel 61

A-20 Elite Sentiment Toward China Grouped by Project Treatment Status: Balanced Panel 66

A-21 Stacked Event Study on Elite Social Media Posts about China: constituency-date level 71

A-22 Stacked Event Study on Elite Social Media Posts about China: politician-date level 73

A-23 Stacked DID on Elite Social Media Posts about China: Constituency-date level . . . 75

A-24 Stacked DID on Elite Social Media Posts about China: Politician-date level 77

A-25 Stacked Event Study on Elite Social Media Posts about China (SpaCy Sentiment): constituency-date level . . . 79
A-26 Stacked Event Study on Elite Social Media Posts about China (SpaCy Sentiment): politician-date level . . . 81
A-27 Stacked DID on Elite Social Media Posts about China (SpaCy Sentiment): Constituency-date level . . . 83
A-28 Stacked DID on Elite Social Media Posts about China (SpaCy Sentiment): Politician-date level . . . 85
A-29 Stacked Event Study: With Only Constituency Fixed Effect . . . 87
A-30 Stacked Event Study: With Only Politician Fixed Effect . . . 89
A-31 Stacked Difference-in-difference: With Only Constituency Fixed Effect . . . 91
A-32 Stacked Difference-in-difference: With Only Politician Fixed Effect . . . 93
A-33 Stacked Event Study: 28-day event window . . . 95
A-34 Stacked Event Study at Politician-date Level: 28-day event window . . . 97
A-35 Stacked Difference-in-Difference: 28-day event window . . . 99
A-36 Stacked Difference-in-Difference at Politician-date Level: 28-day event window . . . 101
A-37 Stacked Event Study: One Day After the Event Date . . . 103
A-38 Stacked Event Study At Politician-Date Level: One Day After the Event Date . . . 105
A-39 Stacked Difference-in-Difference: Placebo Event Date = one day after the shock . . . 107
A-40 Stacked Difference-in-Difference at Politician-date Level: Placebo Event Date = one day after the shock . . . 109
A-41 Stacked Event Study: Placebo Event Date = one day before the shock . . . 111
A-42 Stacked Event Study at Politician-date Level: Placebo Event Date = one day before the shock . . . 113
A-43 Stacked Difference-in-Difference: Placebo Event Date = one day before the shock . . . 115
A-44 Stacked Difference-in-Difference at Politician-date Level: Placebo Event Date = one day before the shock . . . 117
A-45 Stacked Difference-in-Difference at Constituency-date Level: Logged Cumulative Value of Chinese Projects . . . 119

A-46 Stacked Difference-in-Difference at Politician-date Level: Logged Cumulative Value of Chinese Projects 121
A-47 Stacked Difference-in-Difference at Constituency-date Level: BRI Value Threshold of Chinese Projects 123
A-48 Stacked Difference-in-Difference at Politician-date Level: BRI Value Threshold of Chinese Projects 125
A-49 Incumbent vs. Challengers: Politician-date level 127
A-50 Public Transportation vs. Powerplants: Constituency-date level 129
A-51 Power Plots: Minimum Detectable Effect 131

1 The Dataset of Chinese Investment and Politician Sentiment

1.1 Events: South China Sea Disputes

Table A-1: List of South China Sea Disputes

Country	Date	Event
The Philippines	April 11, 2012	Scarborough Shoal standoff
	July 12, 2016	South China Sea arbitration results
	May 4, 2021	The tweet of the Philippine Foreign Affairs Secretary
Malaysia	June 2, 2015	Chinese Coast Guard Vessels at Luconia Shoals
	June 1, 2021	Chinese Air Force Flyby
	April 10, 2020	Standoff on island Borneo
Indonesia	March 19, 2016	Detention of Chinese fisherman around Natuna island
	January 1, 2021	Chinese fishing boat trespassing

Table A-1 lists the prominent events of South China Sea disputes for each country. Events are selected based on three criteria: 1) There were clashes or confrontations between China and either country in our sample about South China Sea authority; 2) The event triggered diplomatic moves and nationwide discussion in the country of dispute; 3) The event received international attention and coverage (i.e., there were at least three articles from international news sources). As discussions on social media platforms are prompt to news about territorial disputes, we documented the event dates by the day of the first disclosures from the national media.

Although there is a possibility that Southeast Asian countries choose to disclose information about conflict over the South China Sea with the Chinese government at a particular time to advance the national agenda, we argue that for each constituency at the local level, the incidences and the announcement of the conflicts are exogenous to whether the constituency had received BRI projects or not.

1.2 Treatment: Chinese Projects

Our main treatment variable measures China's commercial activities in respective countries. We define Chinese projects as any infrastructure projects that are either financed by the Chinese government and banks or constructed and implemented by Chinese companies.

We study concrete infrastructure projects instead of financial transactions for several reasons. First, unlike one-time financial transfers, public infrastructure projects take years to implement. Second, constructions take place in specific areas with clearly-staged goals, while the usage of

financial flows is generally vague and not open to the public. With larger political salience, local residents and politicians are more likely to associate Chinese infrastructure projects in their regions with BRI investment and attribute potential economic and political benefits to China. Figure A-1 displays major types of BRI projects in each country. Figure A-2 and A-3 visualize the spending and time span of all BRI projects respectively.

1.2.1 Validation of Chinese projects

We cross-validated the information about the infrastructure projects recorded in the Reconnecting Asia Database (Center for Strategic and International Studies, 2020), AidData's Global Chinese Development Finance Dataset Version 2.0 (Dreher et al., 2022), and the Geolocated Dataset of Chinese Overseas Development Finance by Boston University (Ray et al., 2021), and obtained a comprehensive list of projects that are either Chinese-funded or Chinese-built in the Philippines, Malaysia, and Indonesia.

Table A-2: Comparison of Datasets on Chinese Investment in Southeast Asia

	Our Validated Dataset	AidData	Boston University
The Philippines			
Total number of projects	46	40	8
Time frame of projects	2000-2022	2000-2017	2009-2019
Projects with geolocations	44	26	7
Major types of projects	Powerplants Transportation	Education Agriculture	Transportation Irrigation
Malaysia			
Total number of projects	34	20	1
Time frame of projects	1994-2020	2004-2017	2016
Number of projects with geolocations	34	17	1
Major types of projects	Powerplants Transportation	Construction Education	Transportation
Indonesia			
Total number of projects	52	182	30
Time frame of projects	2003-2019	2002-2017	2008-2017
Number of projects with geolocations	52	107	30
Major types of projects	Powerplants Transportation	Powerplants Transportation	Powerplants Transportation

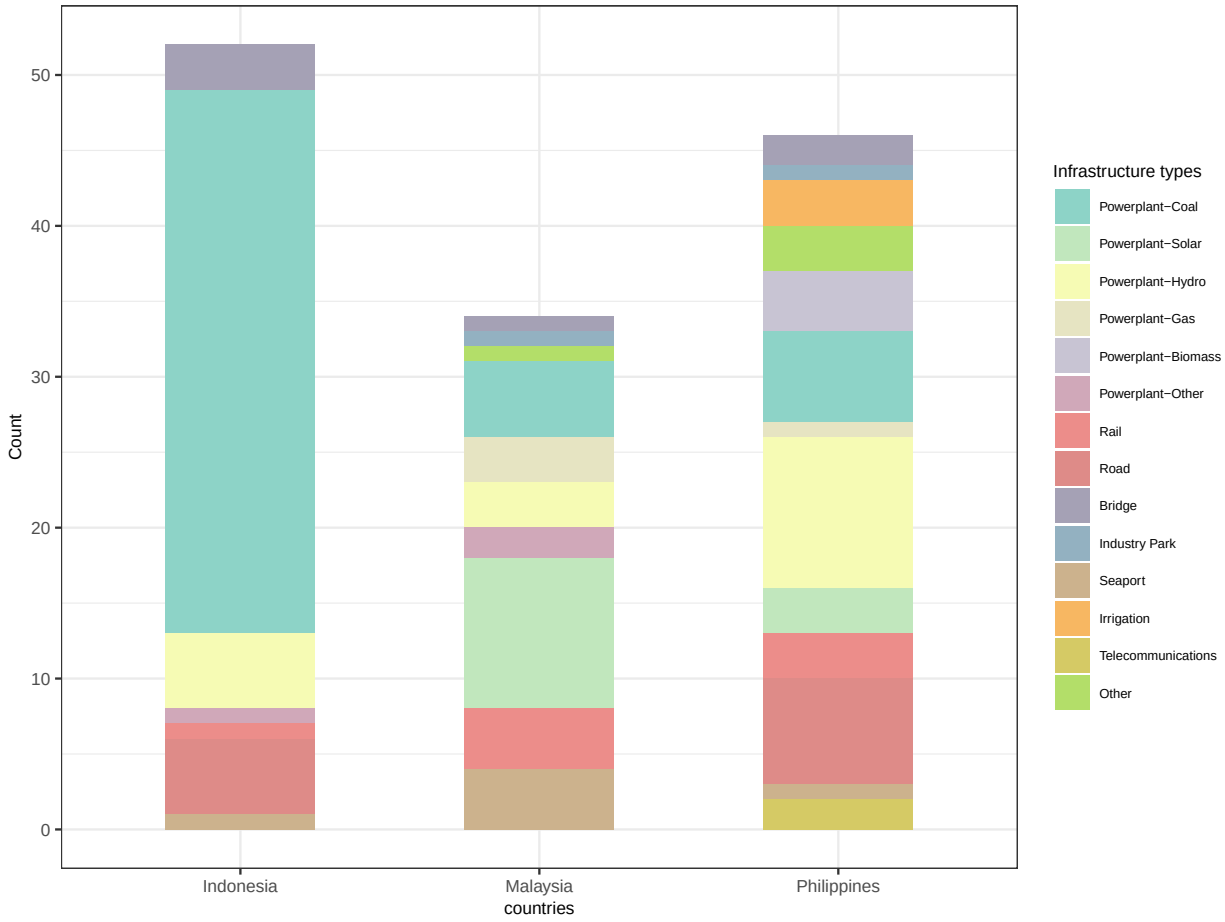
Data sources: The metrics for AidData is from AidData's Global Chinese Development Finance Dataset Version 2.0 (Dreher et al., 2022), after applying the filters of "non-umbrella project" and "recommended for aggregates" and excluding all emergency reliefs. The metrics for Boston University is calculated from the raw data of the Geolocated Dataset of Chinese Overseas Development Finance by Boston University (Ray et al., 2021). Our validated dataset is based on infrastructure projects recorded in the Reconnecting Asia Database (Center for Strategic and International Studies, 2020). We kept projects that are either financed or implemented by Chinese-related entities, and complemented that with infrastructure projects from AidData and the dataset of Boston University.

Our local research assistants then conducted thorough desk research to fill in details about the project type, total costs, financing and implementation agencies, locations, and start and finish dates. We also filtered out duplicated entries of projects and combined multiple phases of the same project into one unique entry. As shown in the Table A-2, our measurement of China's commercial activities in Southeast Asian countries covers more Chinese activities in the Philippines and Malaysia. In Indonesia, there are 182 entries identified by the AidData excluding emergency relief. After manual cleaning to remove duplicated projects, there are 44 projects in Indonesia from AidData, while we identified 52 projects after cross-validation.

1.2.2 Types of Chinese projects

Most of the Chinese projects in Indonesia are Coal Powerplants, followed by Hydro and Biomass Powerplants. There are also 6 railway or road projects and 1 seaport project in Indonesia. Powerplants also also dominates the projects in Malaysia, with a half of the powerplants using solar energy. In the Philippines, around half of the projects are powerplants, and the other half spans from railway, road, bridge, industry parks, irrigation projects, and telecommunications. Figure A-1 shows the details.

Figure A-1: Types of Chinese Projects



1.2.3 Total value of Chinese projects

Figure A-2: Total Value of BRI Projects

(a) The Philippines

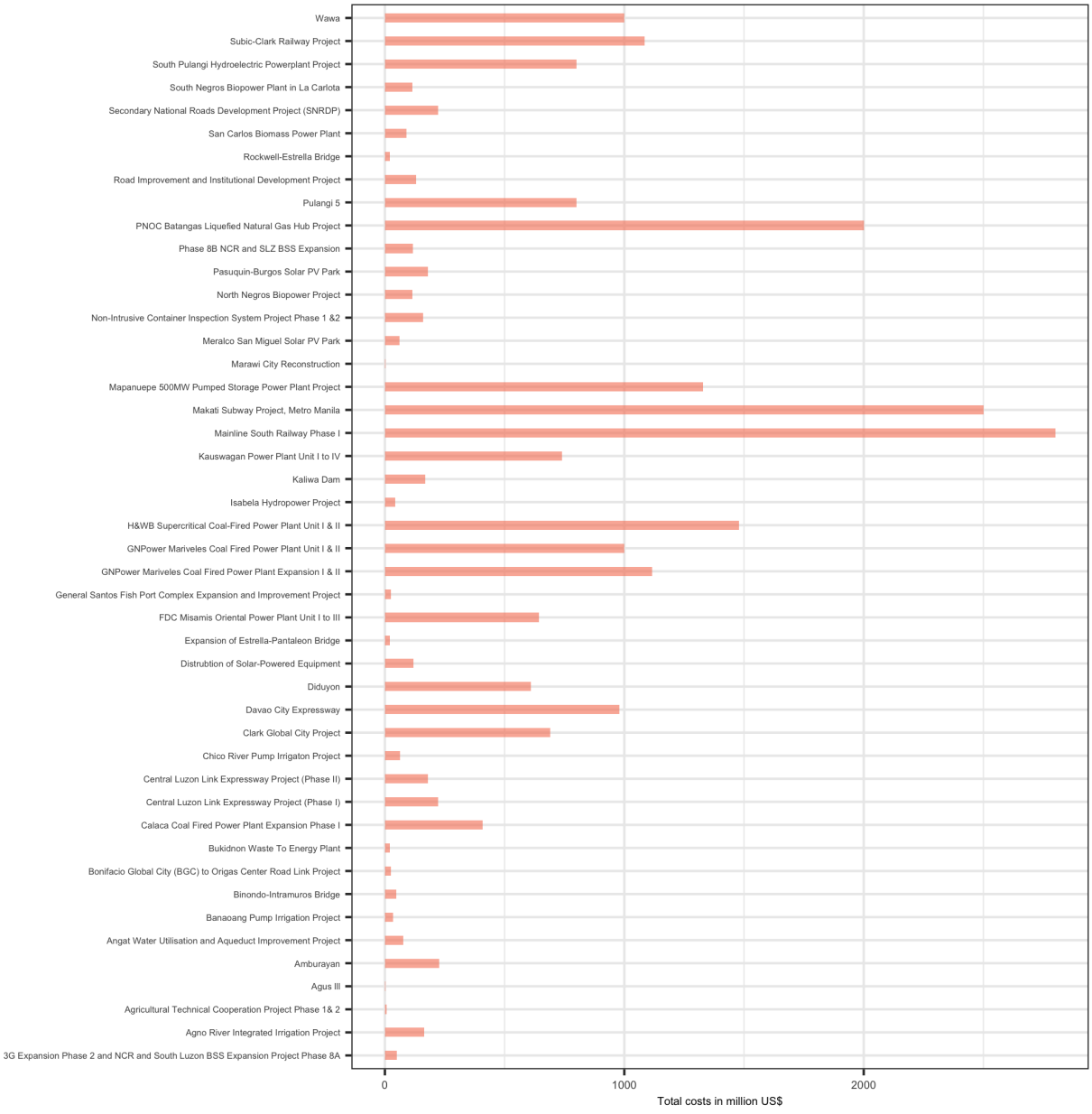


Figure A-2: Total Value of BRI Projects (continued)

(b) Malaysia

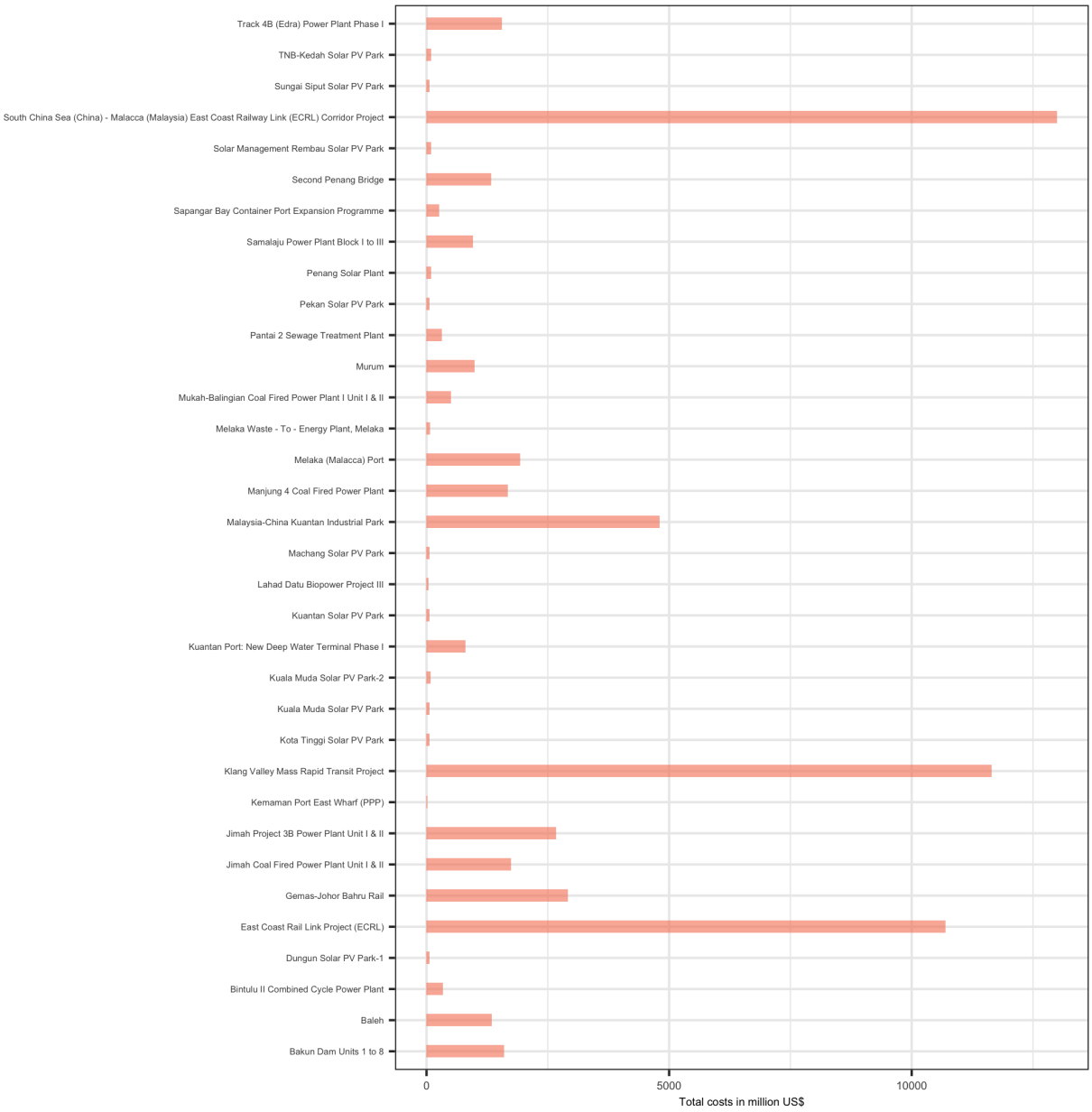


Figure A-2: Total Value of BRI Projects (continued)

(c) Indonesia

1.2.4 Time range of Chinese projects

Figure A-3: Time Distribution of BRI Projects

(a) The Philippines

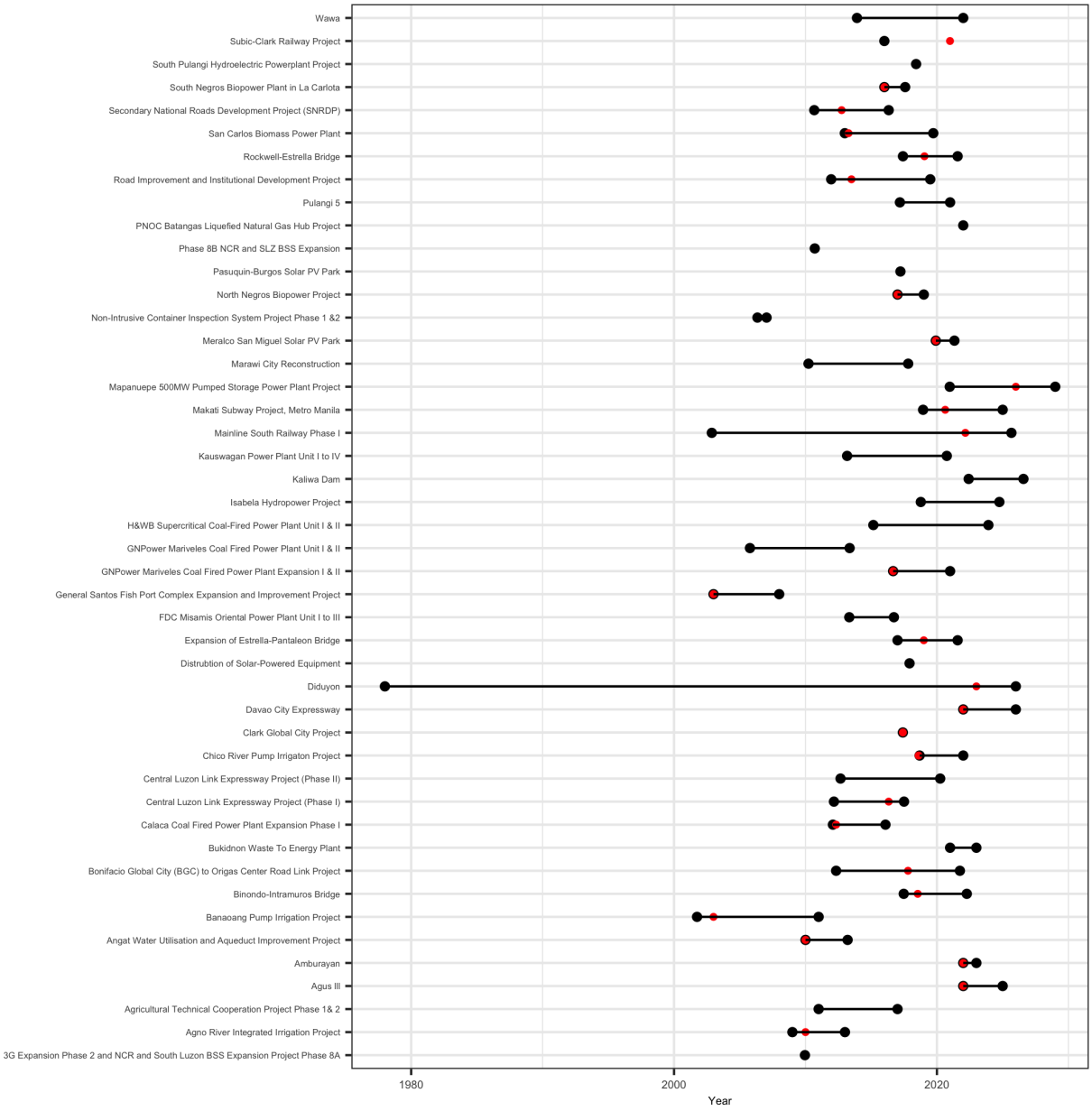


Figure A-3: Time Distribution of BRI Projects (continued)

(b) Malaysia

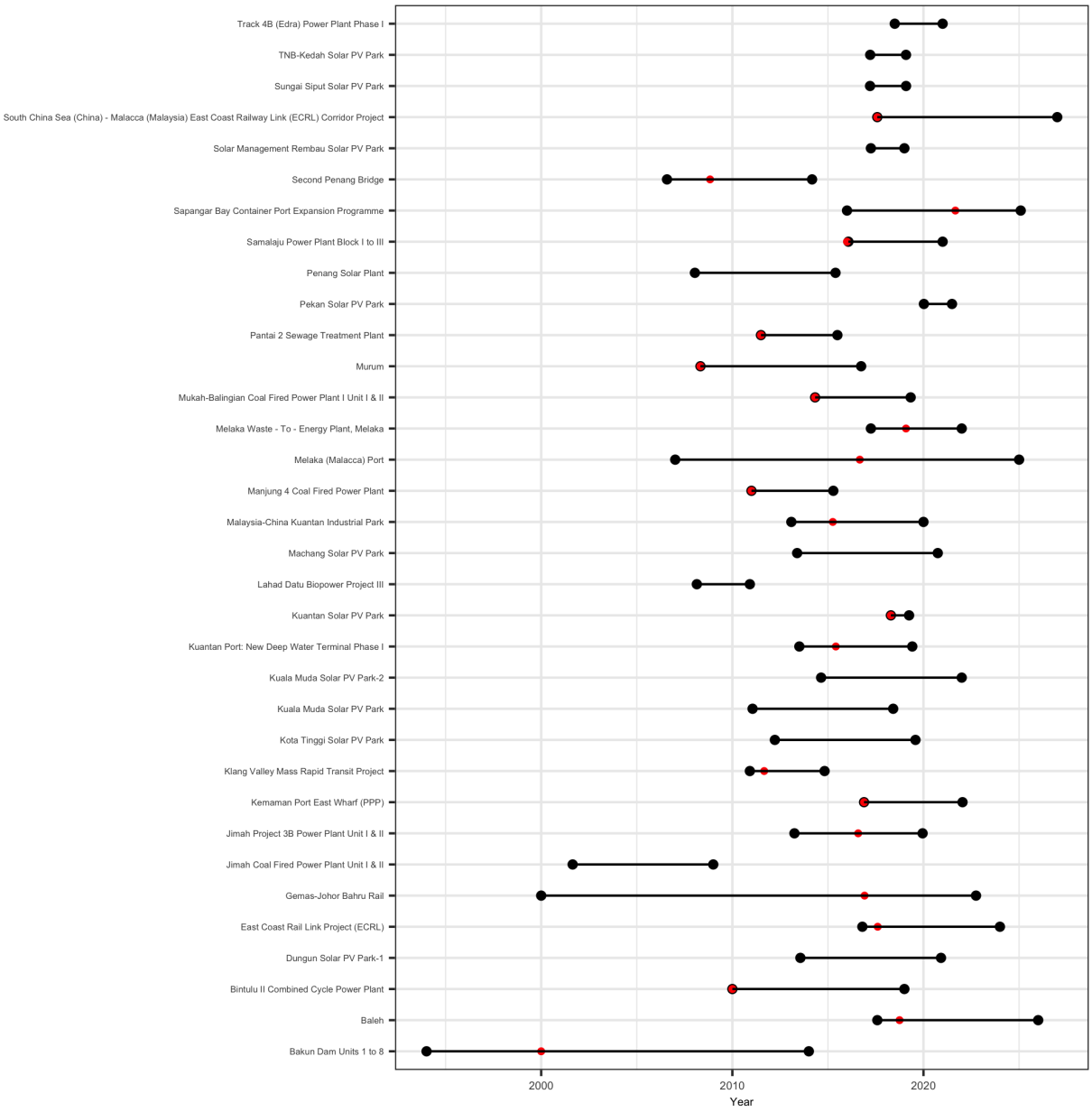
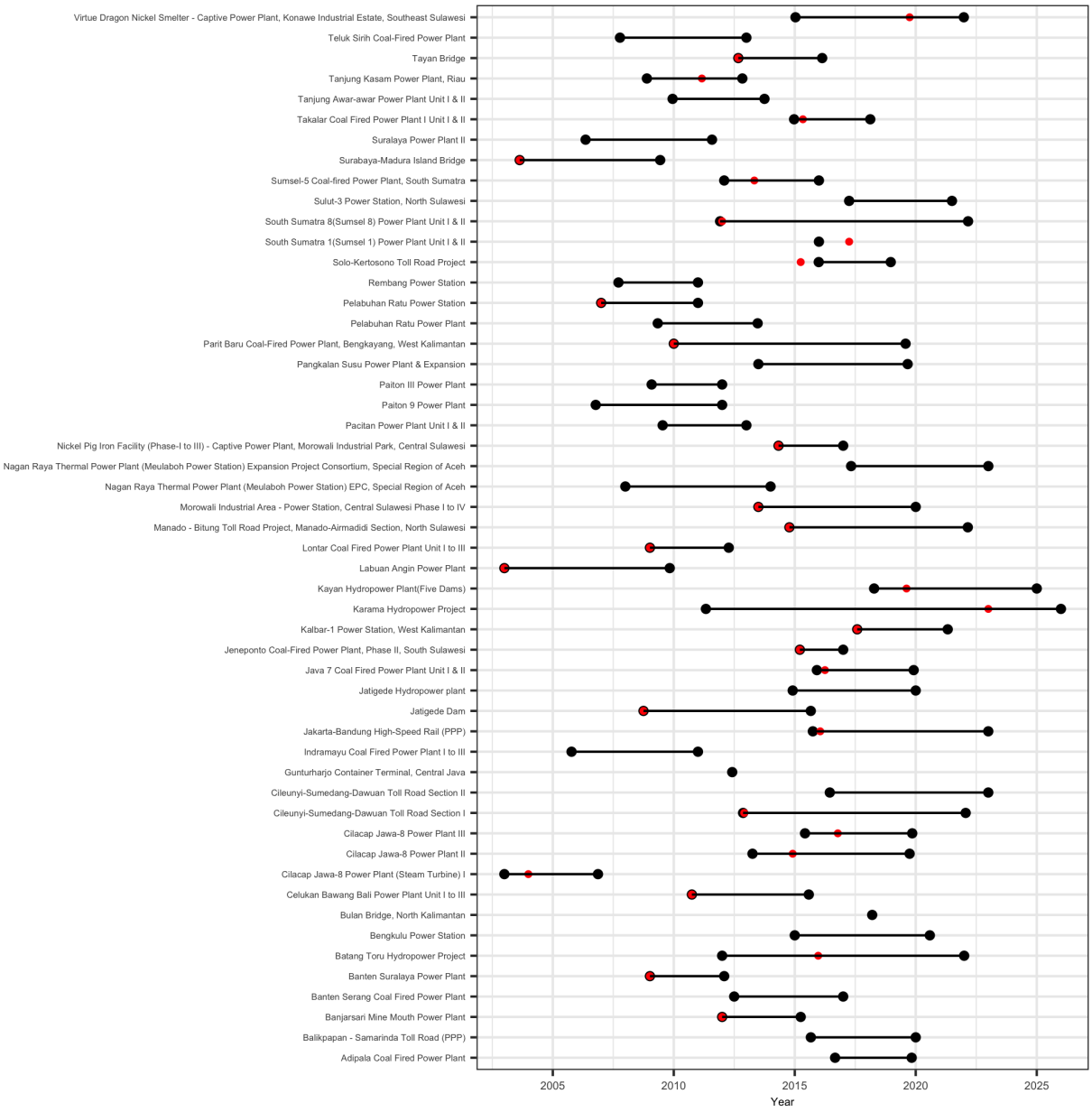


Figure A-3: Time Distribution of BRI Projects (continued)

(c) Indonesia

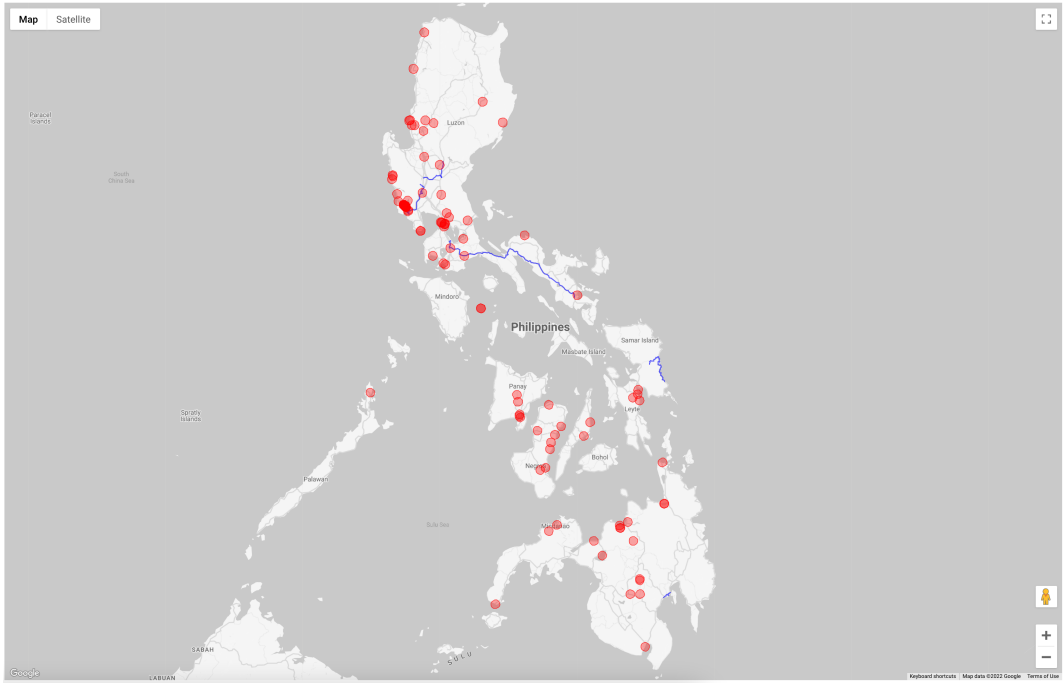


1.2.5 Geolocations of Chinese projects

To correctly map the influenced areas of BRI projects such as roads, railways, and bridges, we estimated the project paths based on publicly available construction plans using Google Direction API [1]. After visualizing the planned construction routes, the resulting Polylines are overlayed with constituency maps to determine units under the influence. Figure A-4 shows all the BRI project's geolocations in each country. Red dots represent the location of projects limited to one city (e.g., powerplants, industry park, seaport), and the blue lines approximate to the paths of railways or roads.

Figure A-4: Geolocations of Chinese Projects

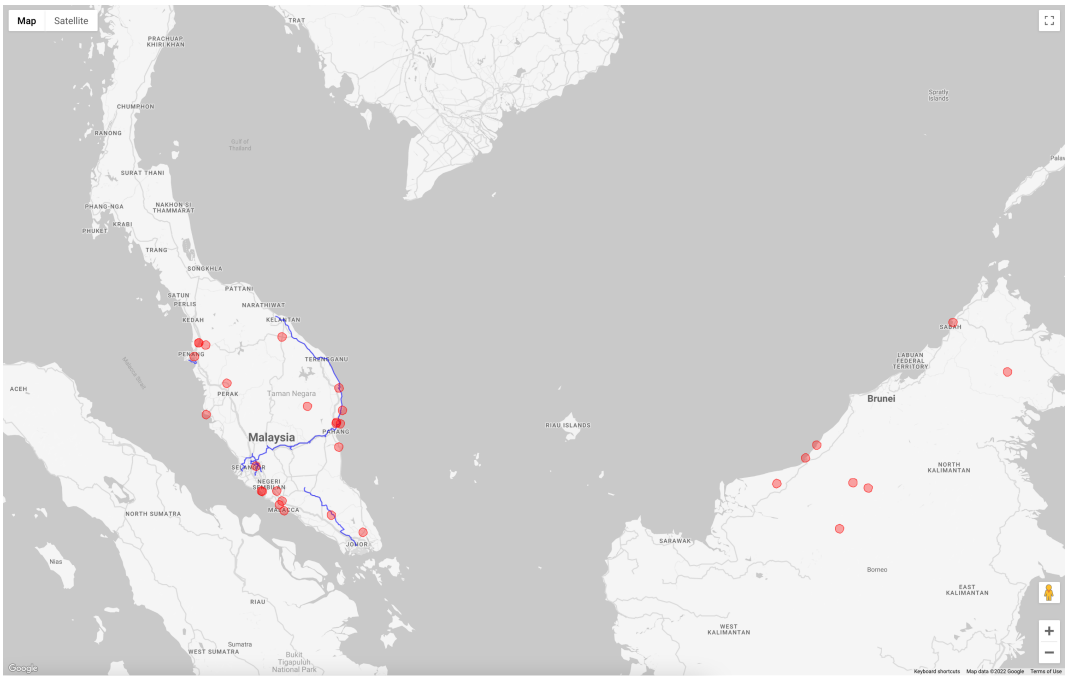
(a) The Philippines



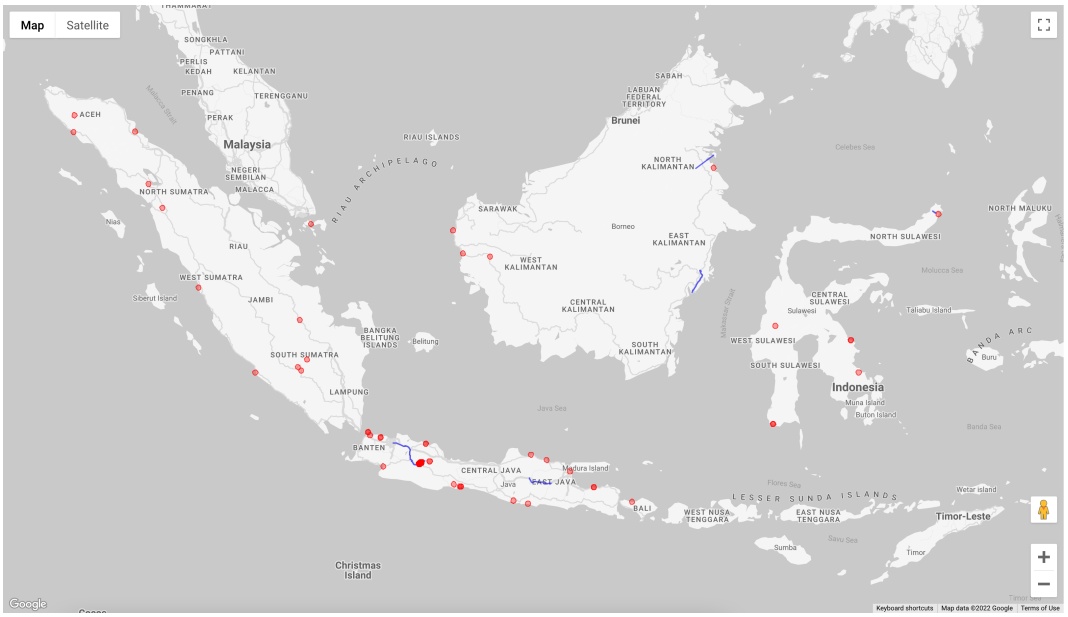
[1] Google Direction API enables us to obtain the approximate shortest route from one city or station to the other, and thus covering all connected regions enroute that can potentially benefit from these projects. Details can be found at `https://developers.google.com/maps/documentation/directions/overview`

Figure A-4: Geolocations of BRI Projects (continued)

(b) Malaysia



(c) Indonesia

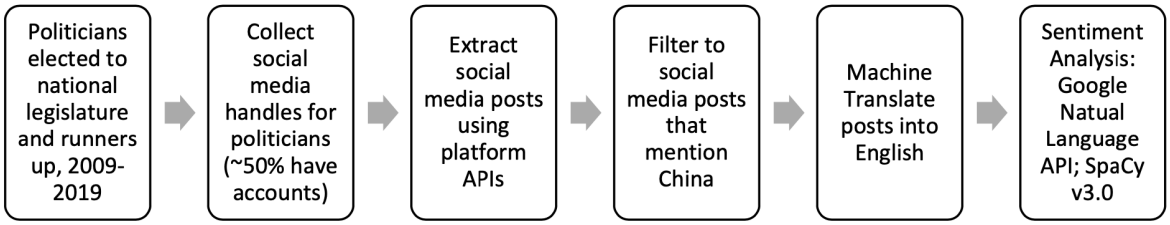


1.3 Outcome: Elite Sentiment

1.3.1 Text processing pipeline

In principle social media expressions could provide a near real-time measurement of elite attitudes toward a range of issues, including China. Figure A-5 shows our pipeline to derive the outcome variable.

Figure A-5: Text Processing Pipeline



We define political elites broadly, focusing on political candidates for the lower chamber of the parliament (or its equivalent, the House of Representatives) in each country from 2009 to 2019. The data is obtained from the Constituency-Level Elections Archive (CLEA) (Kollman et al., 2019). The reason to focus on lower chamber elections is two-folded: First, the lower house, elected directly and more frequently, has more seats than the upper house and the seats are usually proportional to population, thus providing better representation of voter preferences. Second, the lower chamber is elected across all democracies consistently, which facilitates "the comparison across countries and within countries over time"[2].

We keep elected candidates together with their runner-up as our base of political elites. Note that we purposefully add runner-up to our base sample, so that we can take into account not only the legislative members in power but also their challengers (possibly with different policy preferences). Below is the list of political elites in the original sample for each country.

- Philippines: elected political candidates and their runner-up of the 2010, 2013, 2016, and 2019 elections for the House of Representatives.

[2] "Why do you focus on lower chamber elections?", Constituency-Level Elections Archive (CLEA), `https://electiondataarchive.org/about-clea/faq/`

- Malaysia: elected political candidates and their runner-up of 2008, 2013, 2018 general elections for the Dewan Rakyat, the lower chamber of parliament.
- Indonesia: elected political candidates and their runner-up of 2009, 2014, and 2019 legislative elections for the People's Representative Council (DPR), the lower chamber of parliament.

With this representative list of politicians, our locally based research assistants manually searched for their social media handles. Table A-3 reports the total number of politicians in each country and those with identified social media handles. Politicians entered into our analysis are those with either Facebook or Twitter accounts in each country, which is more than half of the total sample of politicians.

Table A-3: Politicians with Identified Social Media Handles

Country	Total number	on Facebook	on Twitter	Either	Both
The Philippines	2303	889 (38.60%)	412 (17.89%)	1018 (44.20%)	283 (12.29%)
Malaysia	1362	552 (40.53%)	499 (36.64%)	720 (52.86%)	331 (24.30%)
Indonesia	2243	716 (31.92%)	990 (44.14%)	1291 (57.56%)	415 (18.50%)

Note: Exact number of politicians in total, with identified social media handles (at least one) on Facebook, on Twitter, on either Facebook or Twitter, and on both Facebook and Twitter. Numbers in parentheses refer to the percentage of politicians with the identified social media handles out of the total number of politicians in our sample.

Table A-4: List of Filtering Keywords

Country	Keywords about China	Keywords on events
The Philippines	China, Chinese, Cina, Tiongkok, Tsina, Intsik	BRI, Scarborough Shoal, Huangyan, Panatag Shoal, West Philippine Sea, South China Sea.
Malaysia		BRI, Luconia, Betting Patinggi Ali, Borneo.
Indonesia		BRI, Natuna, Pulau Natuna, Laut Natuna, Kepulauan Natuna.

Note: All keywords are case insensitive. For each country, we determine the post is about China if there are at least one or more keywords matched in the first set "keywords about China" *and* at least one or more keywords matched in the second set "keywords on events".

The full archive of Facebook posts and tweets from the selected politicians are extracted using Facebook CrowdTangle API [3] and Twitter's V2 Research API [4].

[3] CrowdTangle Team (2021). CrowdTangle. Facebook, Menlo Park, California, United States. [1612669, 1612670, 1612671] `https://help.crowdtangle.com/en/`

[4] Twitter's V2 Research API, `https://developer.twitter.com/en/products/twitter-api/academic-research`

Table A-5: Summary Statistics of China-related Social Media Posts

Country	on Facebook	Facebook		on Twitter	Twitter	
		Politicians	Facebook posts		Politicians	Tweets
The Philippines	889	384 (43.19%)	32,465 (6.34%)	412	189 (45.87%)	19,681 (2.86%)
Malaysia	552	324 (58.70%)	43,587 (5.31%)	499	384 (76.95%)	57,611 (2.19%)
Indonesia	716	281 (39.25%)	4,578 (1.87%)	990	461 (46.57%)	26,442 (0.89%)

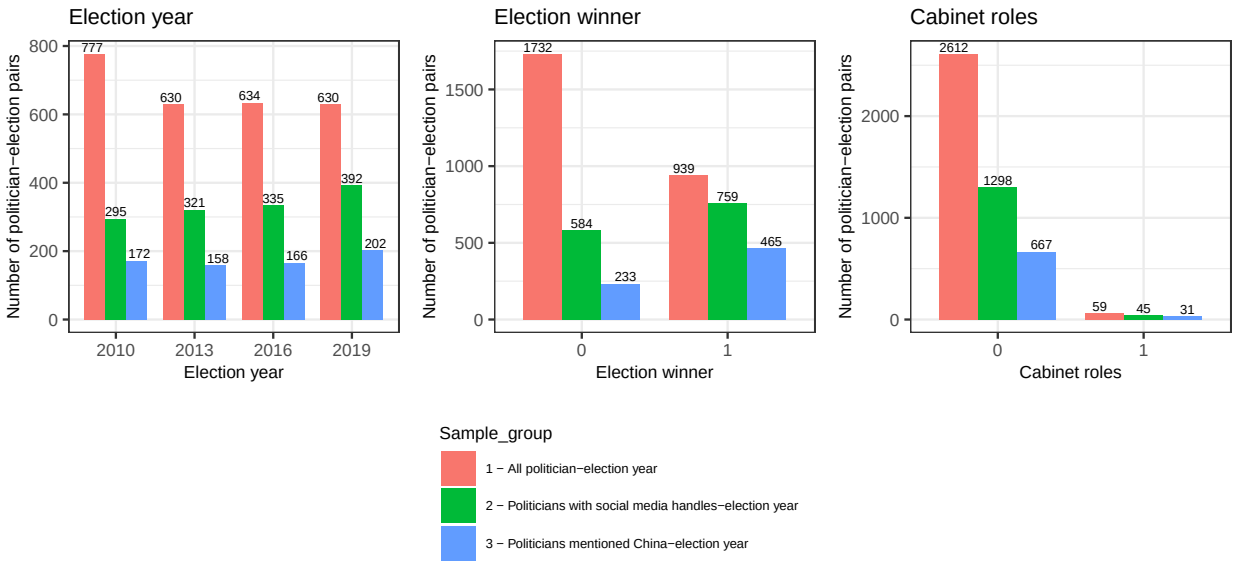
Note: Numbers indicate the number of politicians with identified social media handles that ever posted about China and the number of social media posts about China according to our filtering mechanism. Numbers in parentheses refer to the percentage of politicians that ever posted about China out of the total number of politicians with identified social media handles and the percentage of these China-related posts out of the total number of social media posts that we collected.

To measure politicians' sentiment targeted at China or China-related issues, we employed a series of keywords with a combination of "China/Chinese" and prominent events in four different languages to filter the social media posts (see A-4 for details). Table A-5 shows the summary statistics of the selected political elites and social media posts that enter the final analysis. In general, more than half of the politicians with identified social media handles (from 39.25% to 76.95%) posted about China at least once from 2009 to 2021. However, only about 0.89% to 6.34% of the total social media conversations from the identified social media handles directly mentioned China or prominent South China Sea disputes.

Figure A-6 compares the characteristics of politicians that have social media handles and mentioned China, to those of all politicians in the original sample. The descriptive analysis demonstrates that the number of politicians engaging in China-related conversations is balanced across the election years, but they are more likely to be elected or hold cabinet roles relative to those without social media handles or those who never talked about China.

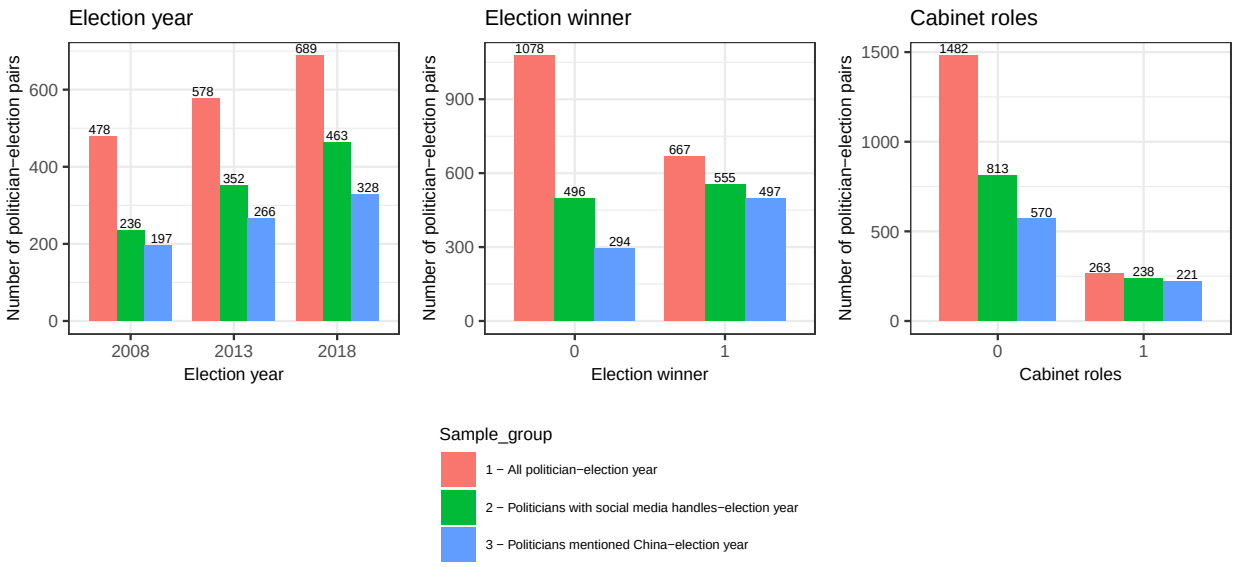
Figure A-6: Descriptive Analysis of Politicians that Mentioned China on Social Media

(a) The Philippines



Note: Observations are at the politician−election year level. Each politician can associate with multiple election years.

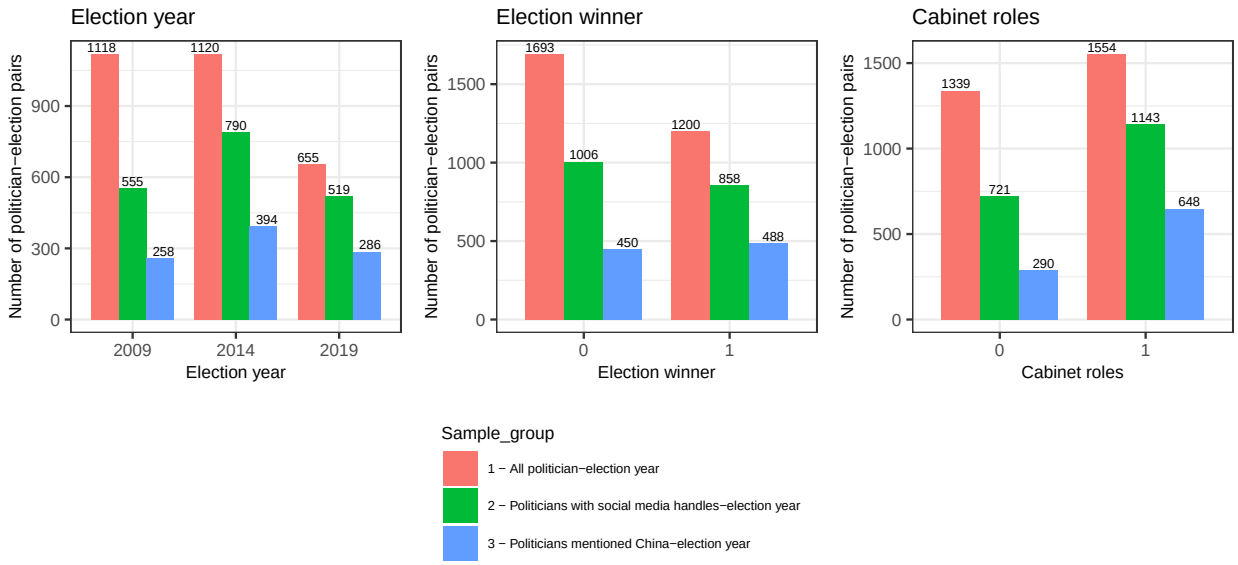
(b) Malaysia



Note: Observations are at the politician−election year level. Each politician can associate with multiple election years.

Figure A-6: Descriptive Analysis of Politicians that Mentioned China on Social Media (continued)

(c) Indonesia



Note: Observations are at the politician−election year level. Each politician can associate with multiple election years.

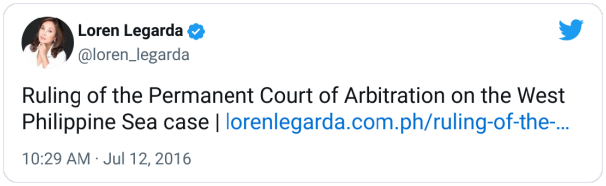
1.3.2 Sample social media posts and machine translation

Figure A-7 displays some examples of tweets in the Philippines that mentioned China. There are tweets with positive, negative, and neutral sentiments across constituencies. Politicians' language usage on social media does not differ that much across BRI treatment statuses.

Figure A-7: Sample tweets in the Philippines

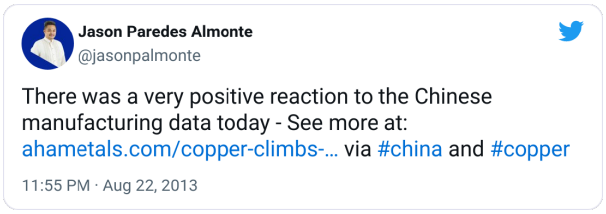
(a) Constituency without BRI: sentiment = 0.2

(b) Constituency without BRI: sentiment = 0



(c) Constituency without BRI: sentiment = -0.6

(d) Constituency with BRI: sentiment = 0.7



(e) Constituency with BRI: sentiment = 0

(f) Constituency with BRI: sentiment = -0.6

Indonesian, Malay, and Tagalog are the most used language other than English in the three countries of research. Direct text processing with posts in these languages is difficult as these languages are less resourced. Therefore, we first translate all the posts into English using Google Translation API and then conduct sentiment analysis.

Based on its Neural Machine Translation (NMT) model, Google Translation API is one of the major machine translation services (Wu et al., 2016). According to a recent evaluation of Google

Translation accuracy (Aiken et al., 2019), the overall performance of the NMT model (4.263 out of 6) is close to human translation quality (4.636 out of 6). The BLEU score (100 in total)[5] of the NMT model with Indonesian text translated to English reaches 82, Malaysian to English 76, and Tagalog to English 70. These scores are comparable to some most well-resourced and accurately translated language pairs (e.g., from German to English 81, from Chinese to English 78). Figure A-8 shows some examples of original tweets in Indonesian, Malay, and Tagalog and machine translation results returned by the Google Translation API.

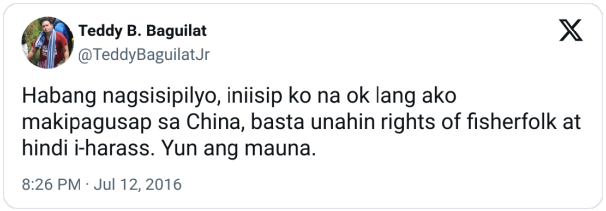
[5]BLEU (Bilingual Evaluation Understudy) is a metric for evaluating machine-translated text.

Figure A-8: Machine translation examples

(a) Indonesian: sentiment = 0

Translation provided by Google Translation API: *"I'm a member of the same DPR with the BRI tagline, serving all layers of society ich ro"*

(b) Tagalog: sentiment = 0.2



Translation provided by Google Translation API: *"while brushing I think it's ok for me to talk to china as long as I prioritize the rights of fisherfolk and don't harass that's the first thing"*

(c) Indonesian: sentiment = -0.5

Translation provided by Google Translation API: *"An underwater drone allegedly belonging to China was found by fishermen near the Selayar Islands, South Sulawesi, the government needs to investigate immediately to uncover the drone's origins"*

(d) Tagalog: sentiment = -0.8

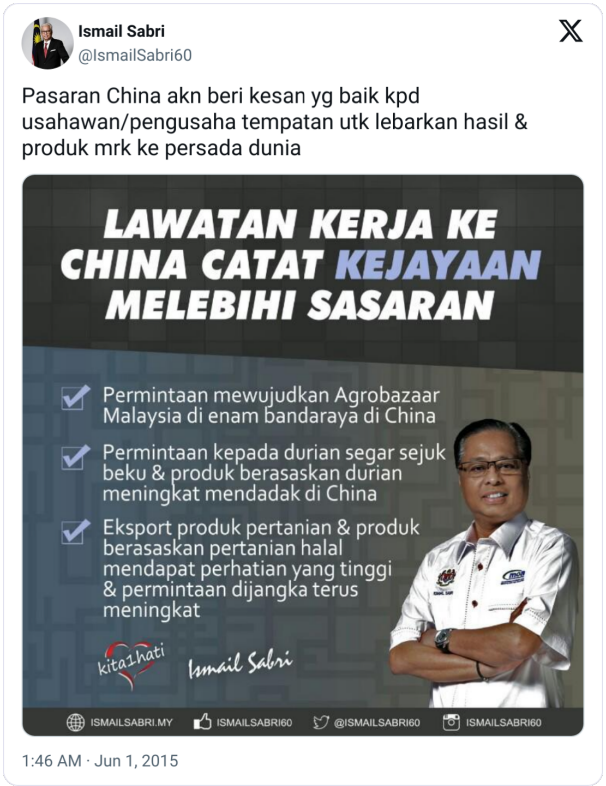
Translation provided by Google Translation API: *"there is a big problem in the west philippine sea the next president needs to know what to do especially with the conflict between america and china kqezi"*

Figure A-8: Machine translation examples (continued)

(e) Malay: sentiment = -0.7

Translation provided by Google Translation API: *"invasion of chinese planes into malaysian airspace diplomatic protest note to chinese government call chinese ambassador to malaysia to explain friendly diplomatic relations with any country never means we compromise on security"*

(f) Malay: sentiment = 0.8



Translation provided by Google Translation API: *"The Chinese market will give a good impression to local entrepreneurs to spread the results of their products to the world stage"*

1.3.3 Sentiment analysis algorithms: Google NLP API and SpaCy

We used two approaches to obtain elite sentiment of social media posts about China. First, we utilized Google Natural Language API, a Machine Learning based language model, to calculate the overall sentiment score for each post [6]. The Google Natural Language API utilizes pre-trained language models and returns a sentiment score for each post. The sentiment score that ranges from -1 to 1, with -1 being negative and 1 being positive. Posts assigned neutral scores (around 0) may contain "low-emotion" or "mixed emotions".

The second approach follows a simple lexicon-based method using the SpaCy v3.0 package [7]. Based on a pre-defined dictionary classifying positive and negative words, the general semantic orientation for each post is calculated by the proportion of positive words as opposed to the proportion of negative words. The final polarity score for each post is a normalized continuous variable, and ranges from -1 to 1, with -1 being completely negative, 1 being completely positive, and 0 being neutral.

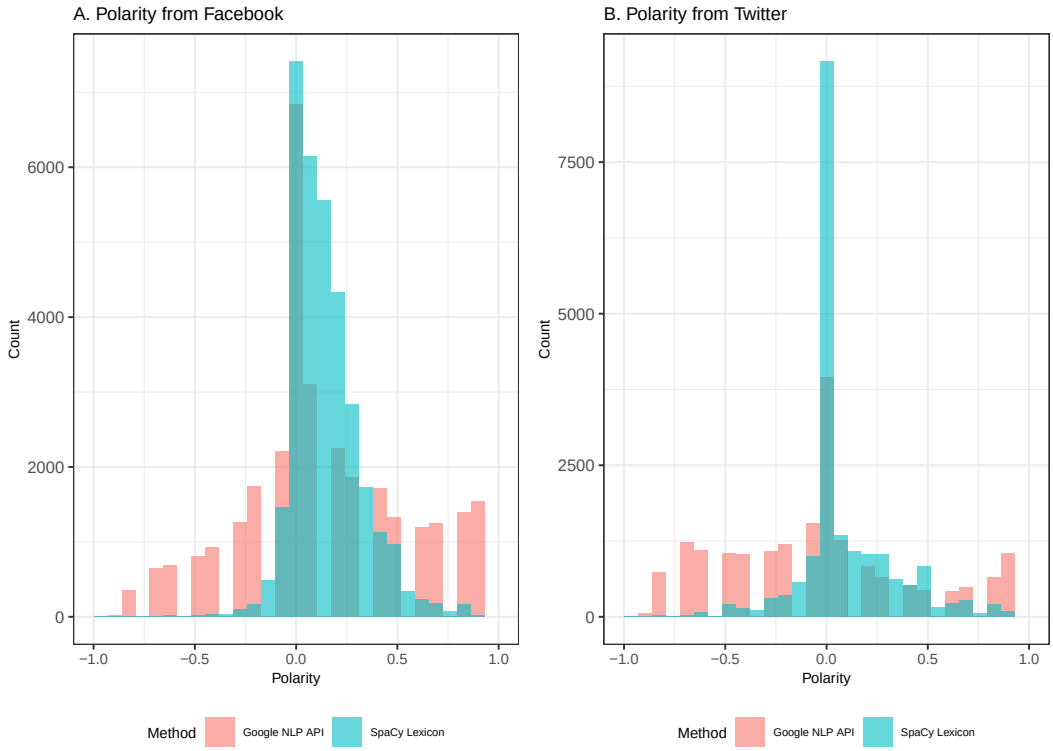
1.3.4 Histogram of politician sentiment

Figure A-9 compares the distribution of politician sentiment derived by Google Natual Language API and SpaCy v3.0 package. The distribution of both sentiment measures conform to a normal distribution, with most of the posts clustering at the neutral stance. The overall sentiment from the lexicon-based method (SpaCy v3.0) is slightly right-skewed for Facebook posts and tweets, with a larger amount of posts assigned neutral scores. The sentiment distribution from the Google Natural Language API spans wider, with more posts identified with clear negative and positive tones.

[6] Google Natural Language API: `https://cloud.google.com/natural-language/docs/sentiment-tutorial`
[7] Sentiment Analysis using spaCy-textblob package: `https://spacy.io/universe/project/spacy-textblob`

Figure A-9: Histogram of Politician Sentiment

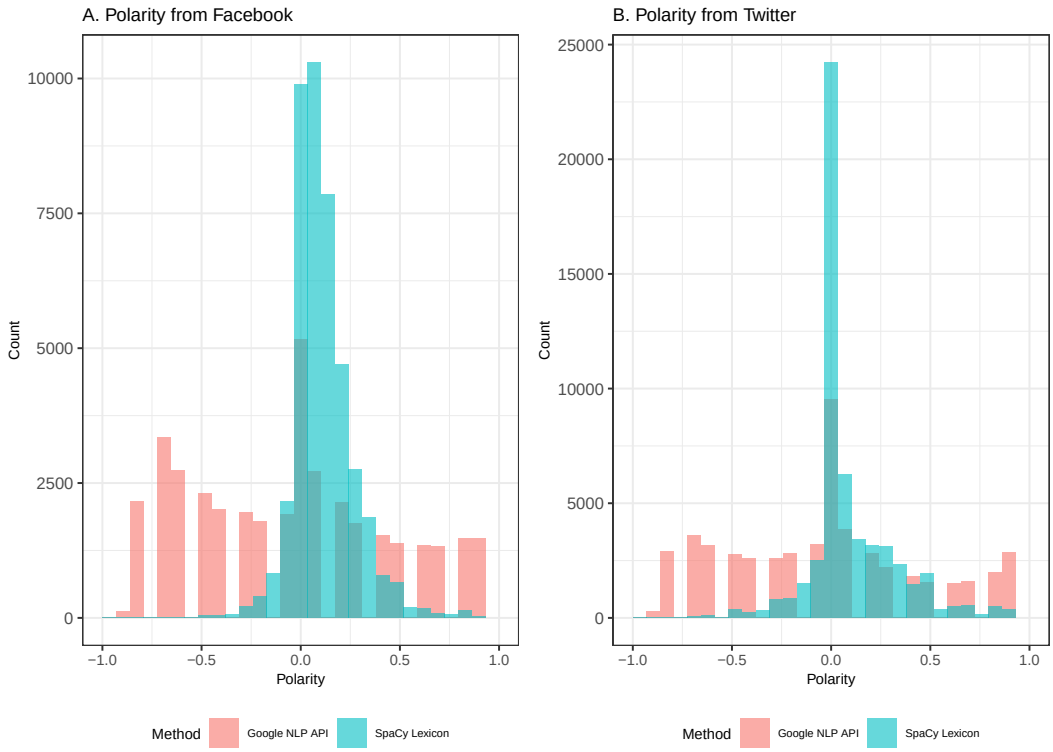
(a) The Philippines



Notes: Total number of Facebook posts = 33459.

Notes: Total number of tweets = 19542.

(b) Malaysia

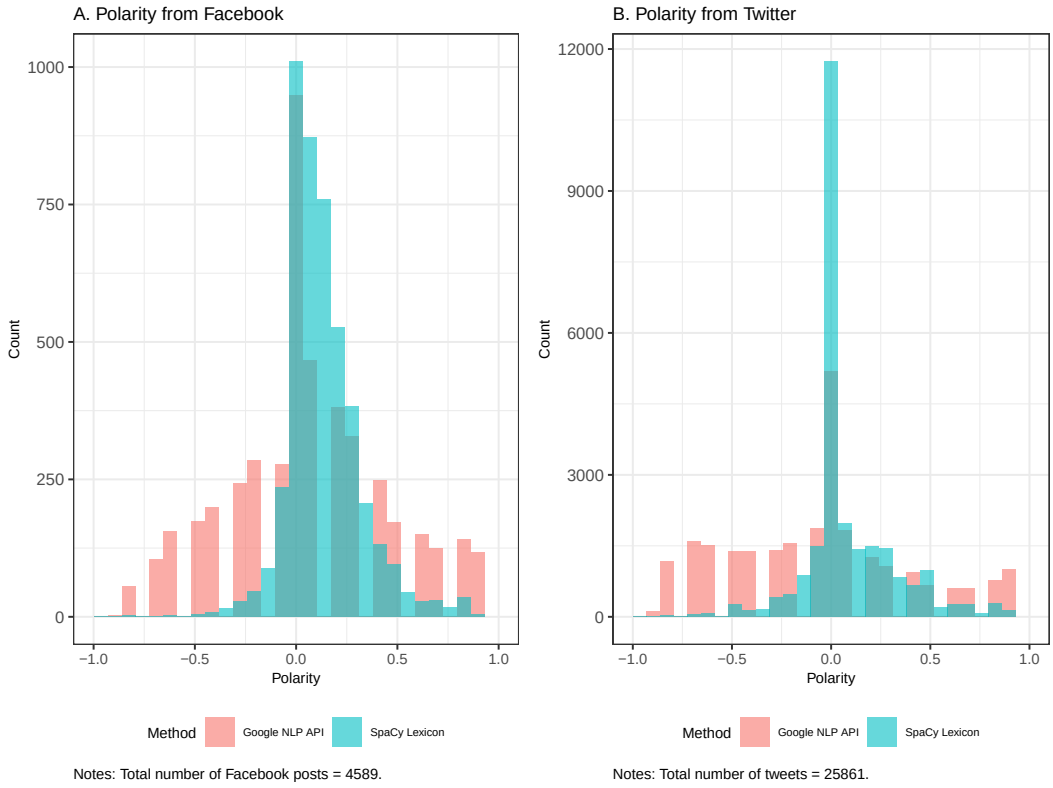


Notes: Total number of Facebook posts = 43321.

Notes: Total number of tweets = 55424.

Figure A-9: Histogram of Politician Sentiment (continued)

(c) Indonesia



1.3.5 Time trend of politician sentiment toward China

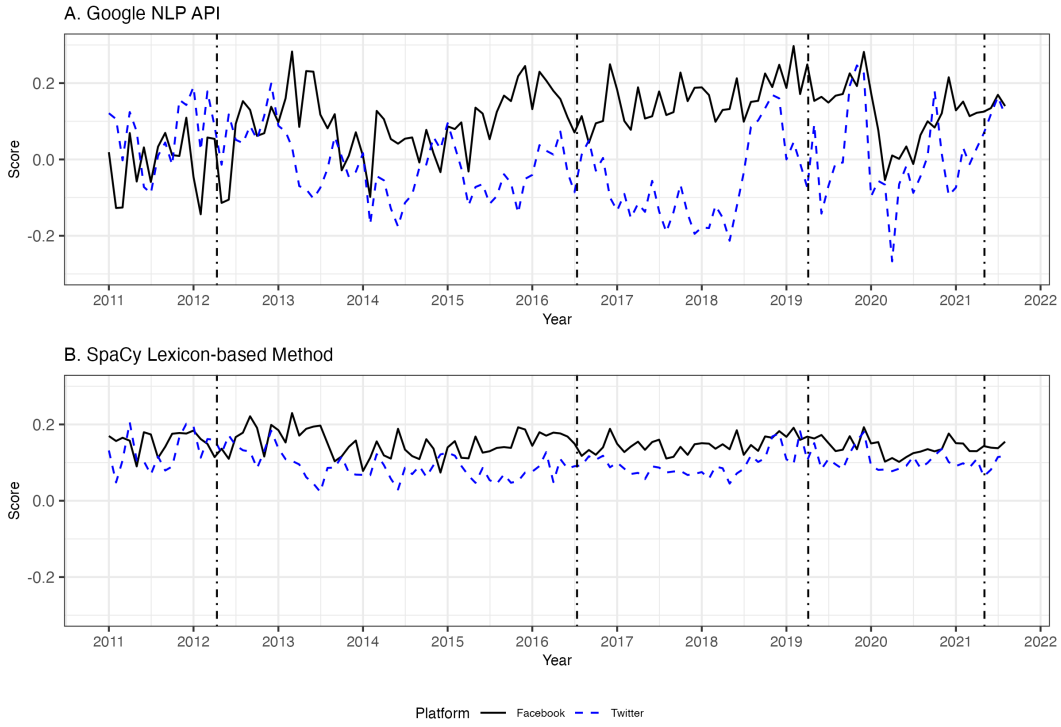
Figure A-10 displays the general time trend of the elite attitude toward China from 2011 to 2021 in each country. Sentiment scores are aggregated at the monthly level and averaged out for a smoother time trend. For each country, the panel A shows the average sentiment scores calculated using Google Natural Language API, and the panel B shows the average sentiment scores derived by the lexicon-based method using SpaCy v3.0 package. Both panels are put under the same scale for each country to facilitate the comparison. The average sentiment scores from SpaCy are slightly positive and fairly smooth across years with smaller variations than that from Google NLP API. The average sentiment scores from Google NLP API generally indicate larger variation centering around the neutral stance, with favorable opinions toward China slightly increasing across the years in Malaysia, Indonesia, and on the Facebook platform in the Philippines.

Politicians might prefer one social media platform to another and strategically tailor posts to follower interests on a specific platform. In the Philippines, the Facebook platform observed more

positive opinions toward China while the tweets are generally less favorable toward China. On the contrary, politicians in Malaysia expressed slightly more positive opinions toward China on Twitter than Facebook, though the differences narrowed after 2016. We do not see substantial sentiment difference across social media platforms in Indonesia.

Figure A-10: Time Trend of Politician Sentiment toward China

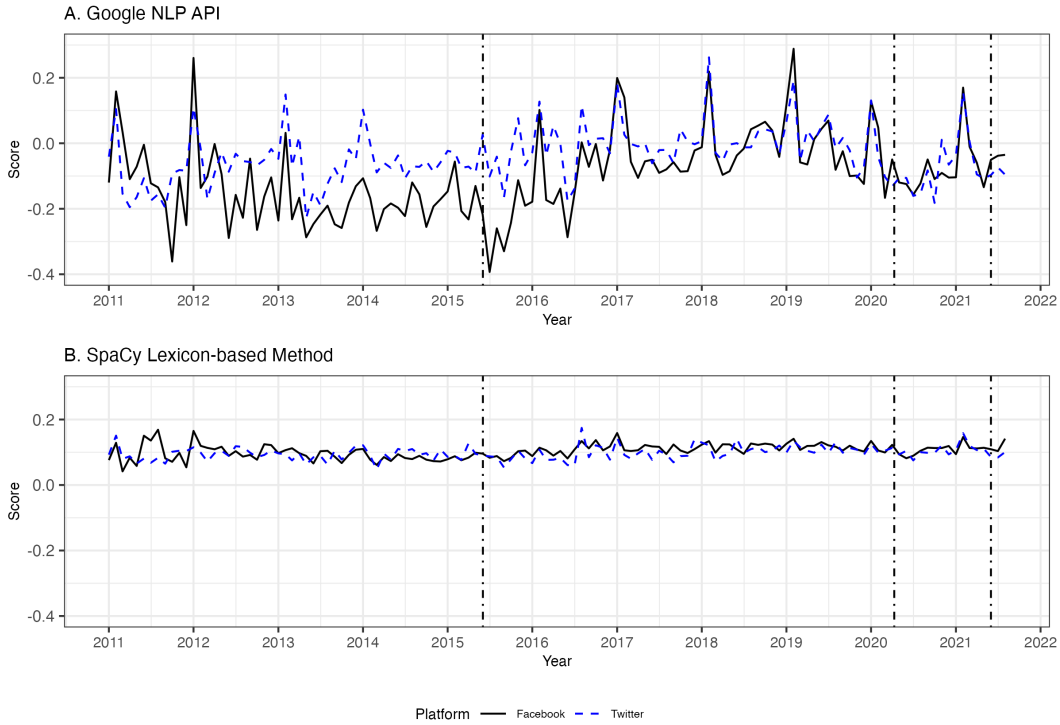
(a) The Philippines



Note: Score averaged at the monthly level. Dashed vertical lines represent the exogenous shocks of South China Sea disputes on 2012-04-11, 2016-07-12, 2019-04-05, and 2021-05-04.

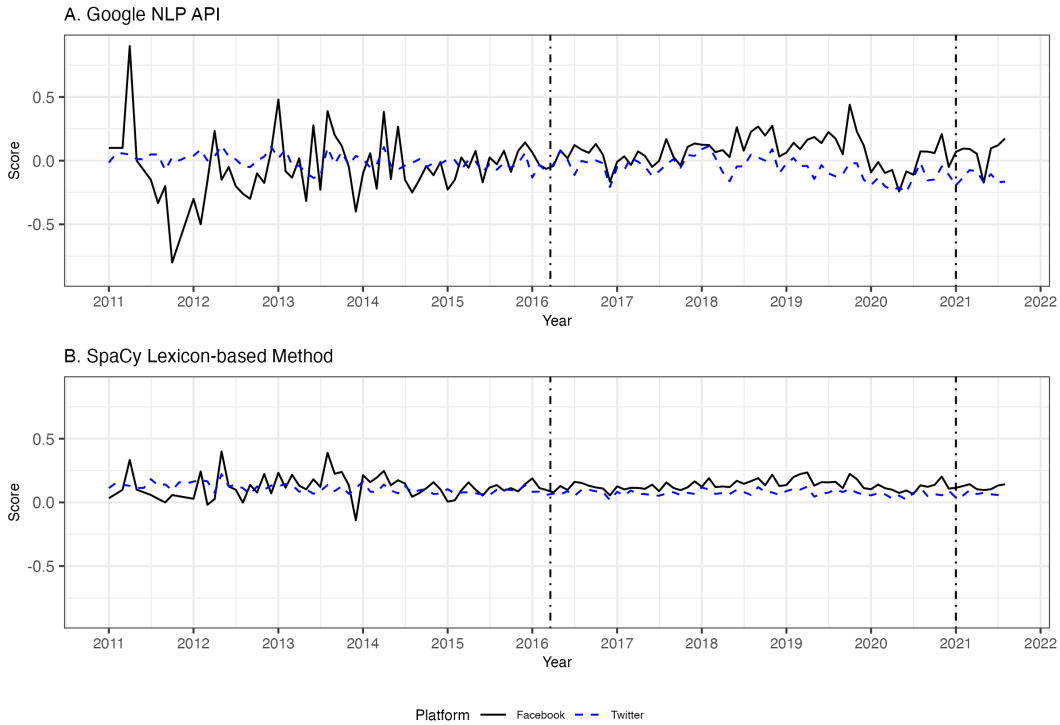
Figure A-10: Time Trend of Politician Sentiment toward China (continued)

(b) Malaysia



Note: Score averaged at the monthly level. Dashed vertical lines represent the exogenous shocks of South China Sea disputes on 2015-06-02, 2021-06-01, and 2020-04-10.

(c) Indonesia

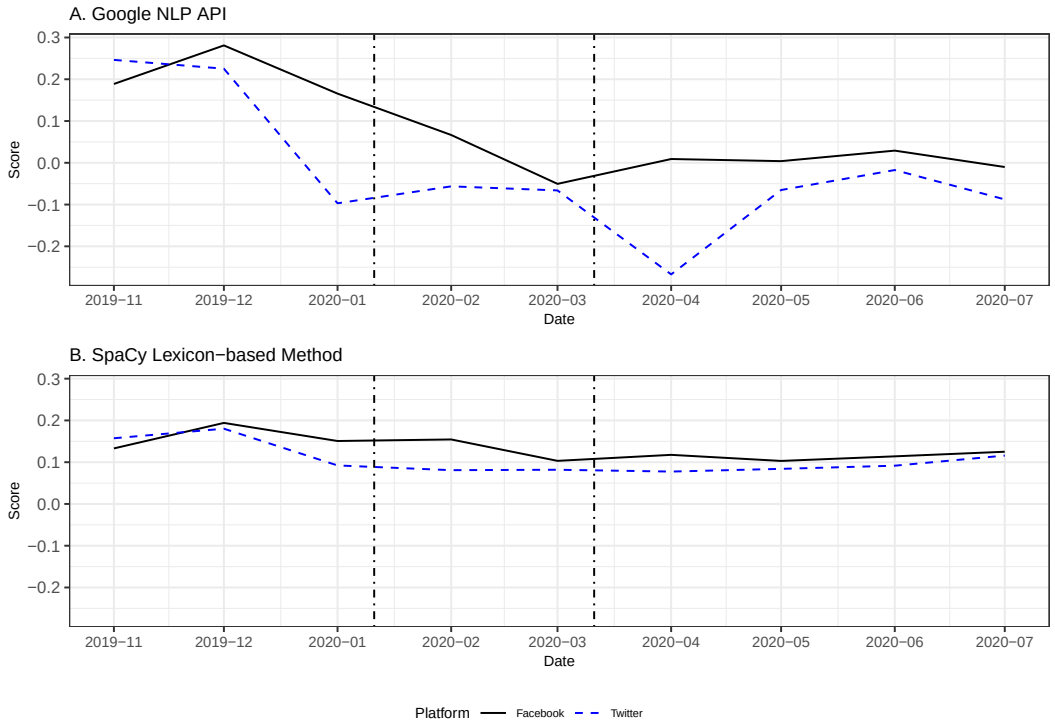


Note: Score averaged at the monthly level. Dashed vertical lines represent the exogenous shocks of South China Sea disputes on 2016-03-19 and 2021-01-01.

To demonstrate the validity of the sentiment analysis, we plot the politicians' valence toward China around other influential world events about China. We chose the event shocks of COVID in 2020 for its clear-cut dates, with the first shock on 2020-01-11, when China reported the first death from COVID, and the second shock on 2020-03-11 when the WHO declared COVID-19 a pandemic. Figure A-11 shows that sentiments toward China well correspond to real-world events, starting to plunge in January 2020 on both platforms in all countries, with another dip around March, especially in the graphs using Google sentiment scores. Figure A-12 breaks down the sentiment by BRI receiving status, and there is not much difference between BRI recipients and non-recipients in the reactions. These graphs provide further evidence of the validity of the sentiment scores (especially the Google sentiment scores) and our main finding on the lack of strong evidence of BRI influence on local politician sentiment.

Figure A-11: Politician Sentiment Toward China Before and After COVID Shock

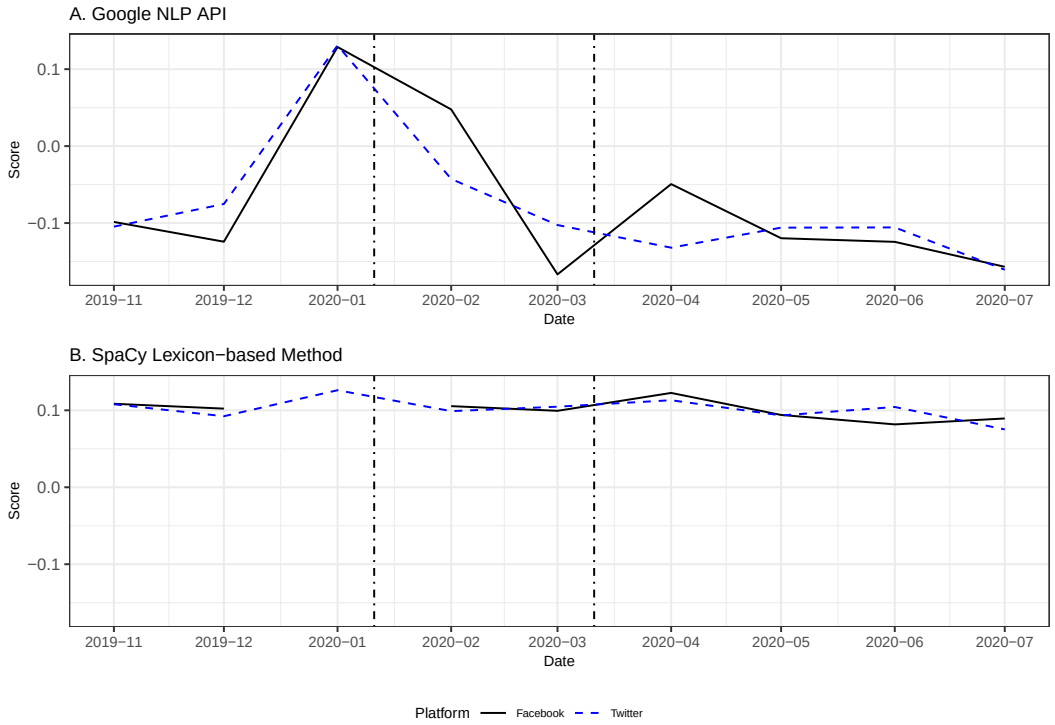
(a) The Philippines



Note: Sentiment scores averaged at the monthly level. Dashed vertical lines represent the date when China reported the first death from COVID (2020−01−11) and the date when the WHO declares COVID−19 a pandemic (2020−03−11)

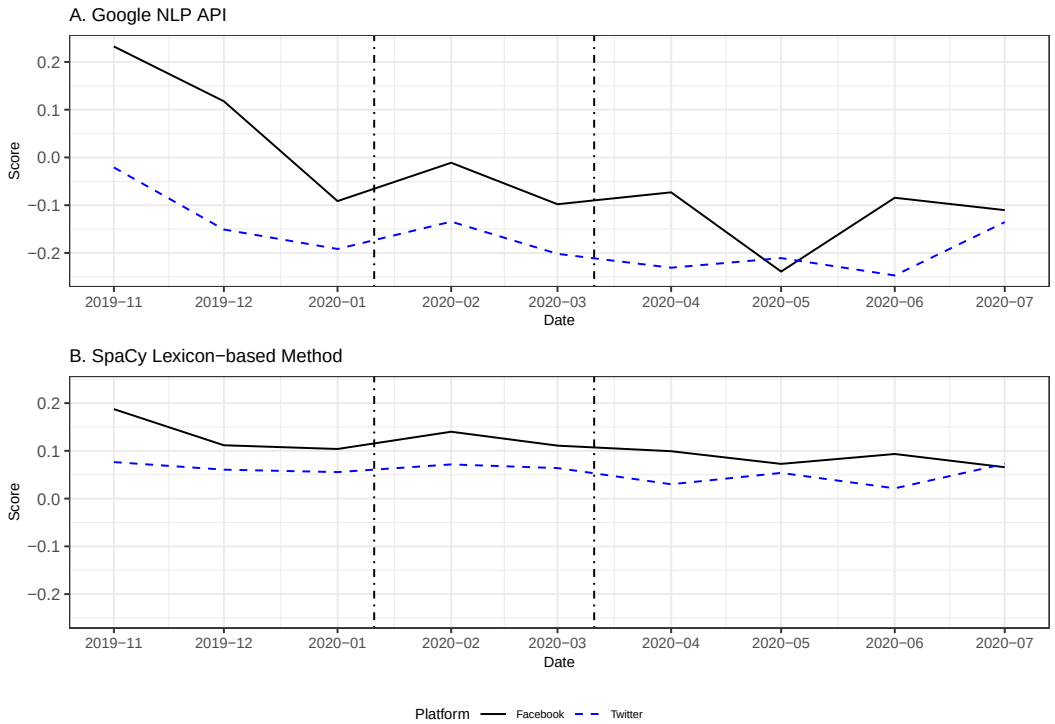
Figure A-11: Politician Sentiment Toward China Before and After COVID Shock (continued)

(b) Malaysia



Note: Sentiment scores averaged at the monthly level. Dashed vertical lines represent the date when China reported the first death from COVID (2020−01−11) and the date when the WHO declares COVID−19 a pandemic (2020−03−11)

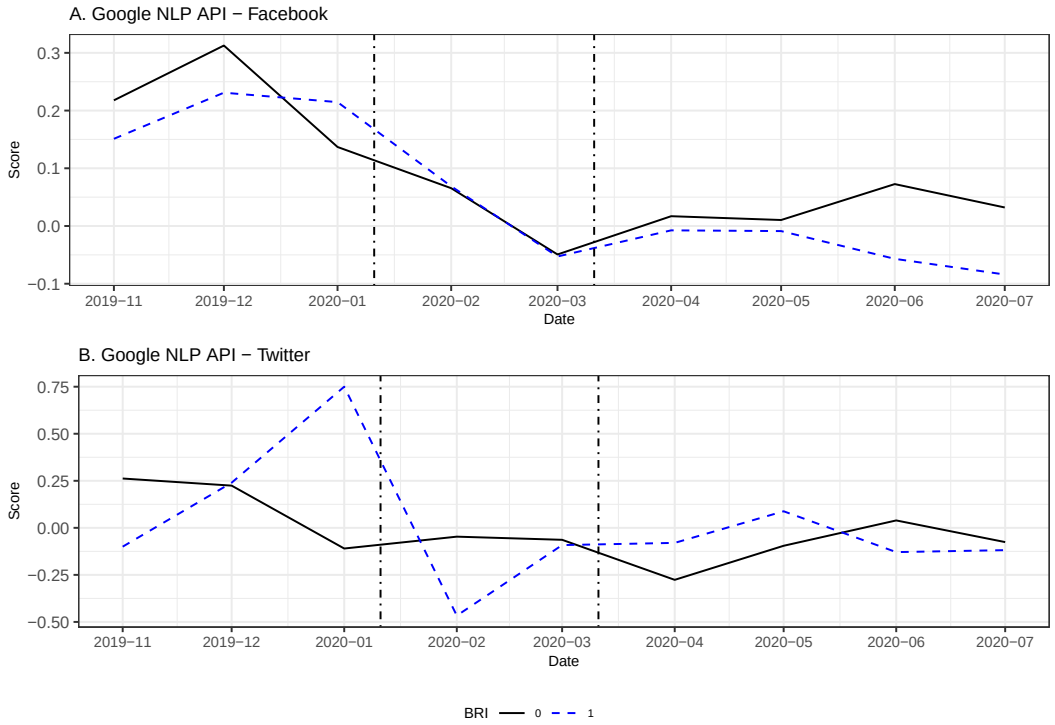
(c) Indonesia



Note: Sentiment scores averaged at the monthly level. Dashed vertical lines represent the date when China reported the first death from COVID (2020−01−11) and the date when the WHO declares COVID−19 a pandemic (2020−03−11)

Figure A-12: Politician Sentiment Toward China Before and After COVID Shock: Condition on BRI Treatment

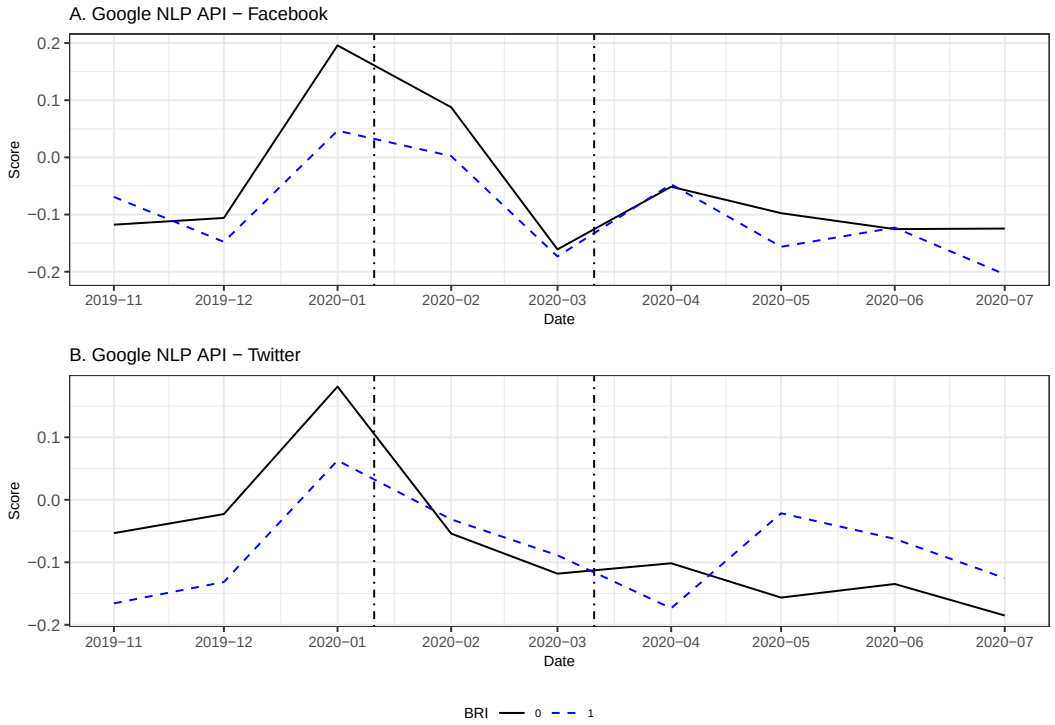
(a) The Philippines



Note: Sentiment scores averaged at the monthly level. Dashed vertical lines represent the date when China reported the first death from COVID (2020−01−11)and the date when the WHO declares COVID−19 a pandemic (2020−03−11)

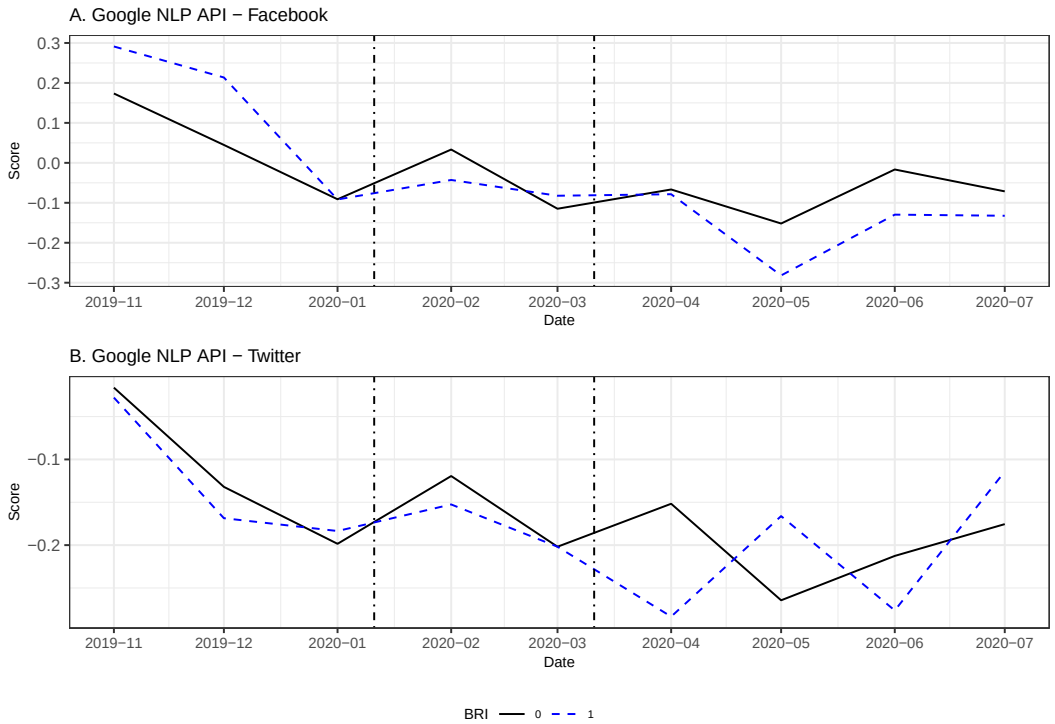
Figure A-12: Politician Sentiment Toward China Before and After COVID Shock (continued)

(b) Malaysia



Note: Sentiment scores averaged at the monthly level. Dashed vertical lines represent the date when China reported the first death from COVID (2020−01−11)and the date when the WHO declares COVID−19 a pandemic (2020−03−11)

(c) Indonesia



Note: Sentiment scores averaged at the monthly level. Dashed vertical lines represent the date when China reported the first death from COVID (2020−01−11)and the date when the WHO declares COVID−19 a pandemic (2020−03−11)

1.3.6 Spatial distribution of identified politicians that engage in China-related conversation

The sentiment for each post is linked to a certain constituency in order to merge with the treatment variables of Chinese projects. First, we associated each post with the political candidates by their social media handles. Second, using the information of the Constituency-Level Elections Archive (Kollman et al., 2019), we match the post and the politician to the constituency that they were serving or running for at the time when they generated the post. [8]

As the boundaries of electoral districts in Indonesia and the Philippines have some minor changes during our time of interest, we manually assigned politicians to the correct constituency numbers according to the baseline constituency maps of choice (2014 constituency boundaries for Indonesia and Malaysia, and 2019 constituency boundary for the Philippines).

Table A-6: Summary Statistics of Politicians that Switch Constituencies

Country	Total number	Ran for multiple elections	Constituency switching
A. Politicians with identified social media handles			
The Philippines	1018	238	15
		(23.38%)	(1.47%)
Malaysia	720	230	50
		(31.94%)	(6.94%)
Indonesia	1291	456	128
		(35.32%)	(9.91%)
B. Politicians ever engaged in China-related conversation			
The Philippines	469	158	7
		(22.69%)	(1.49%)
Malaysia	492	207	46
		(42.07%)	(9.35%)
Indonesia	591	266	85
		(45.01%)	(14.38%)

Note: Panel A lists the number of all politicians with identified social media handles. Panel B restricts politicians to those who post about China on social media. Numbers indicate the number of politicians. Numbers in parentheses refer to the percentage of politicians that run for multiple elections or switch constituencies out of the total number of politicians.

Another challenge during the spatial-matching process is that political candidates often serve or campaign for different constituencies over the years. Table A-6 shows the number of politicians that run for multiple elections and the ratio of constituency switching. Most of the political candidates only ran for one election during our period of research, and the ratio of politicians

[8] We did not use the automatic detection of geolocation by Twitter due to its inconsistency and large number of missing values.

that switch constituencies is relatively low across the countries (1.5% to 15% of all politicians). Therefore, we believe that the approach of matching posts to the constituency should not affect our results much.

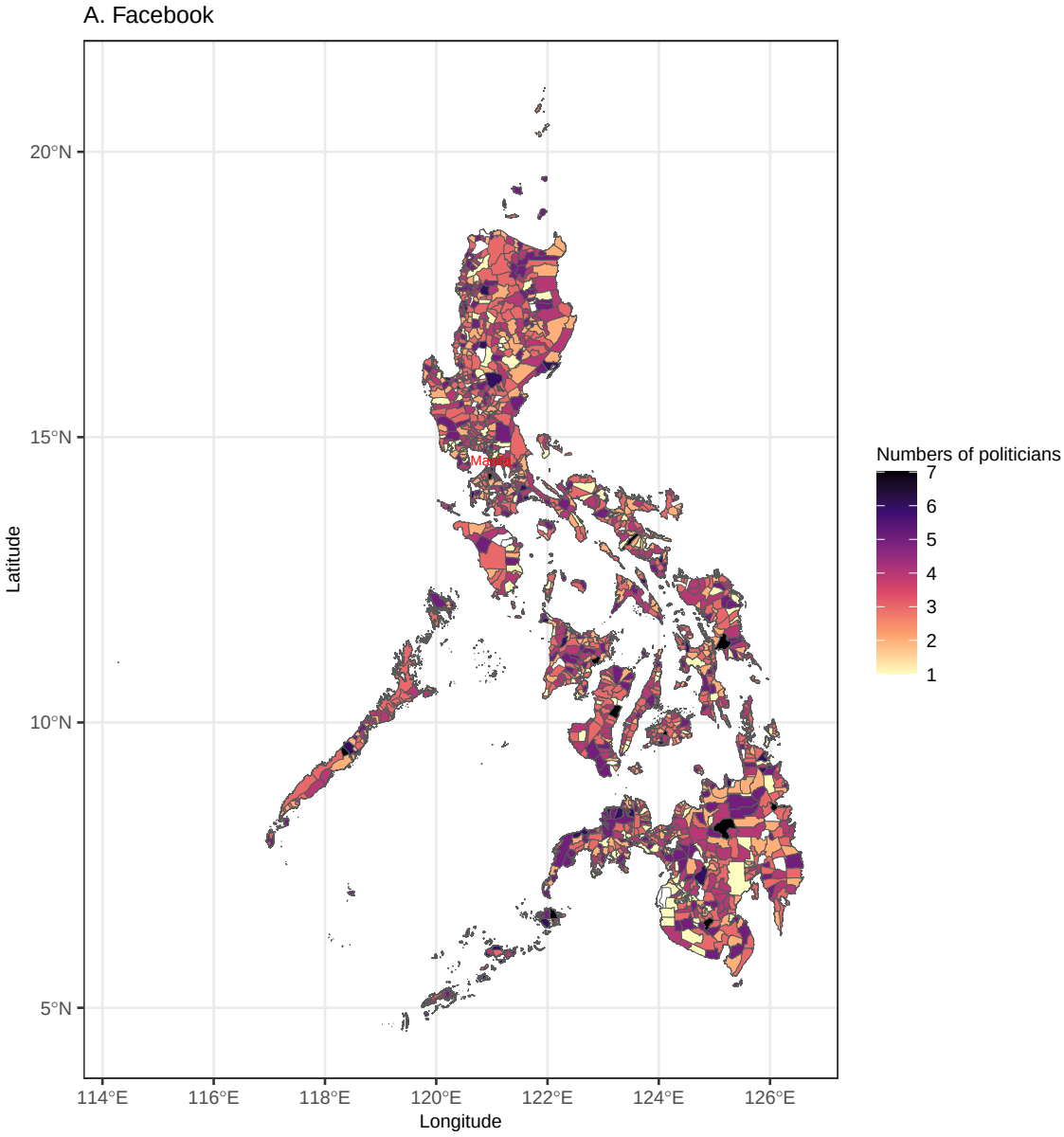
We can either 1) assign the relevant posts by only the constituencies they were serving or running for when the posts were generated (the approach we adopted), or 2) assign the relevant posts to all the constituencies for which they have ever served or campaigned. We selected the former method as it is more intuitive to implement. To illustrate the matching mechanism, suppose a candidate made a post about China on August 1, 2015, and she was an elected official in constituency No. 24 after the election in 2014. In the 2018 election, she no longer served in No. 24 but campaigned for office in constituency No. 25. In this case, her post on August 1, 2015, was assigned to constituency No. 24 but not No. 25. However, all her posts after 2018 are assigned to constituency No. 25. It would require stronger assumptions to implement the latter approach, for example, we need to argue that the politician had the thought of campaigning for constituency No. 25 from 2009 (the start of the period of interest).

For politicians that only campaigned for one constituency from 2009 to 2021 but authored many social media posts before their elections, we assign all their posts to the constituency they have ever served or campaigned for.

Figure A-13 visualize the geolocations of the political elites engaged in China-related topics for each country. Politicians who talk about China on social media are widespread across all constituencies in the Philippines, Malaysia, and Indonesia. The relatively balanced spatial distribution of politicians attenuates the endogeneity concern. Political candidates use both Facebook and Twitter in all three countries, although political candidates in the Philippines slightly prefer Facebook to Twitter for posting content about China.

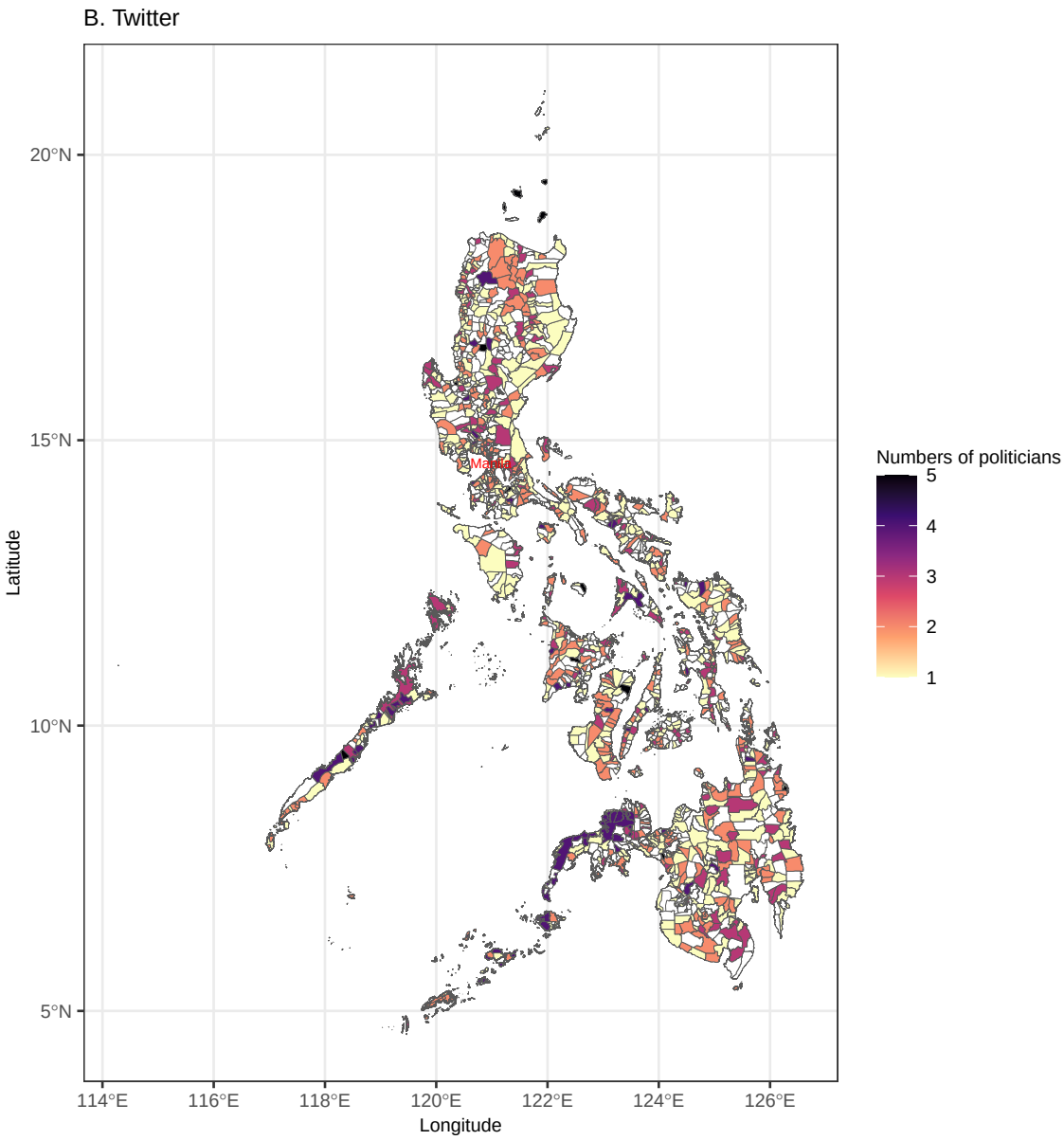
Figure A-13: Map of Politicians Engaged in China-related Topics

(a) The Philippines



Notes: Distribution of politicians in 243 constituencies. Base map uses constituency boundaries in 2019.

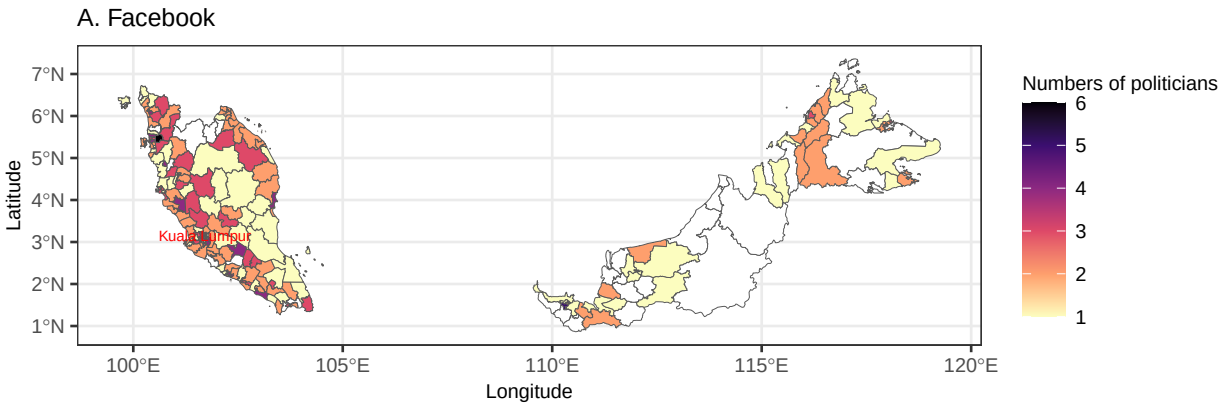
Figure A-13: Map of Politicians Engaged in China-related Topics (continued)



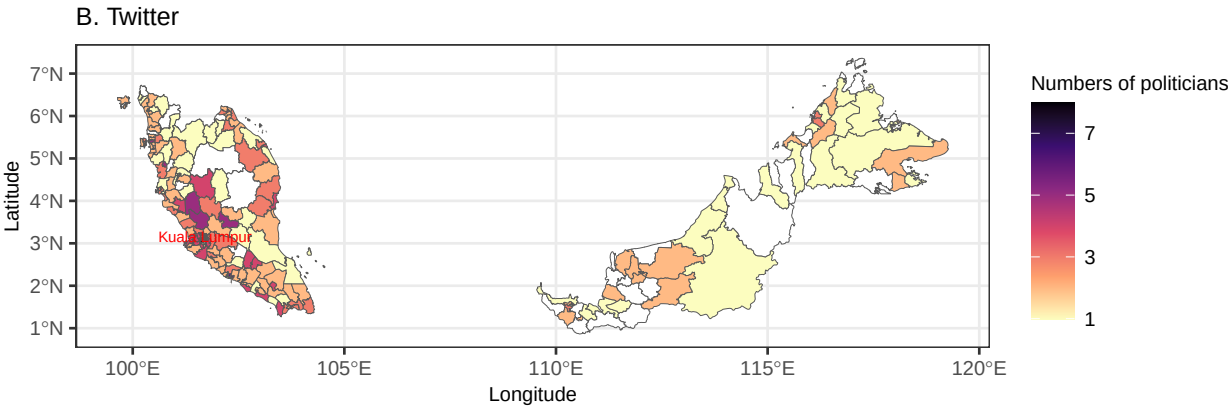
Notes: Distribution of politicians in 243 constituencies. Base map uses constituency boundaries in 2019.

Figure A-13: Map of Politicians Engaged in China-related Topics (continued)

(b) Malaysia



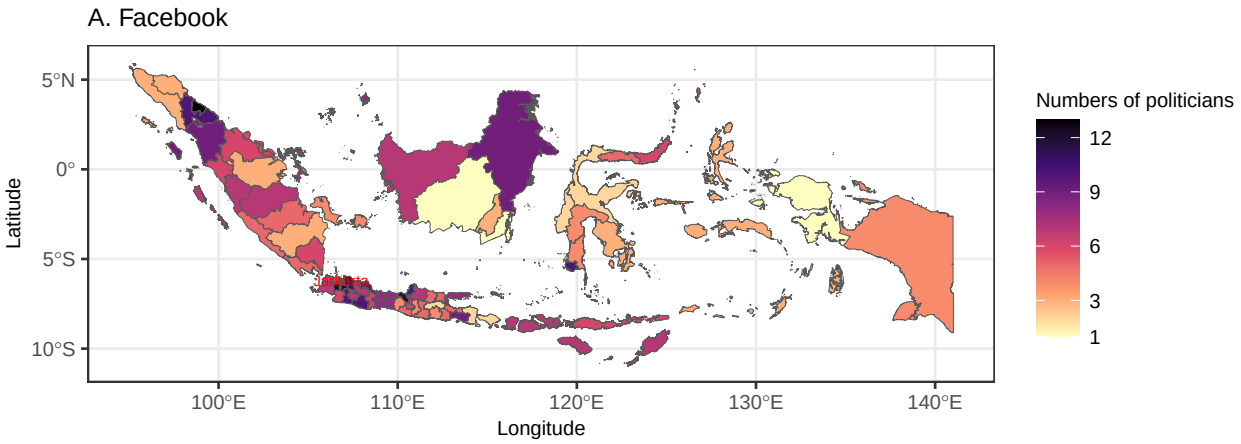
Notes: Distribution of politicians in 222 constituencies. Base map uses constituency boundaries in 2014.



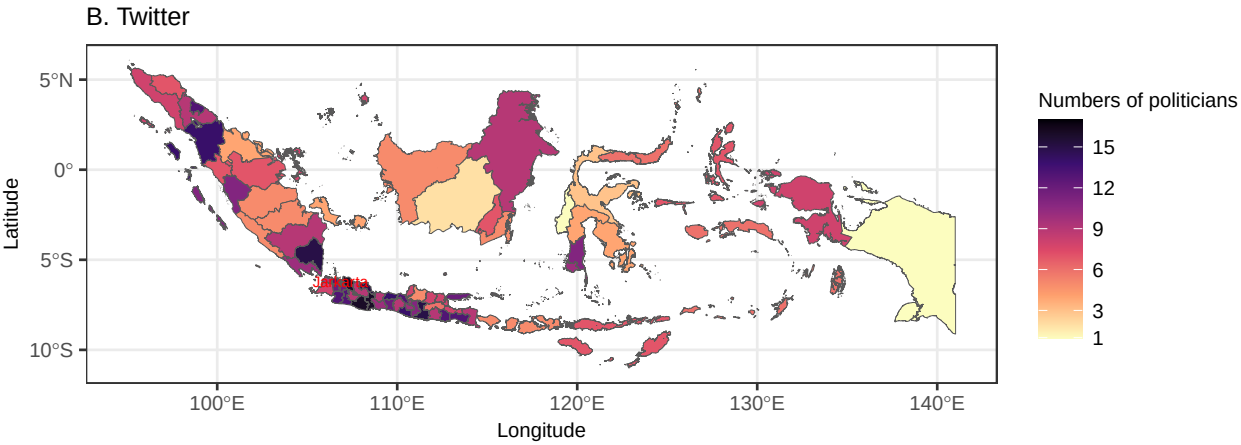
Notes: Distribution of politicians in 222 constituencies. Base map uses constituency boundaries in 2014.

Figure A-13: Map of Politicians Engaged in China-related Topics (continued)

(c) Indonesia



Notes: Distribution of politicians in 77 constituencies. Base map uses constituency boundaries in 2009−2014.



Notes: Distribution of politicians in 77 constituencies. Base map uses constituency boundaries in 2009−2014.

1.4 Unit of Analysis

To construct the full dataset for analysis, we merged the BRI treatment and elite sentiment by constituency and date and aggregated the data into two different units of analysis.

1.4.1 Constituency-date level

We aggregated the number of social media posts and the average sentiment score by each constituency-date pair. For days that did not witness any social media posts in a specific constituency, the number of posts equals 0 and the average sentiment score is assigned 0. The assumption is that if there is no social media conversation about China on that day, the politicians do

not have strong emotions toward China. However, one can argue that politicians may avoid posting negative opinions against China publicly, and the absence of posts about China may indicate negative attitudes. We address this possibility in the discussion section.

We also constructed the cumulative number of BRI projects by constituency-date. For example, if constituency C received its first Chinese project on January 1, 2016, then the cumulative number of BRI projects is coded as 1 for all pairs after January 1, 2016. The cumulative number would turn 2 on the date when constituency C received another BRI project.

1.4.2 Politician-date level

For the constituency-date level analysis, we expect elected political officials and their challengers in the same constituency to share certain traits and treat local politics as a black box. It is also possible that political candidates hold opinions toward China for reasons not related to the region where their voters are. We constructed a second unit of analysis at the politician-date level, where the number of social media posts and the average sentiment score are collapsed by each politician-date pair. Similarly, when the politician did not post anything about China, her number of posts and sentiment were assigned both 0 on that day.

We also constructed the cumulative number of BRI projects by politician-date. For example, imagine candidate A campaigned for office in constituency No.1 in 2015 and campaigned for office in constituency No.5 in 2019. Suppose constituency No.1 received two projects in 2016 and constituency No.5 did not receive any projects up until now, the cumulative number of BRI projects for candidate A from 2016 to 2019 is two, and becomes zero after 2019.

We present results from both units of analysis. As the results aggregated at both levels are overall similar, we include the constituency-date level analysis as the default baseline unless otherwise noted.

2 Descriptive Analysis

2.1 Summary Statistics

Table A-7 reports the summary statistics of the outcome variables in the event-specific dataset, each 61-day event window centered on the shock of South China Sea disputes. For each outcome measurement, the first row lists the statistics of social media behaviors toward China from all political elites, the second row of politicians in constituencies that had received BRI investment before the shock, and the third row of politicians in constituencies that had not yet received any BRI investment before the particular event.

2.2 Content Classification

We applied the "content classification" feature of the Google Natural Language API to investigate what political elites talk about China across years [9]. The API returns one or more content categories for each social media post with a confidence score for each category [10]. Exploiting that feature, we constructed a frequency variable to filter the most engaging topics about China on social media ($Frequency_i = \sum confidence_i$).

Figure A-14 shows the top 12 categories that politicians talk about from 2011 to 2021. The time trend of each topic can be found in figure A-15. For politicians in the Philippines, their discussion about China is mostly about the Covid-19 pandemic and international news. Filipino politicians also increasingly discussed K-12 education and law and government when they mentioned China on Facebook starting in 2019. On Twitter, discussions about war and conflict, law and government, and political news peaked in 2017 in the Philippines.

Sharing news about China is more prevalent in Malaysia. Politicians often engaged in discussions about law and government and social issues such as migrant workers. The conversation about most topics gradually increased starting from 2011 on Facebook, though politicians slowly decreased their posting about political news on Twitter over the years. The discussion about Covid-19 and public health surged in 2020 on both Facebook and Twitter.

[9] Google Natural Language API - Content Classification: `https://cloud.google.com/natural-language/docs/classifying-text`

[10] Full list of content categories by Google Natural Language API: `https://cloud.google.com/natural-language/docs/categories`

Table A-7: Summary Statistics of Outcome Variables in the Stacked Event Study Design

		N	Mean	St. Dev.	Min	Median	Max
The Philippines							
Number of Facebook Posts		57,706	0.069	0.447	0	0	18
	BRI = 1	18,178	0.090	0.537	0	0	11
	BRI = 0	39,528	0.060	0.398	0	0	18
Number of tweets		57,706	0.026	0.331	0	0	23
	BRI = 1	18,178	0.005	0.086	0	0	4
	BRI = 0	39,528	0.035	0.395	0	0	23
Facebook Sentiment		57,706	0.005	0.078	−1	0	1
	BRI = 1	18,178	0.007	0.086	−1	0	1
	BRI = 0	39,528	0.005	0.074	−1	0	1
Tweet Sentiment		57,706	0.001	0.050	−1	0	1
	BRI = 1	18,178	0.00005	0.031	−1	0	1
	BRI = 0	39,528	0.001	0.057	−1	0	1
Malaysia							
Number of Facebook Posts		38,735	0.087	0.422	0	0	13
	BRI = 1	11,651	0.116	0.423	0	0	8
	BRI = 0	27,084	0.074	0.421	0	0	13
Number of tweets		38,735	0.066	0.371	0	0	13
	BRI = 1	11,651	0.104	0.459	0	0	8
	BRI = 0	27,084	0.049	0.325	0	0	13
Facebook Sentiment		38,735	−0.007	0.103	−1	0	1
	BRI = 1	11,651	−0.012	0.128	−1	0	1
	BRI = 0	27,084	−0.004	0.090	−1	0	1
Tweet Sentiment		38,735	−0.004	0.096	−1	0	1
	BRI = 1	11,651	−0.006	0.118	−1	0	1
	BRI = 0	27,084	−0.003	0.085	−1	0	1
Indonesia							
Number of Facebook Posts		9,394	0.018	0.203	0	0	14
	BRI = 1	4,392	0.017	0.246	0	0	14
	BRI = 0	5,002	0.019	0.156	0	0	3
Number of tweets		9,394	0.108	0.640	0	0	18
	BRI = 1	4,392	0.096	0.633	0	0	12
	BRI = 0	5,002	0.119	0.646	0	0	18
Facebook Sentiment		9,394	0.0001	0.041	−1	0	1
	BRI = 1	4,392	−0.0002	0.039	−1	0	1
	BRI = 0	5,002	0.0004	0.044	−1	0	1
Tweet Sentiment		9,394	−0.005	0.094	−1	0	1
	BRI = 1	4,392	−0.004	0.088	−1	0	1
	BRI = 0	5,002	−0.006	0.100	−1	0	1

Note: Summary statistics of outcome variables in the event-specific dataset for each country following the identification strategy. Sentiment scores are calculated with the Google Natural Language API. Observations are collapsed at the constituency-date level and subset to the 61-day event window centered at the event dates listed in table A-1 for dates in $[k_{e-30}, k_{e+30}]$. For each variable, the first row lists the statistics of observations that locate in constituencies had ever received BRI investment before the shock of each event, and the second row lists the statistics of observations that locate in constituencies had never received BRI investment before the shock of each event. As there are multiple event shocks for each country, certain constituencies might be untreated in the first shock but become treated in later shocks. In that case, these constituencies are dropped from the control group in the first event episode but included as the treated group in later events. We dropped 16 constituencies from the control group in the Philippines, 31 in Malaysia, and none in Indonesia.

In Indonesia, the discussion about all categories gradually rises from 2011 to 2020 on both platforms. Among all topics, Indonesian politicians are most interested in law and government and news about China. Interestingly, not only discussion on public health peaked in 2020, but also conversations about business and industry, social issues, and news about China in general. The frequency of these discussions returned to the baseline level in 2021.

In figure A-16 we subset the social media posts to the 61-day time window according to the stacked event study design and then centered the frequency by the dates of the shocks. We do not find strong differences in the topics before and after the South China Sea disputes in general. However, discussion on violence and conflict surged on the date of shocks on the Twitter platform in the Philippines.

Figure A-14: Top 12 Categories of Politician Posts about China in General

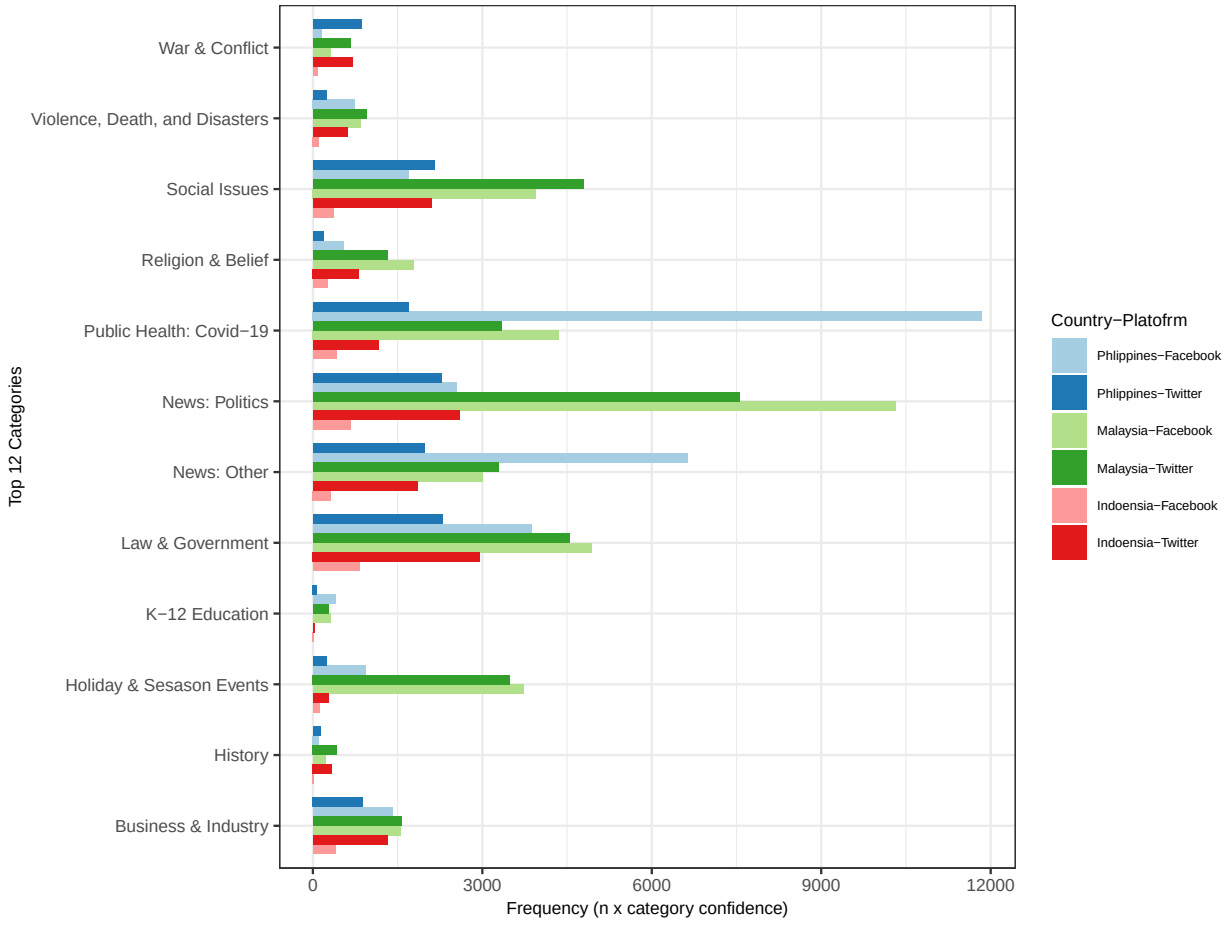


Figure A-15: Time Trend of Top 12 Categories of Politician Posts about China

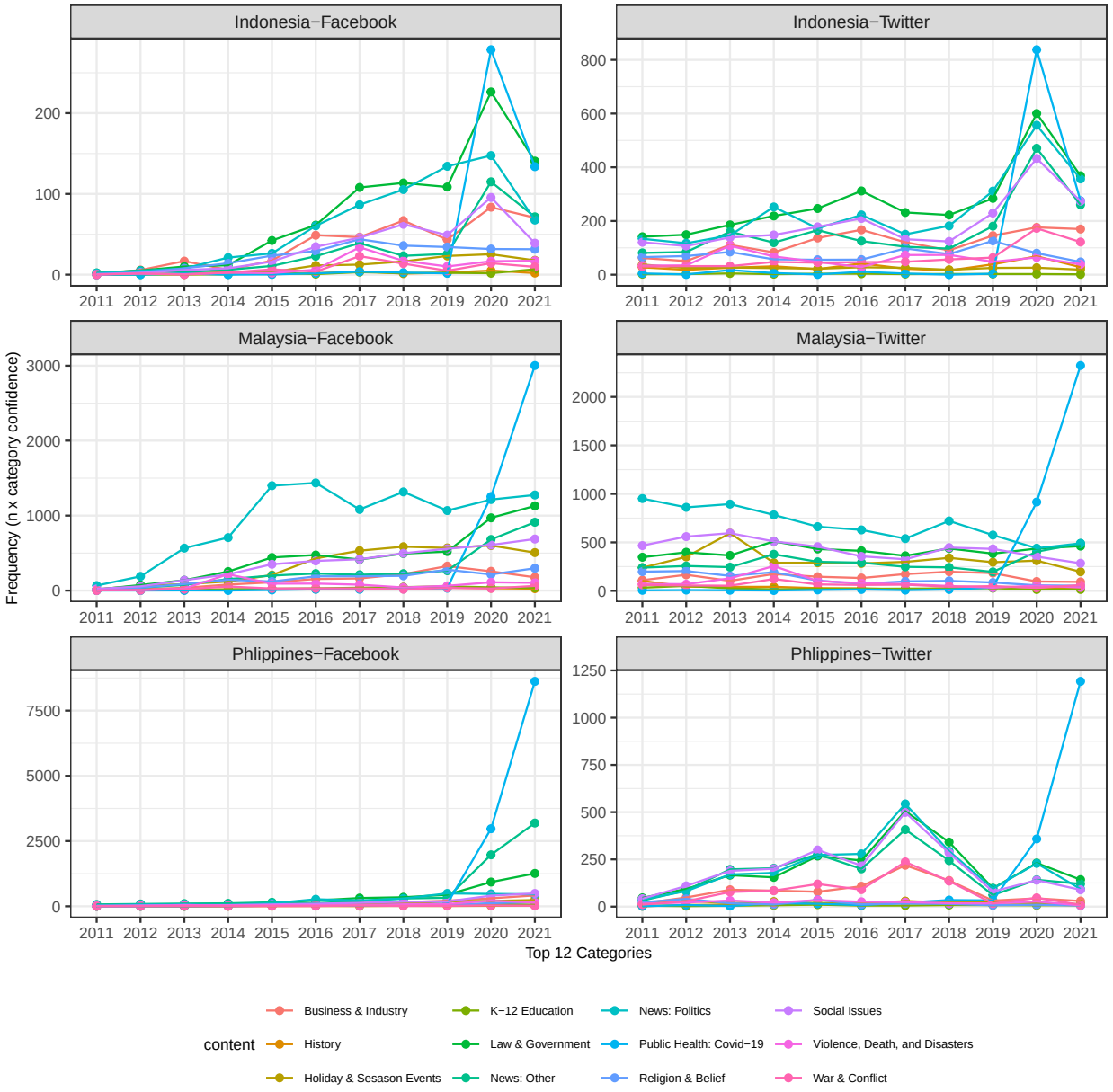
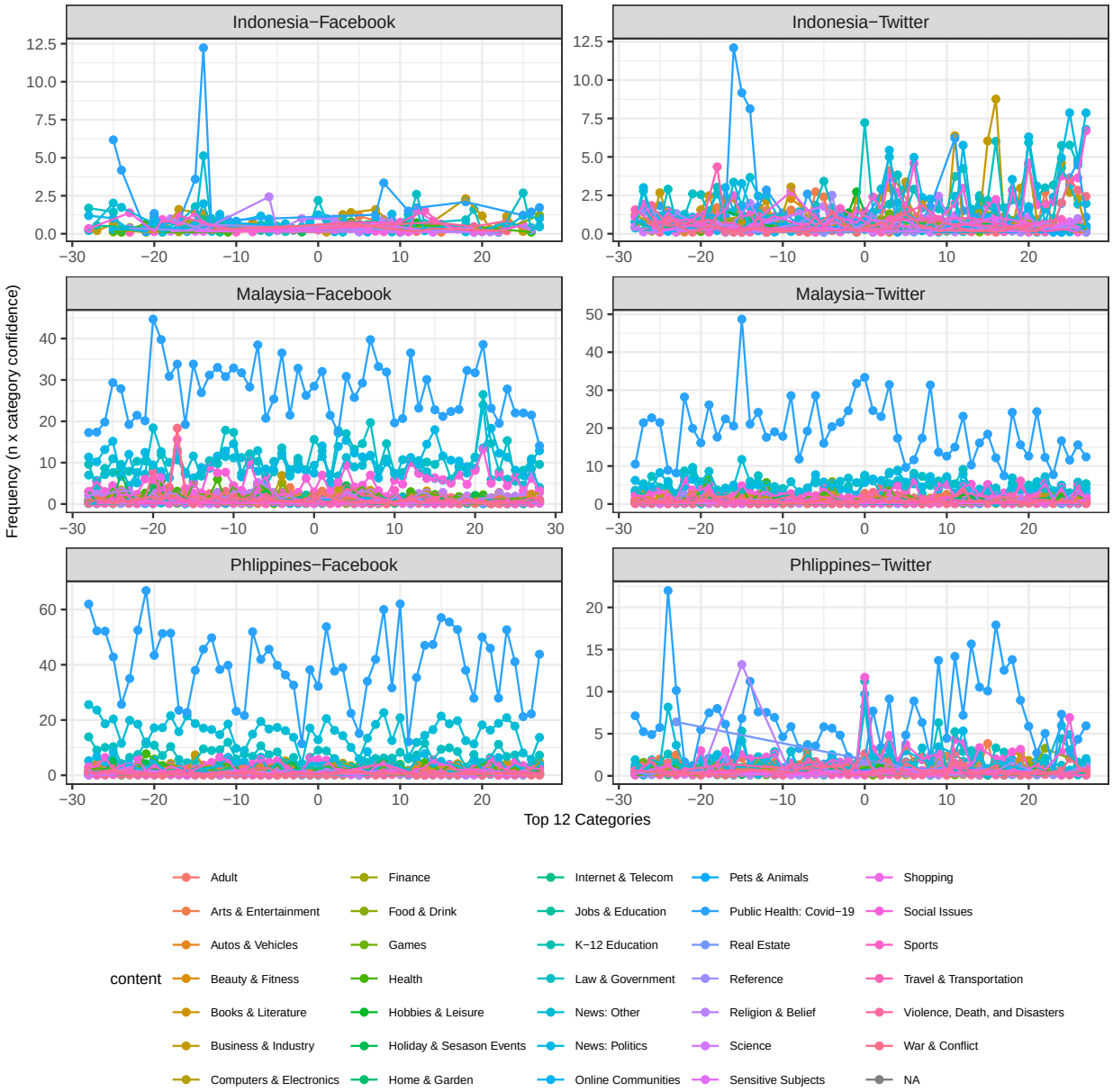


Figure A-16: Categories of Politician Posts about China Around Shocks



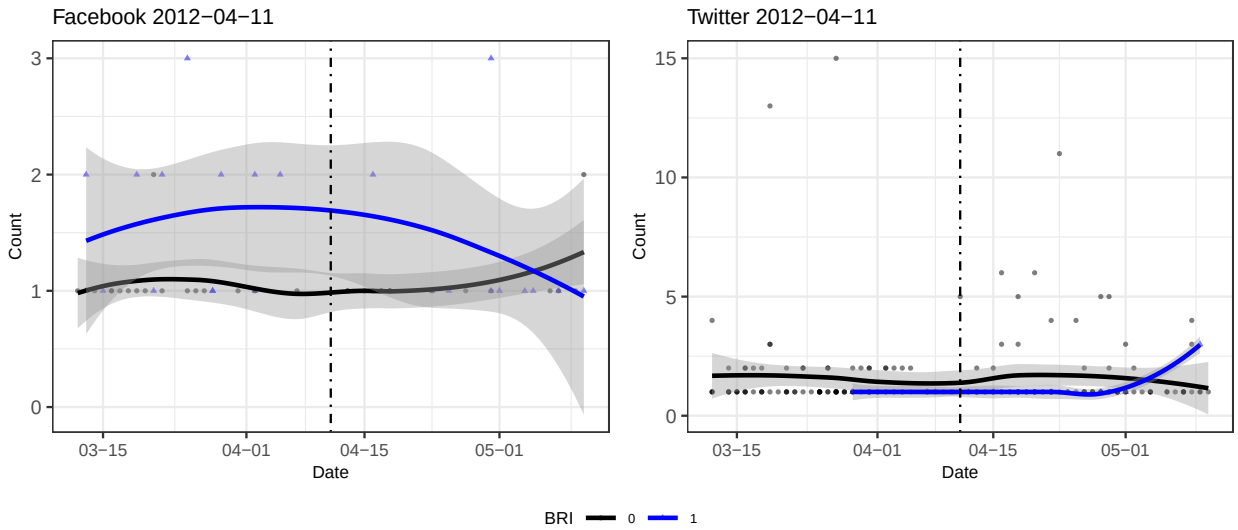
2.3 Number of Posts by Chinese Investment Around Shocks

We compare the number of social media posts before and after shocks by the treatment status of the constituency. Figure A-19 shows the results of the unbalanced panel, where constituencies without social media posts about China on a particular date are dropped on that date. Each dot in the graphs represents one constituency with at least one posting about China on that specific date; the position of the points represents the total number of posts for one constituency on that

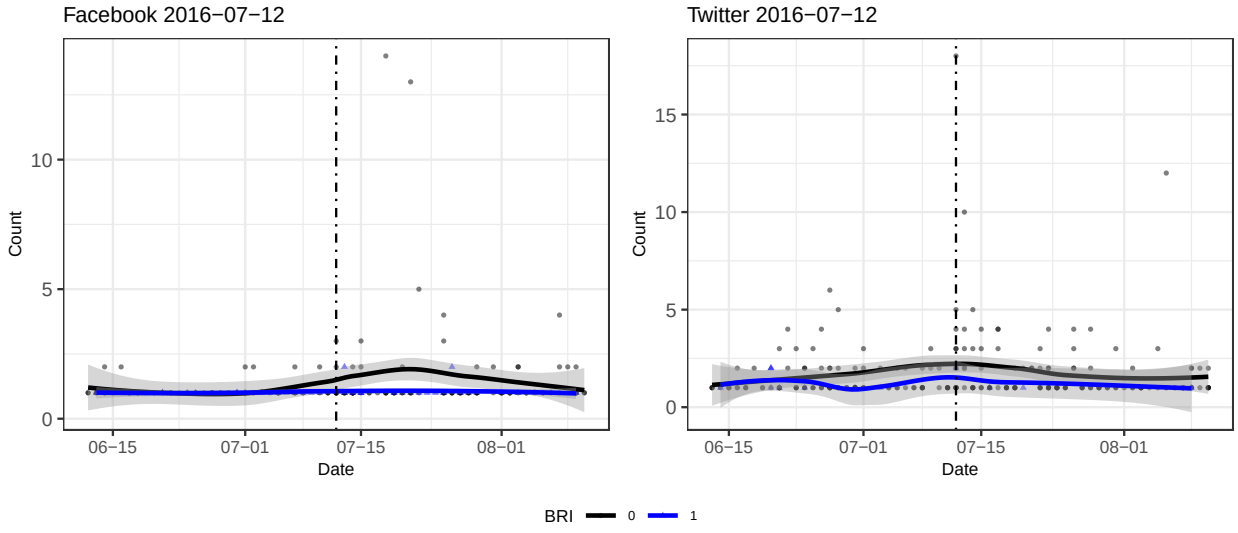
date; and the color of points indicates the BRI treatment status. Figure A-20 shows the descriptive analysis of the balanced constituency-date panel data. For constituencies without any posts about China on a particular date, the number of posts equals 0. In the next section, we present the descriptive analysis of the sentiment scores following these two panel construction approaches and discuss the assumptions behind the balanced panel constructions and the reason why we mainly rely on the balanced panel data for regression analysis.

Figure A-17: Number of Posts about China Grouped by Project Treatment Status: Unbalanced Panel

(a) The Philippines

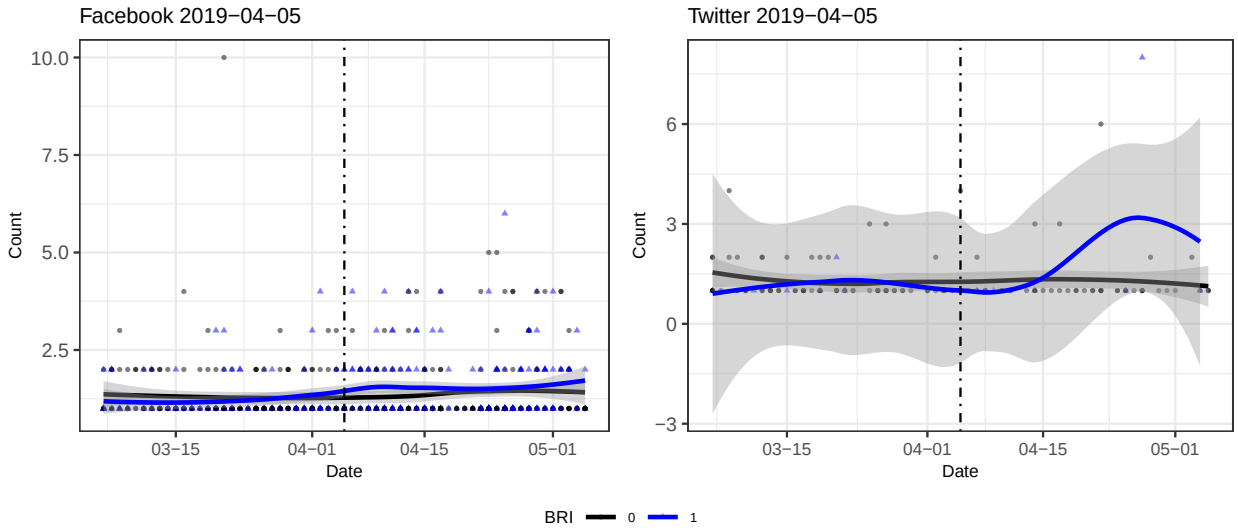


Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped. 243 constituencies and 55 constituency−date level observations of elite sentiment on Facebook, and 199 observations on Twitter within the 61−day event window in total.

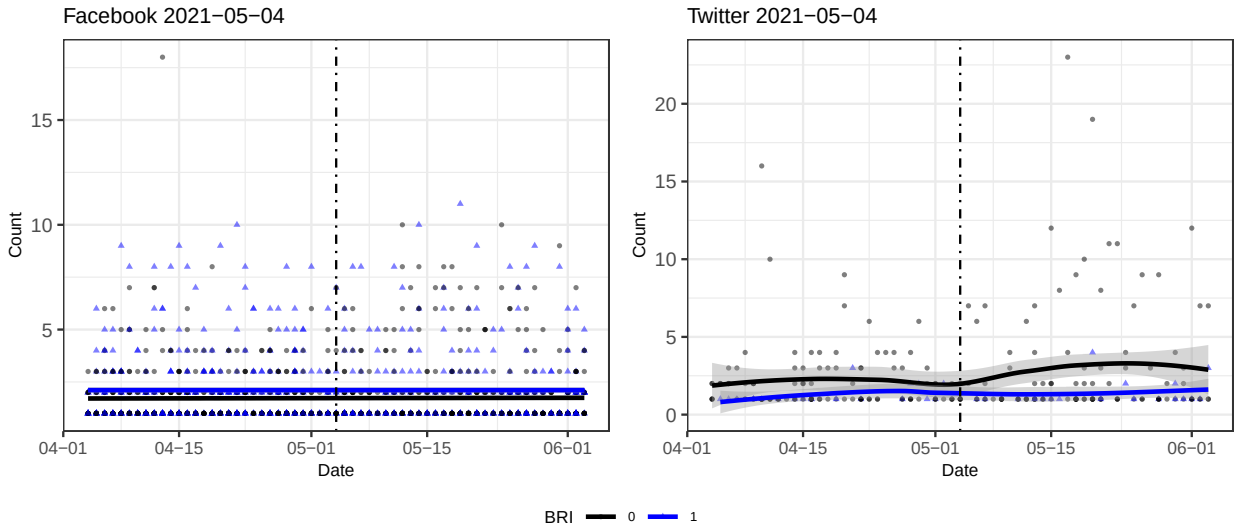


Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped. 243 constituencies and 215 constituency−date level observations of elite sentiment on Facebook, and 218 observations on Twitter within the 61−day event window in total.

Figure A-17: Number of Posts about China Grouped by Project Treatment Status: Unbalanced Panel (continued)



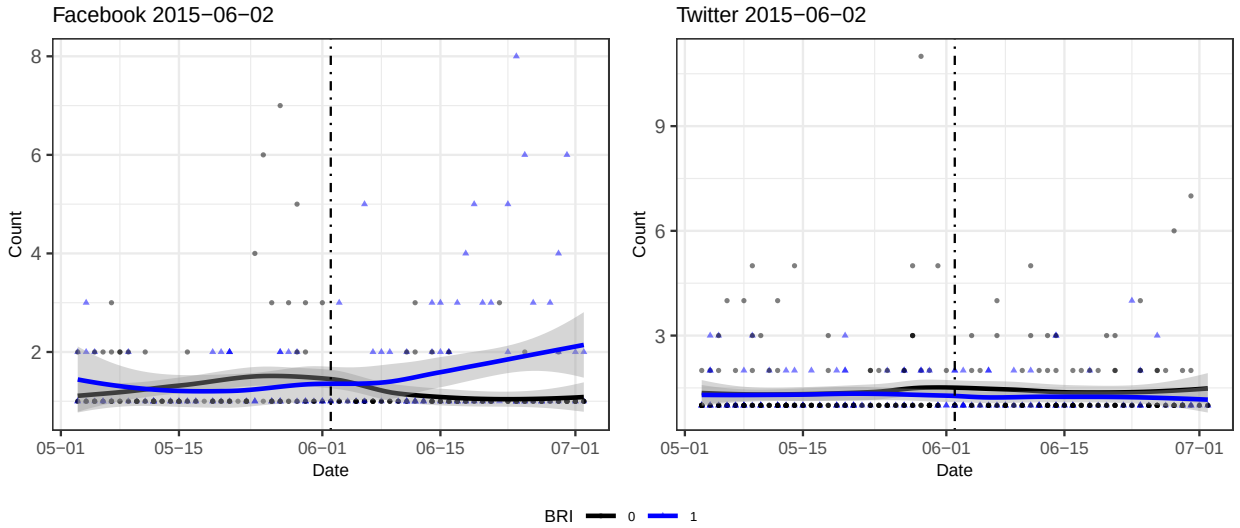
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
243 constituencies and 675 constituency−date level observations of elite sentiment on Facebook, and 127 observations on Twitter within the 61−day event window in total.



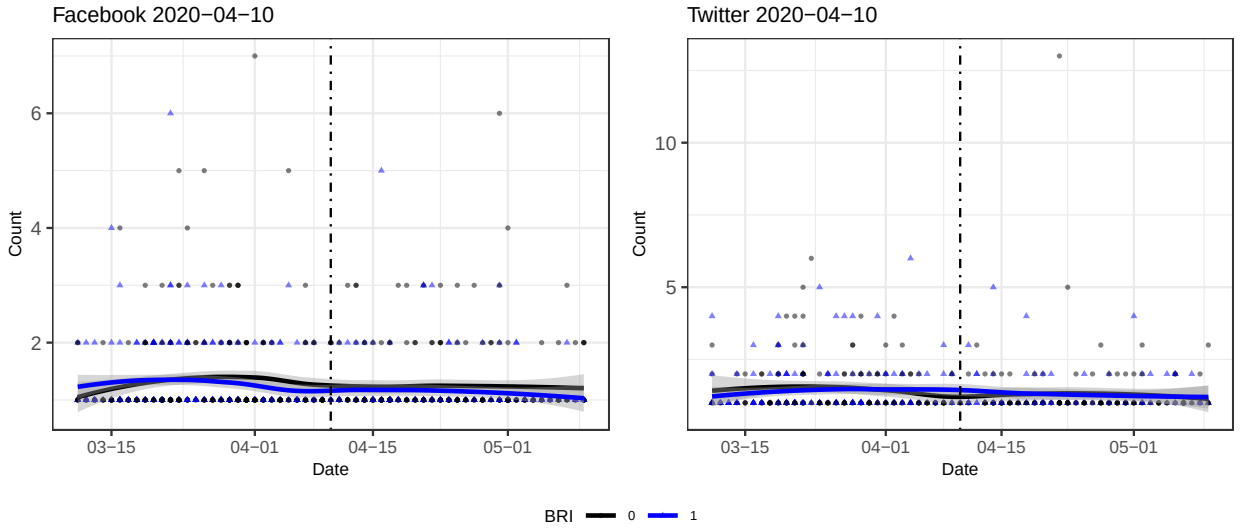
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
243 constituencies and 1664 constituency−date level observations of elite sentiment on Facebook, and 250 observations on Twitter within the 61−day event window in total.

Figure A-17: Number of Posts about China Grouped by Project Treatment Status: Unbalanced Panel (continued)

(b) Malaysia

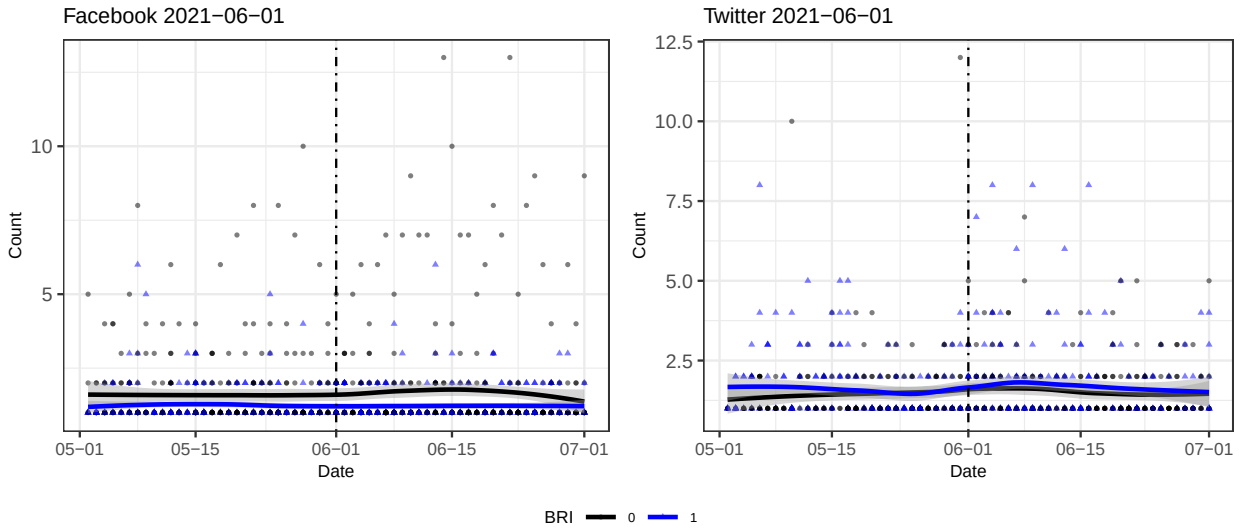


Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
222 constituencies and 376 constituency−date level observations of elite sentiment on Facebook, and 465 observations on Twitter within the 61−day event window in total.



Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
222 constituencies and 915 constituency−date level observations of elite sentiment on Facebook, and 537 observations on Twitter within the 61−day event window in total.

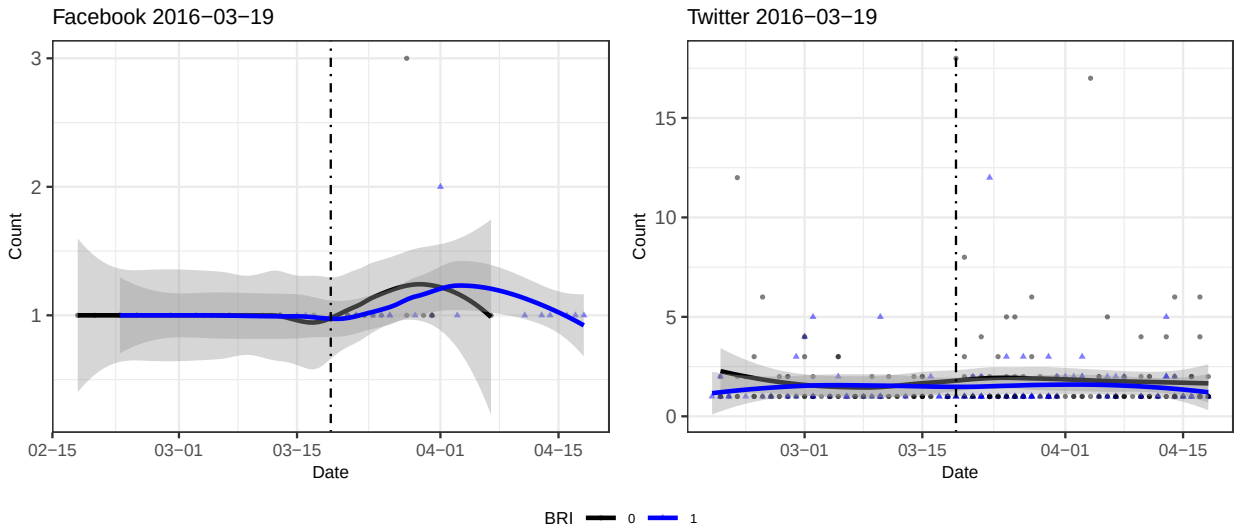
Figure A-17: Number of Posts about China Grouped by Project Treatment Status: Unbalanced Panel (continued)



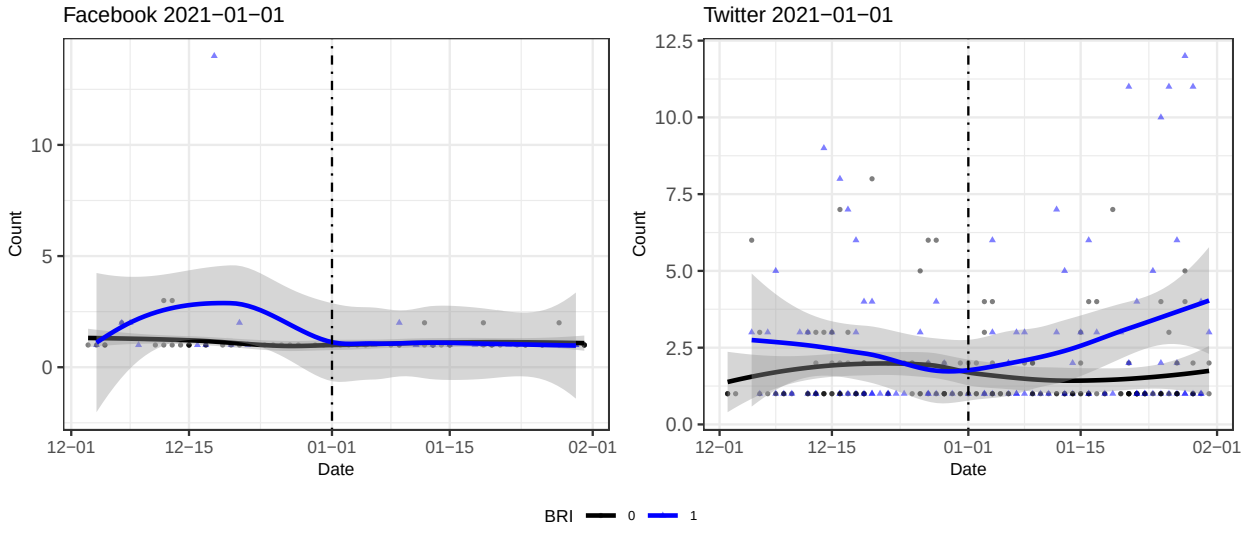
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped. 222 constituencies and 1201 constituency−date level observations of elite sentiment on Facebook, and 831 observations on Twitter within the 61−day event window in total.

Figure A-17: Number of Posts about China Grouped by Project Treatment Status: Unbalanced Panel (continued)

(c) Indonesia



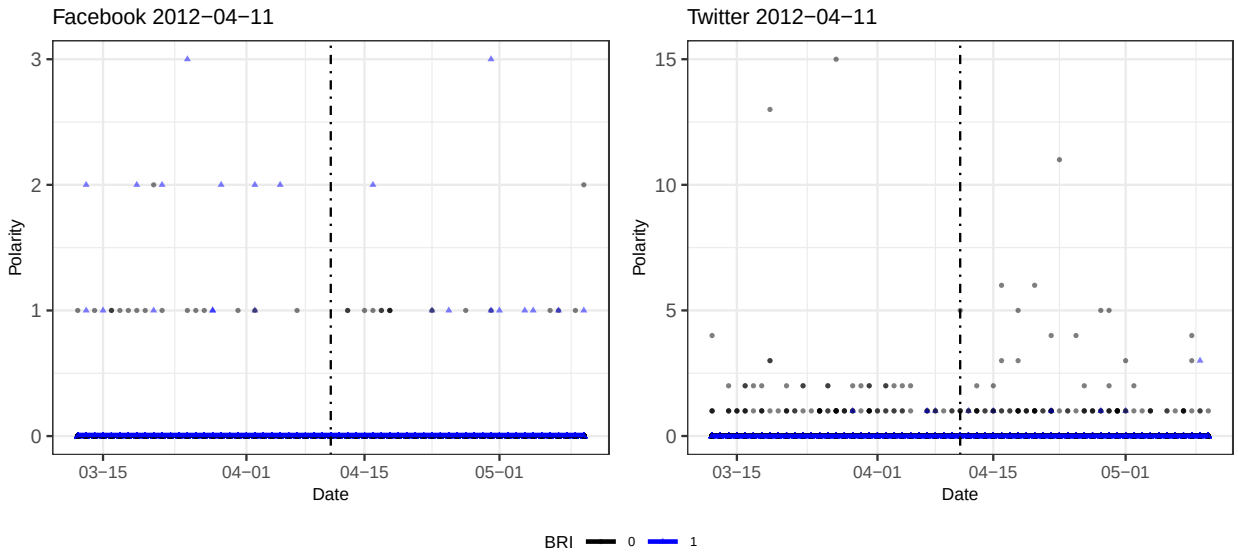
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped. 77 constituencies and 53 constituency−date level observations of elite sentiment on Facebook, and 303 observations on Twitter within the 61−day event window in total.



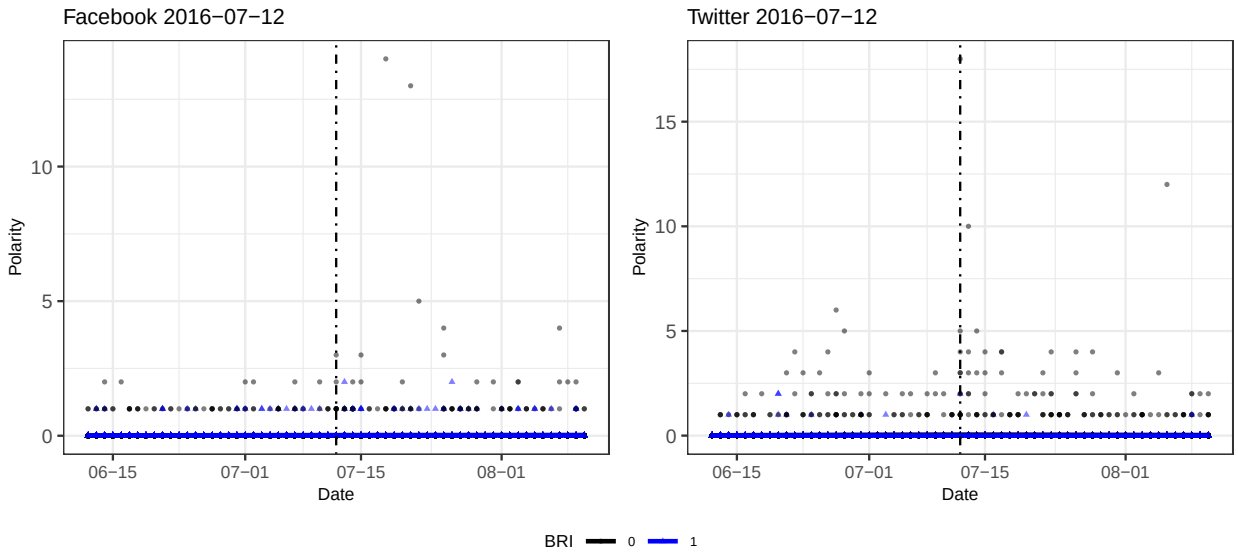
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped. 77 constituencies and 91 constituency−date level observations of elite sentiment on Facebook, and 250 observations on Twitter within the 61−day event window in total.

Figure A-18: Number of Posts about China Grouped by Project Treatment Status: Balanced Panel

(a) The Philippines

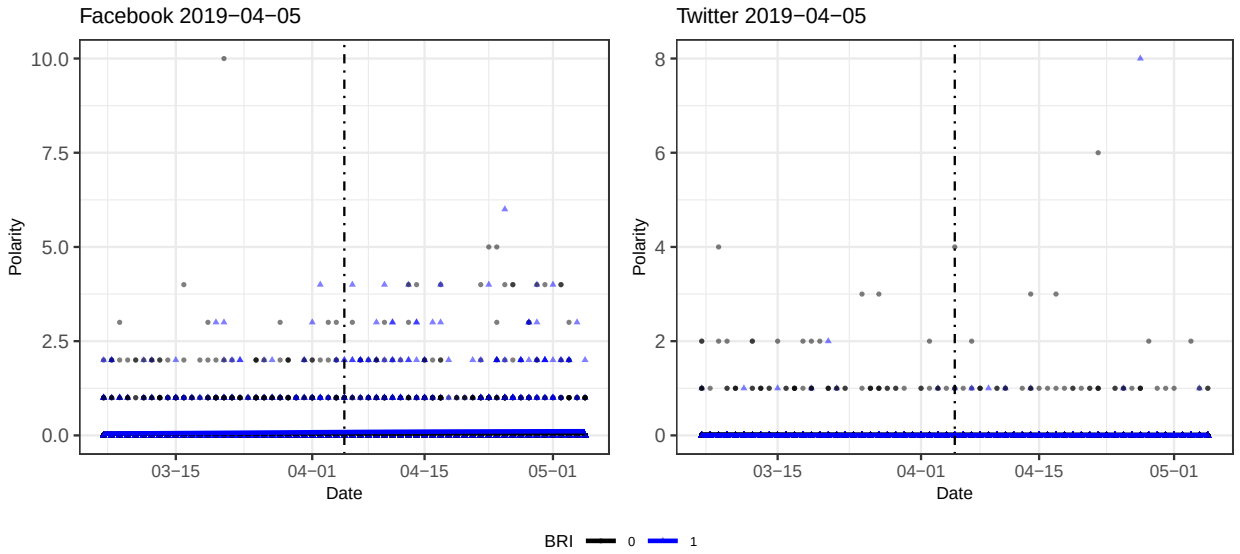


Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0.
243 constituencies and 14832 constituency−date level observations of elite posts within the 61−day event window in total.

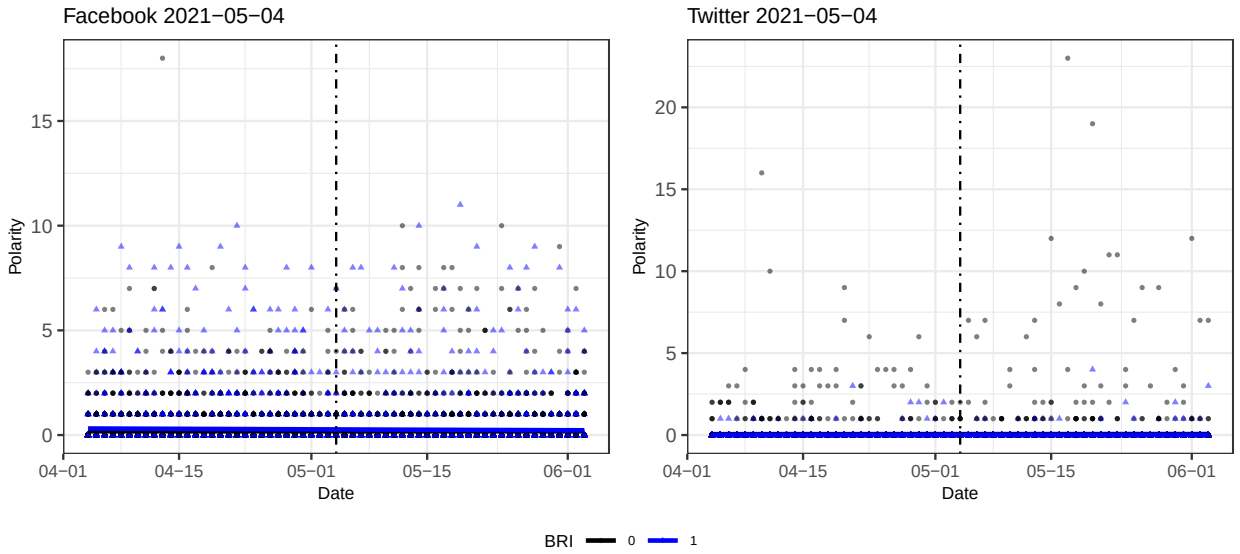


Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0.
243 constituencies and 14832 constituency−date level observations of elite posts within the 61−day event window in total.

Figure A-18: Number of Posts about China Grouped by Project Treatment Status: Balanced Panel (continued)



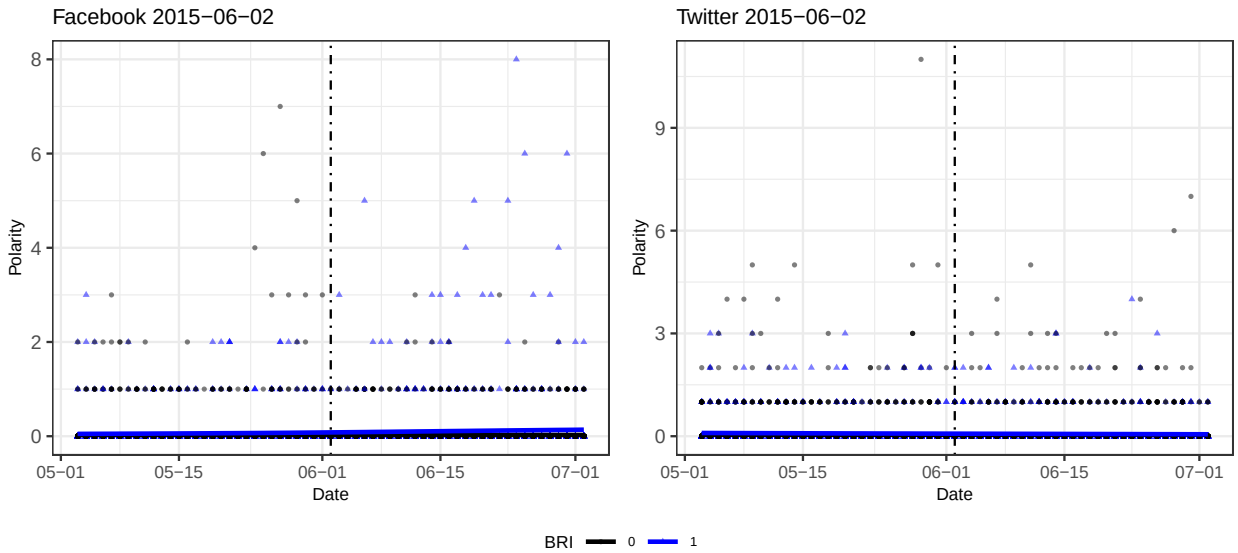
Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0. 243 constituencies and 14832 constituency−date level observations of elite posts within the 61−day event window in total.



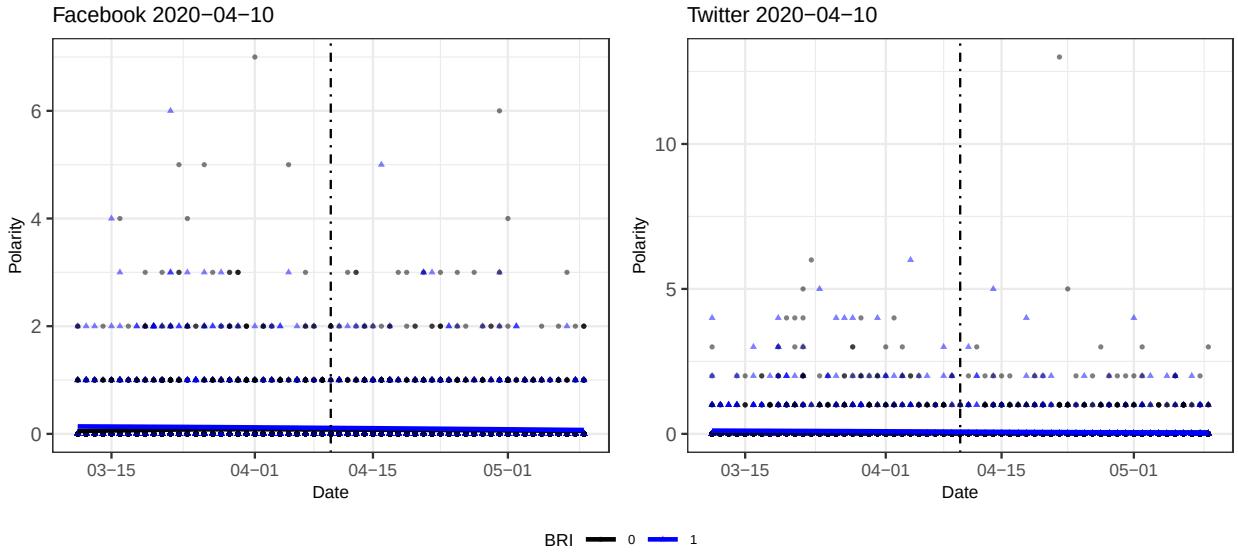
Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0. 243 constituencies and 14832 constituency−date level observations of elite posts within the 61−day event window in total.

Figure A-18: Number of Posts about China Grouped by Project Treatment Status: Balanced Panel (continued)

(b) Malaysia

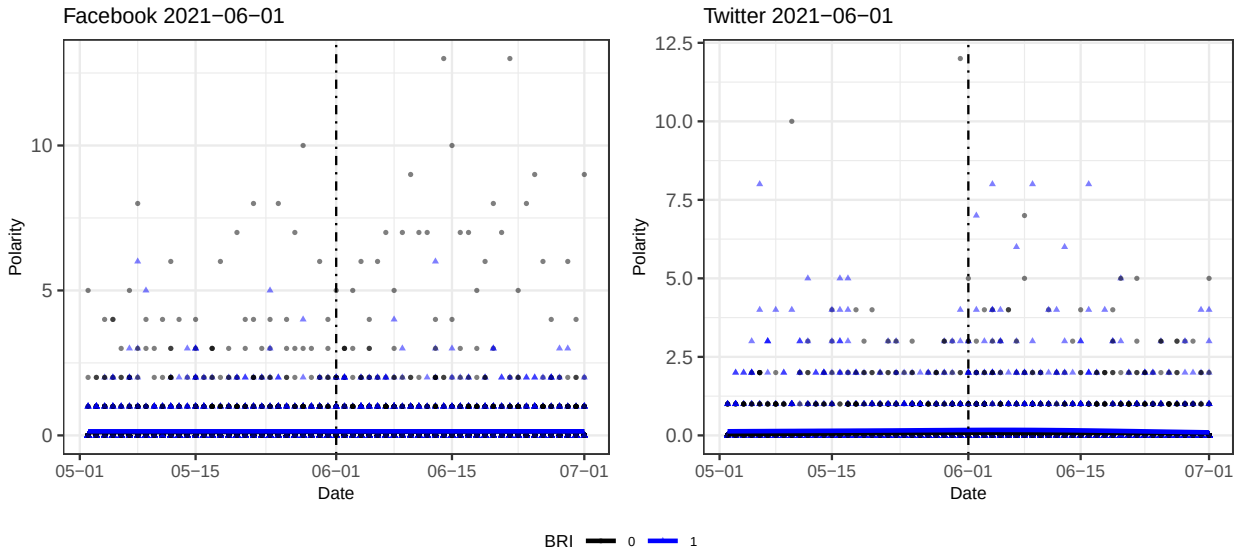


Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0. 222 constituencies and 13542 constituency−date level observations of elite posts within the 61−day event window in total.



Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0. 222 constituencies and 13542 constituency−date level observations of elite posts within the 61−day event window in total.

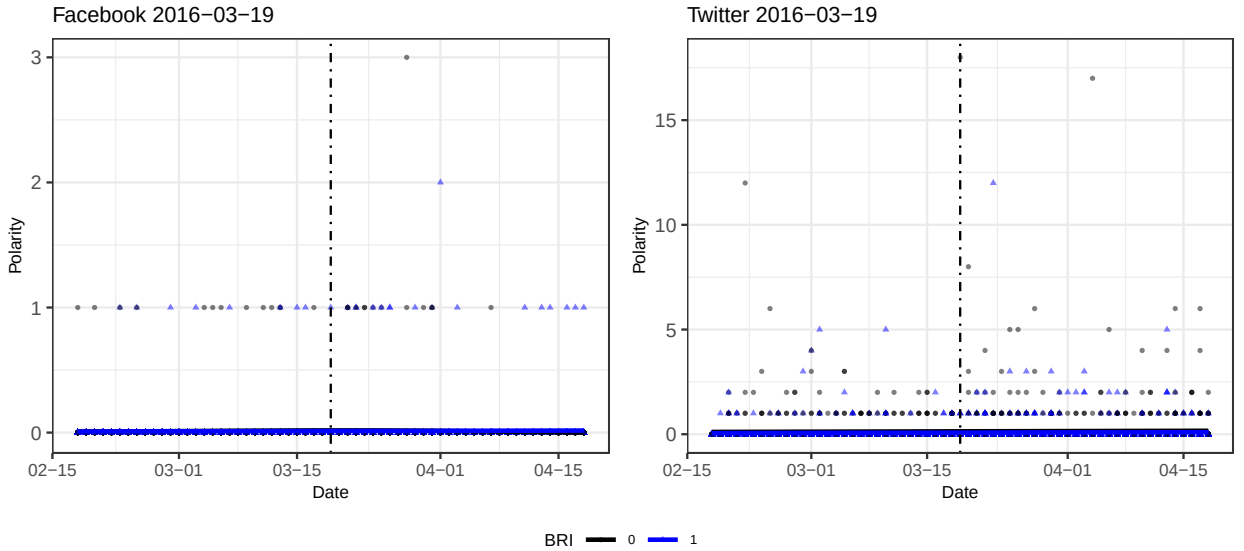
Figure A-18: Number of Posts about China Grouped by Project Treatment Status: Balanced Panel (continued)



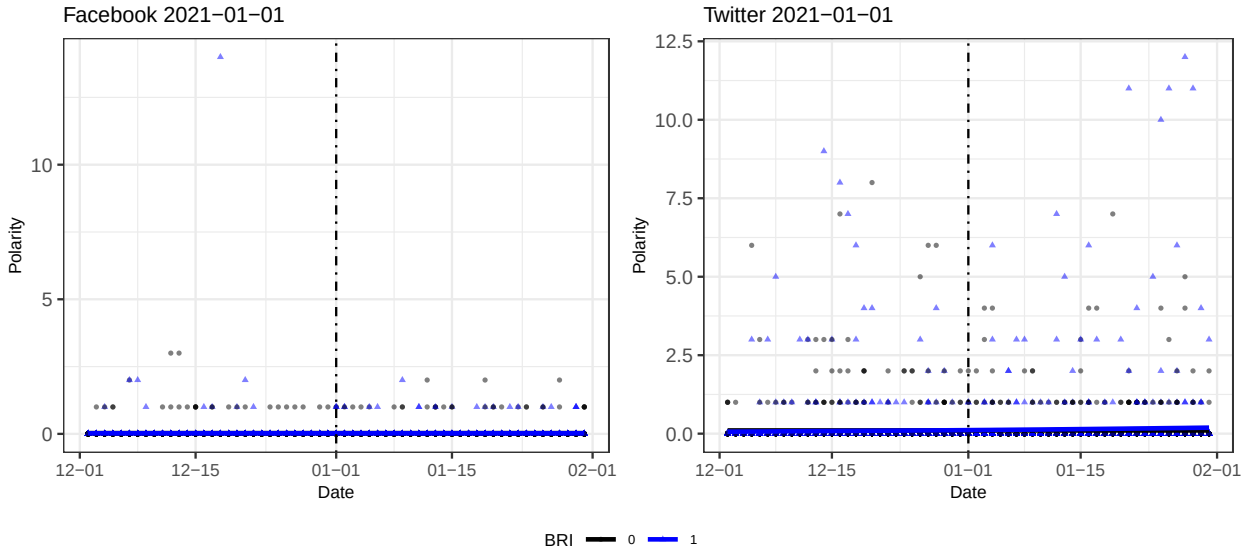
Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0. 222 constituencies and 13542 constituency−date level observations of elite posts within the 61−day event window in total.

Figure A-18: Number of Posts about China Grouped by Project Treatment Status: Balanced Panel (continued)

(c) Indonesia



Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0.
77 constituencies and 4697 constituency−date level observations of elite posts within the 61−day event window in total.



Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the number of posts will be set to 0.
77 constituencies and 4697 constituency−date level observations of elite posts within the 61−day event window in total.

2.4 Elite Sentiment by Chinese Investment Around Shocks

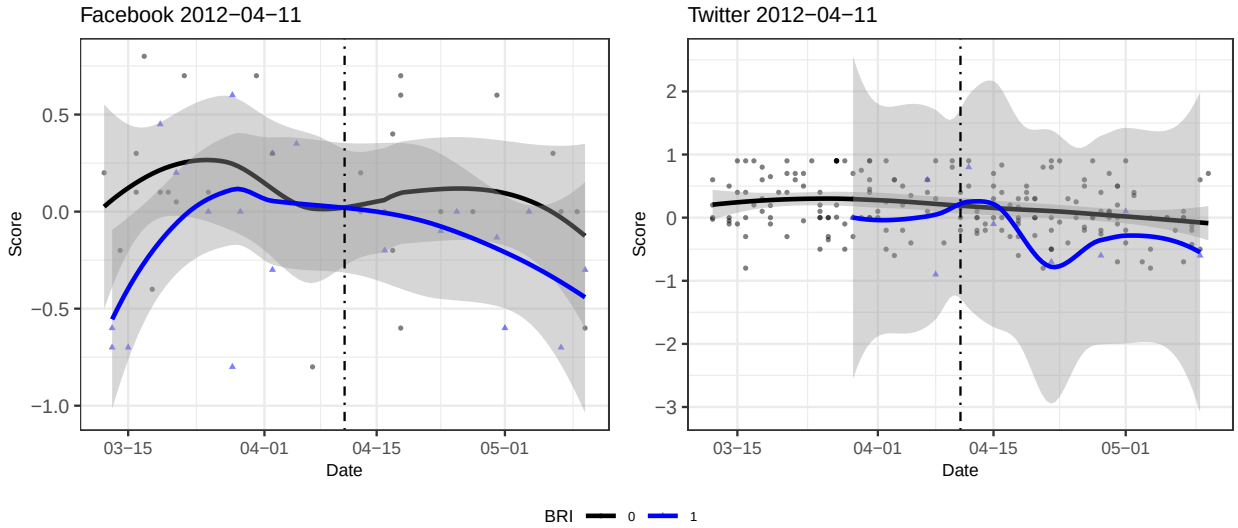
We present the descriptive analysis to compare elite sentiment toward China before and after shocks by the treatment status. Figure A-19 shows the results of the unbalanced panel, where constituencies without social media posts about China on a particular date are dropped on that

date. Each dot in the graphs represents one constituency with at least one posting about China on that specific date; the position of the points represents the calculated sentiment score of that social media post; and the color of points indicates the BRI treatment status. Figure A-20 shows the descriptive analysis of the balanced constituency-date panel data. For constituencies without any posts about China on a particular date, the number of posts equals 0, and the sentiment scores of that constituency-date pair are also assigned 0.

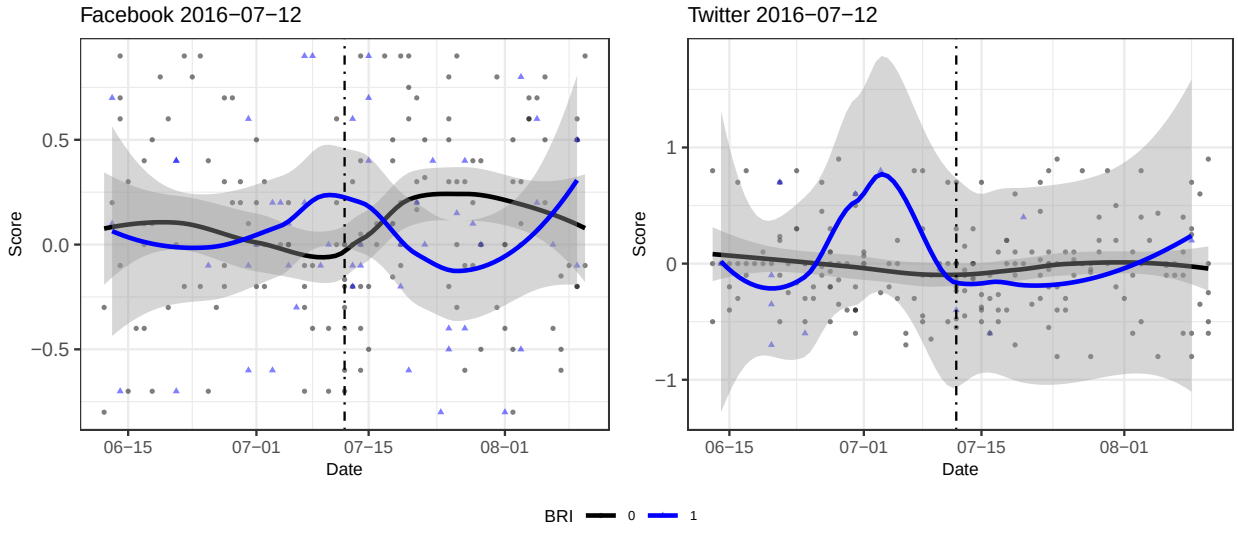
The assumption behind our balanced panel construction is that when a constituency does not have any posts about China, the politicians in that constituency express a neutral sentiment about China. However, the non-existence of postings about China might indicate highly negative or positive opinions that politicians do not feel comfortable expressing on their social media platforms, or indifference about China – a slightly different attitude from neutral sentiment. By dropping constituencies that do not have posts about China, we avoid making assumptions about politician attitudes when there are no posts about China. Nevertheless, the unbalanced panel data structure would distort the representation of the constituencies. Therefore, we use the balanced panel dataset for all regression analysis, and the unbalanced datasets are reserved only for descriptive analysis.

Figure A-19: Elite Sentiment Toward China Grouped by Project Treatment Status: Unbalanced Panel

(a) The Philippines

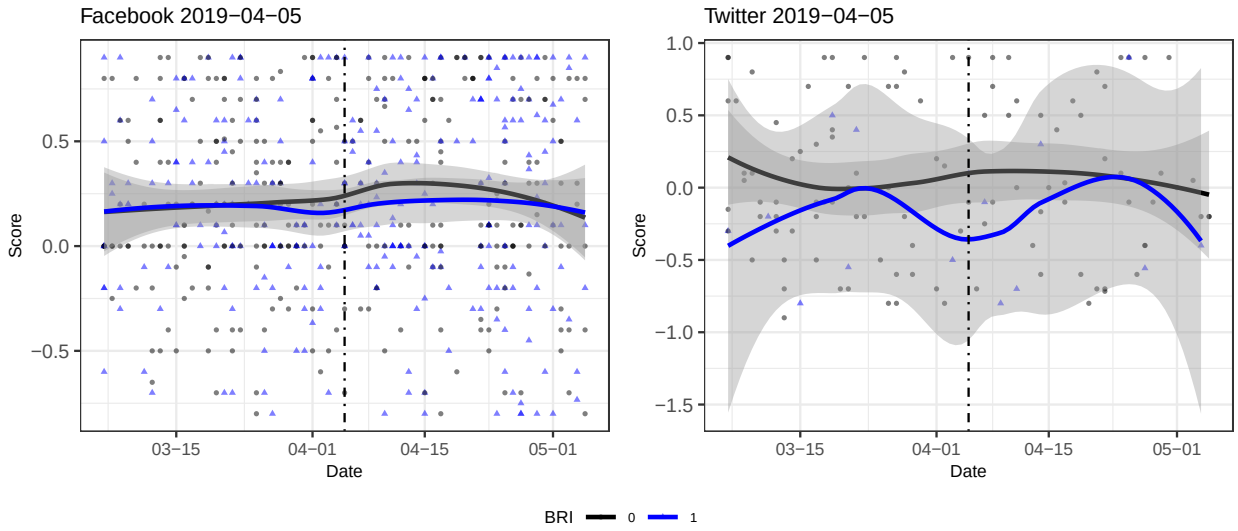


Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped. 243 constituencies and 55 constituency−date level observations of elite sentiment on Facebook, and 199 observations on Twitter within the 61−day event window in total.

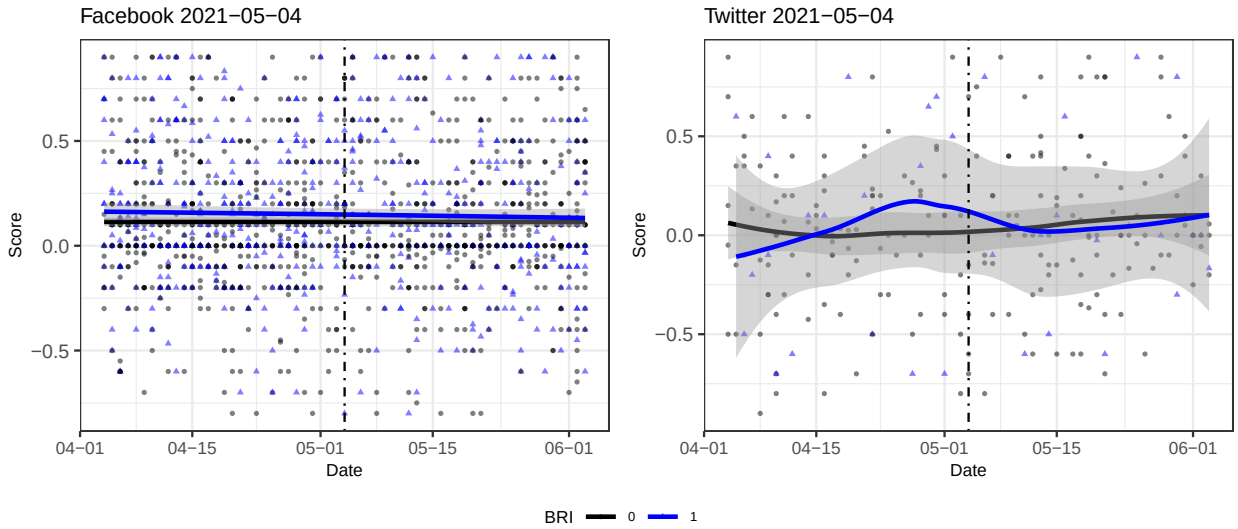


Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped. 243 constituencies and 215 constituency−date level observations of elite sentiment on Facebook, and 218 observations on Twitter within the 61−day event window in total.

Figure A-19: Elite Sentiment Toward China Grouped by Project Treatment Status: Unbalanced Panel (continued)



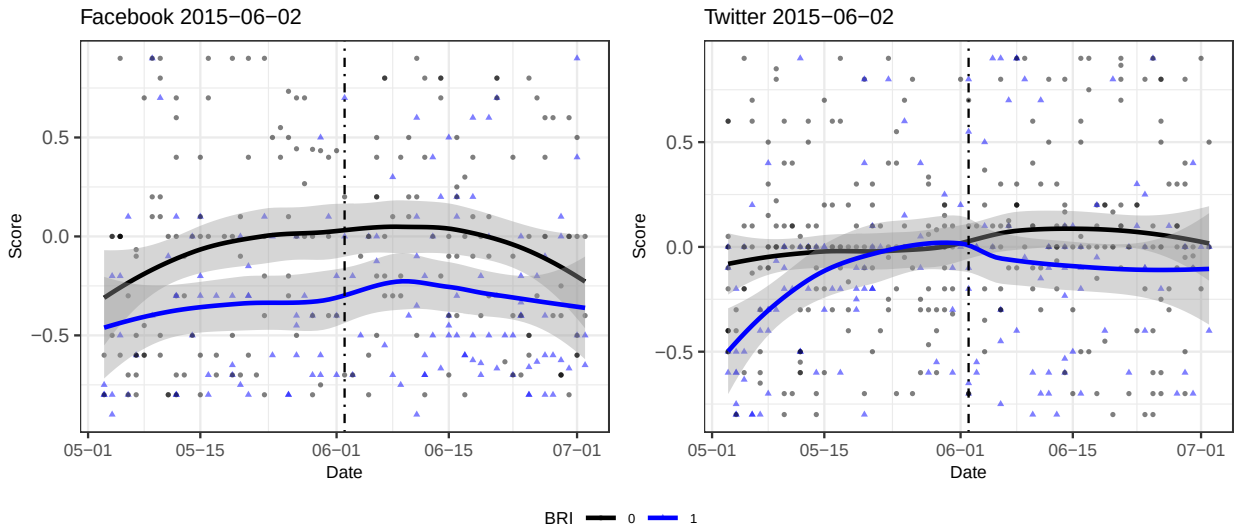
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
243 constituencies and 675 constituency−date level observations of elite sentiment on Facebook, and 127 observations on Twitter within the 61−day event window in total.



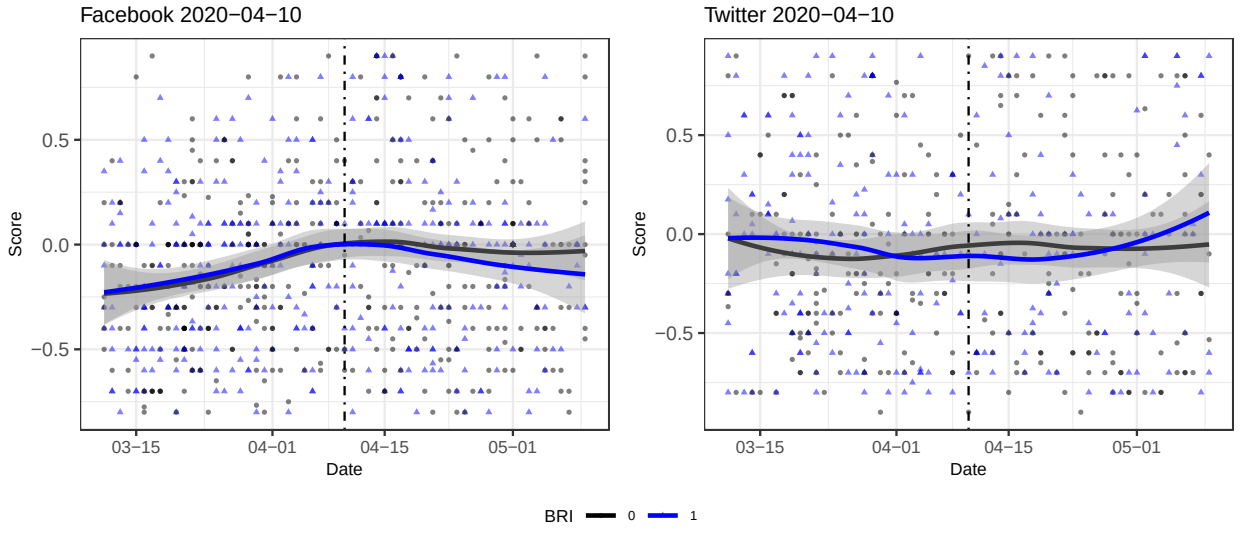
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
243 constituencies and 1664 constituency−date level observations of elite sentiment on Facebook, and 250 observations on Twitter within the 61−day event window in total.

Figure A-19: Elite Sentiment Toward China Grouped by Project Treatment Status: Unbalanced Panel (continued)

(b) Malaysia

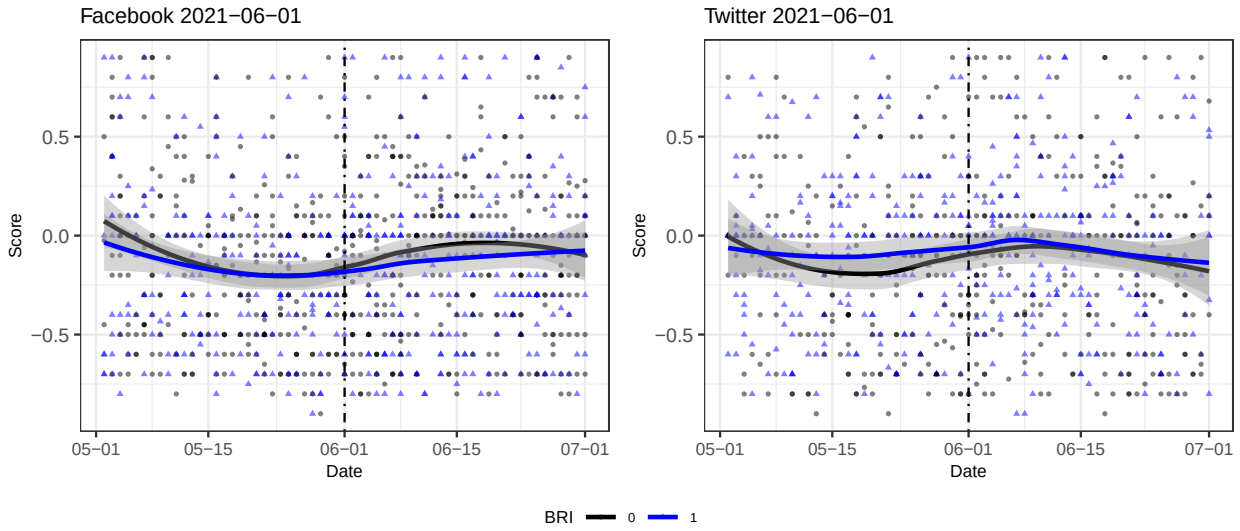


Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
222 constituencies and 376 constituency−date level observations of elite sentiment on Facebook, and 465 observations on Twitter within the 61−day event window in total.



Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
222 constituencies and 915 constituency−date level observations of elite sentiment on Facebook, and 537 observations on Twitter within the 61−day event window in total.

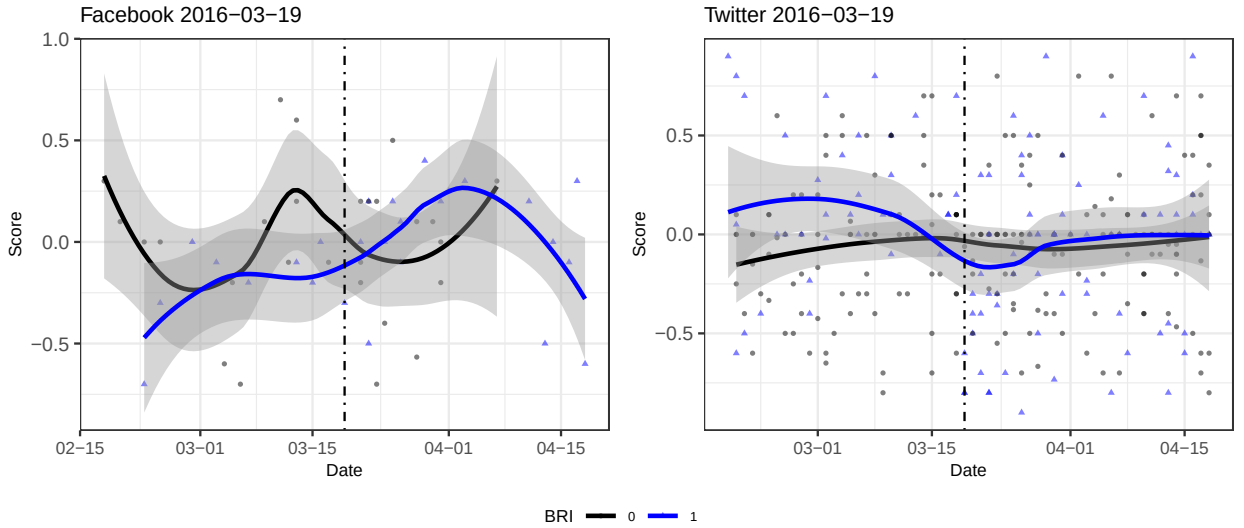
Figure A-19: Elite Sentiment Toward China Grouped by Project Treatment Status: Unbalanced Panel (continued)



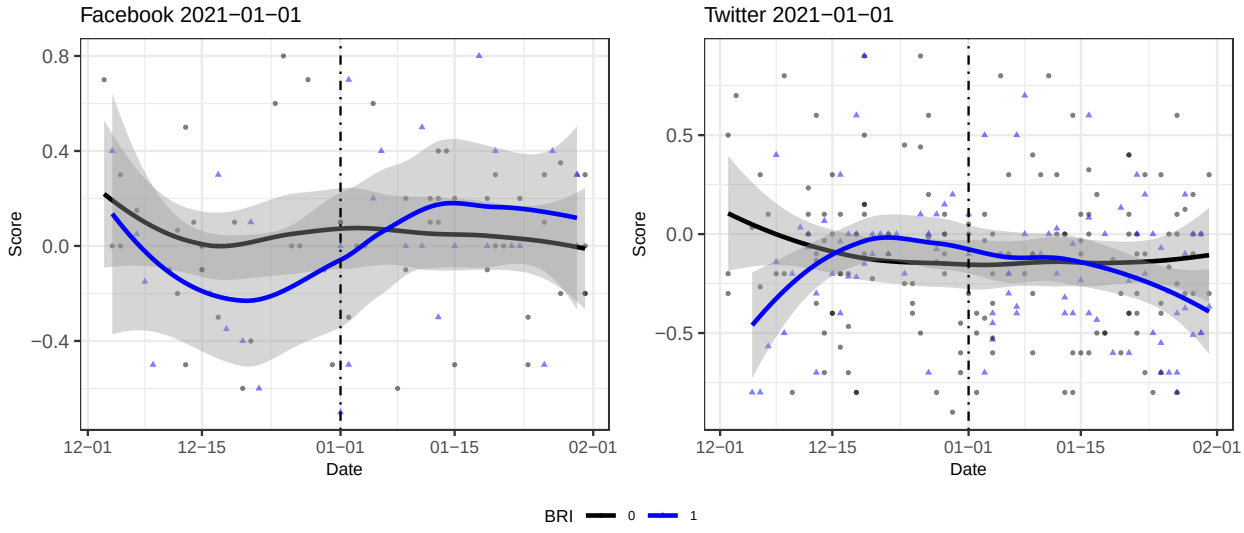
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped. 222 constituencies and 1201 constituency−date level observations of elite sentiment on Facebook, and 831 observations on Twitter within the 61−day event window in total.

Figure A-19: Elite Sentiment Toward China Grouped by Project Treatment Status: Unbalanced Panel (continued)

(c) Indonesia



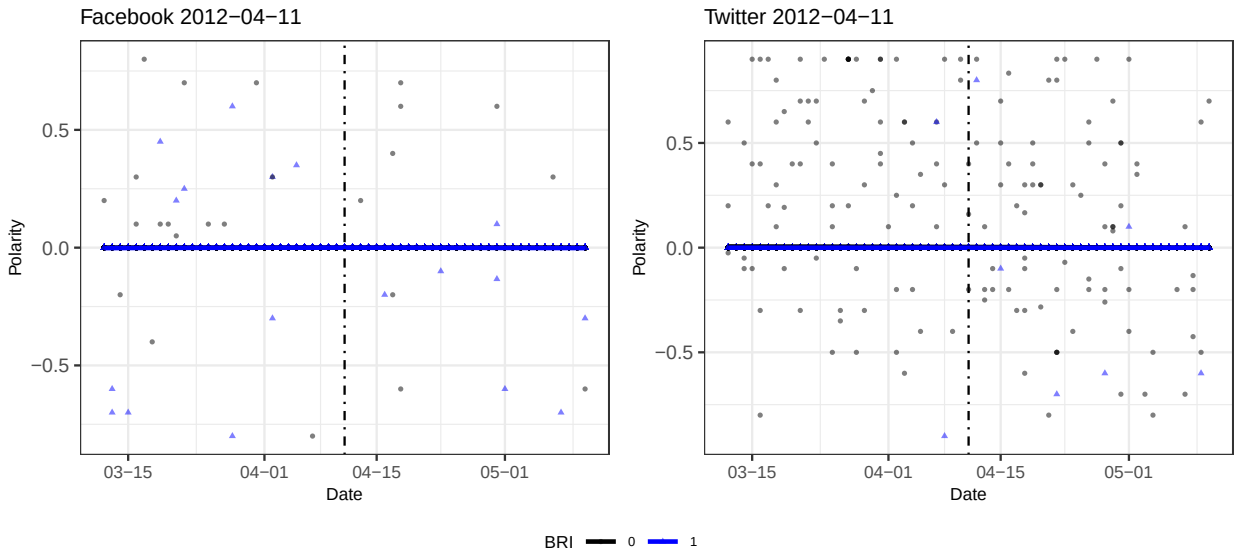
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
77 constituencies and 53 constituency−date level observations of elite sentiment on Facebook, and 303 observations on Twitter within the 61−day event window in total.



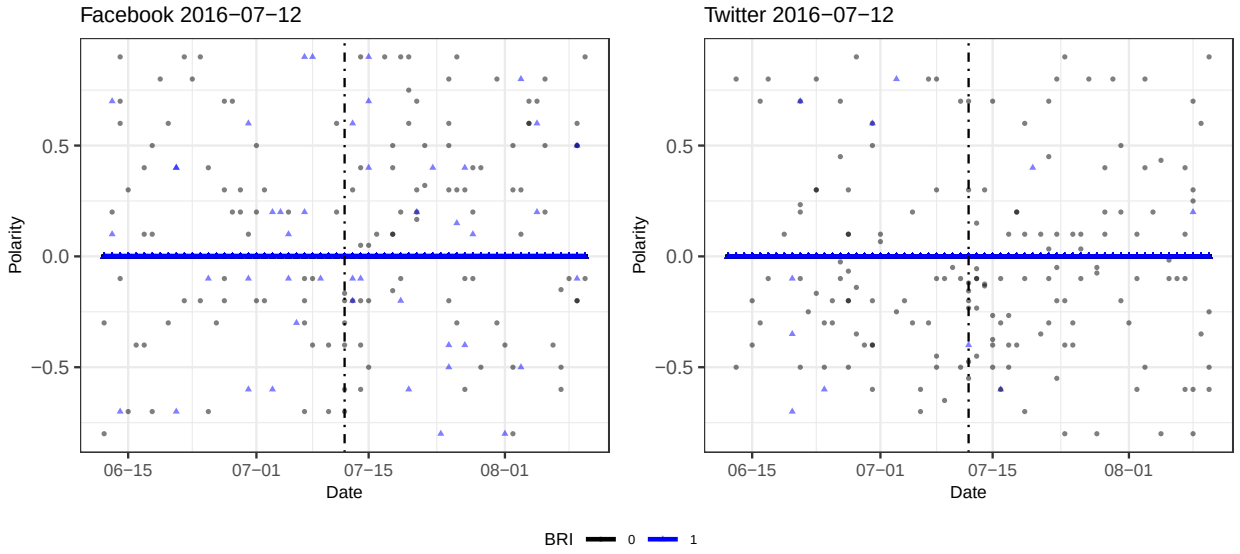
Note: Graphs show the unbalanced panel. Constituency−date pairs without any social media posting are dropped.
77 constituencies and 91 constituency−date level observations of elite sentiment on Facebook, and 250 observations on Twitter within the 61−day event window in total.

Figure A-20: Elite Sentiment Toward China Grouped by Project Treatment Status: Balanced Panel

(a) The Philippines

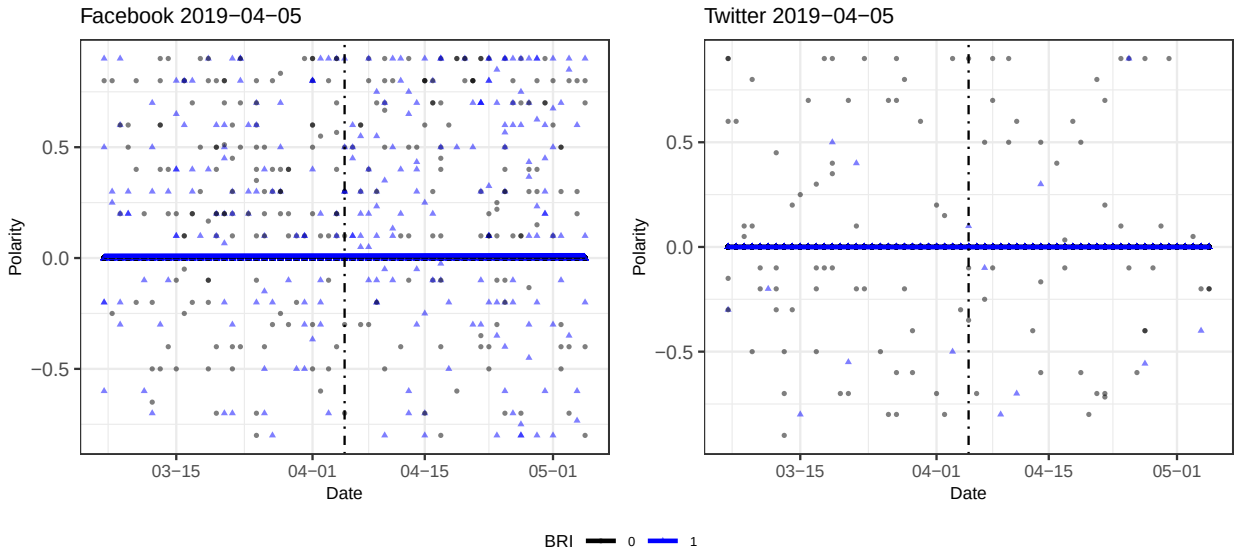


Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0. 243 constituencies and 14832 constituency−date level observation of elite sentiment within the 61−day event window in total.

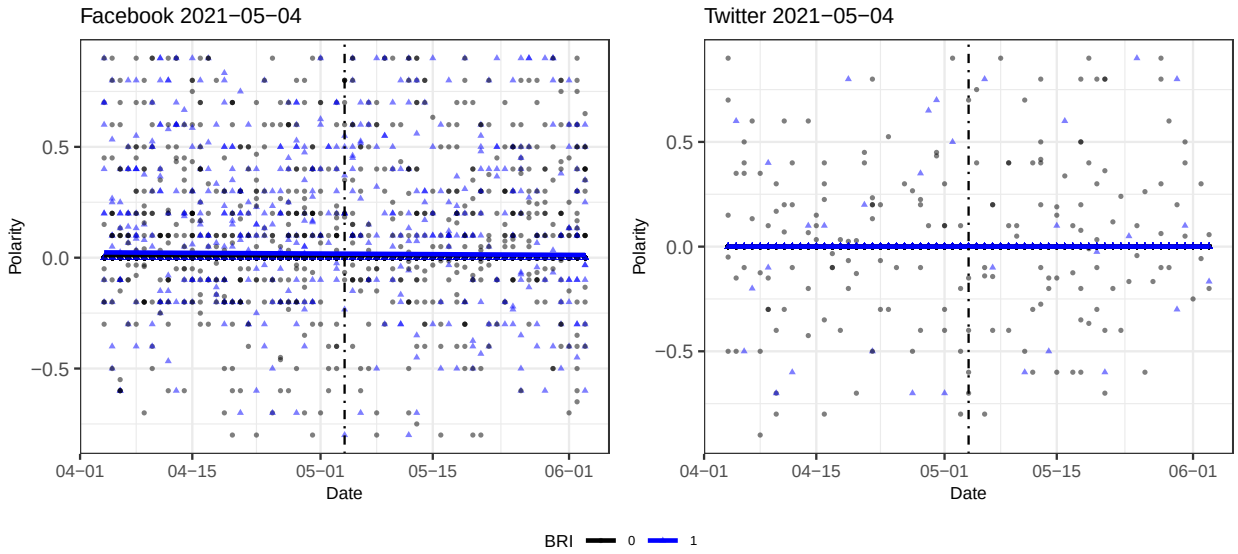


Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0. 243 constituencies and 14832 constituency−date level observation of elite sentiment within the 61−day event window in total.

Figure A-20: Elite Sentiment Toward China Grouped by Project Treatment Status: Balanced Panel (continued)



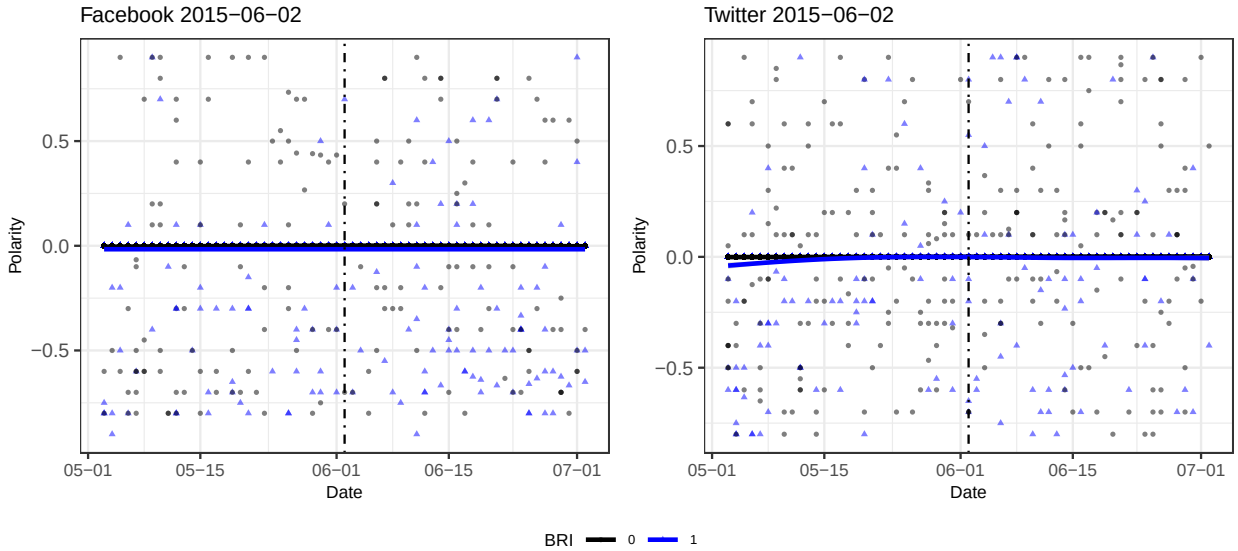
Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0. 243 constituencies and 14832 constituency−date level observation of elite sentiment within the 61−day event window in total.



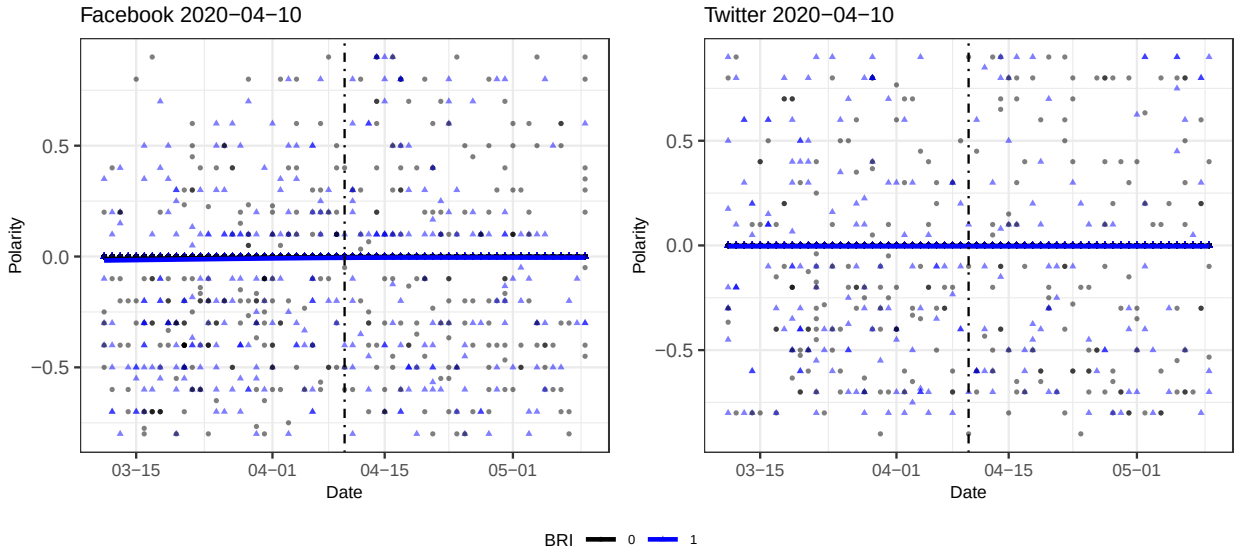
Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0. 243 constituencies and 14832 constituency−date level observation of elite sentiment within the 61−day event window in total.

Figure A-20: Elite Sentiment Toward China Grouped by Project Treatment Status: Balanced Panel (continued)

(b) Malaysia

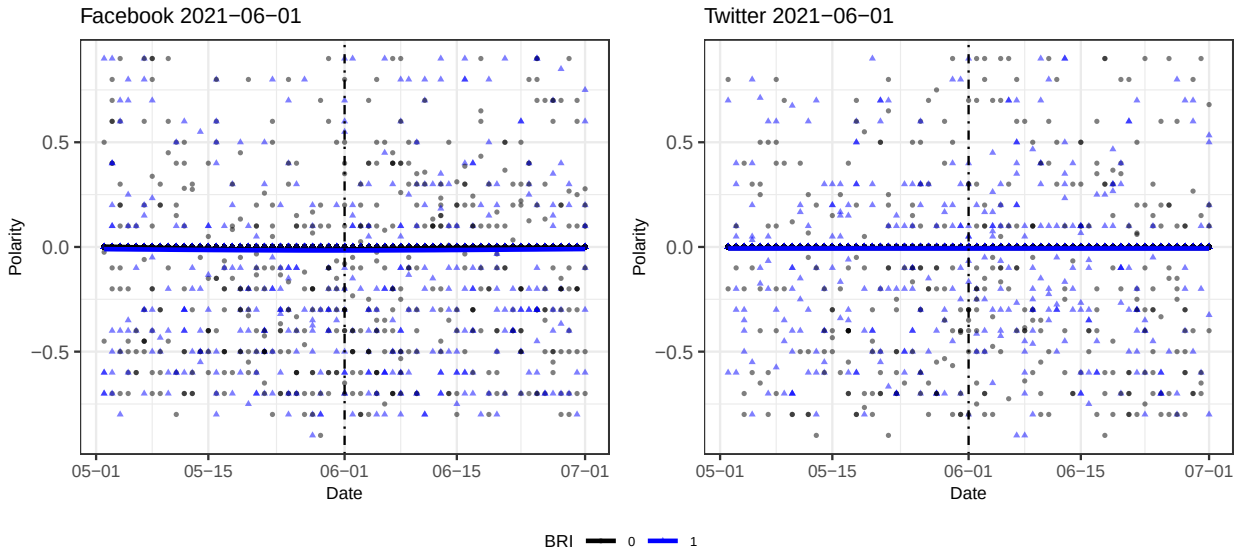


Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0. 222 constituencies and 13542 constituency−date level observation of elite sentiment within the 61−day event window in total.



Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0. 222 constituencies and 13542 constituency−date level observation of elite sentiment within the 61−day event window in total.

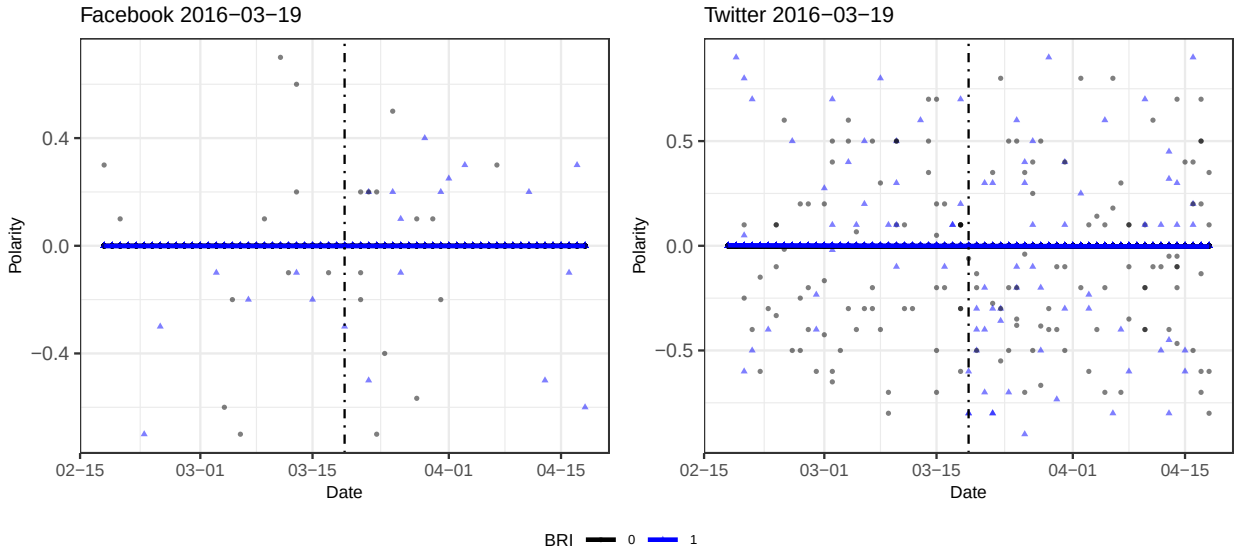
Figure A-20: Elite Sentiment Toward China Grouped by Project Treatment Status: Balanced Panel (continued)



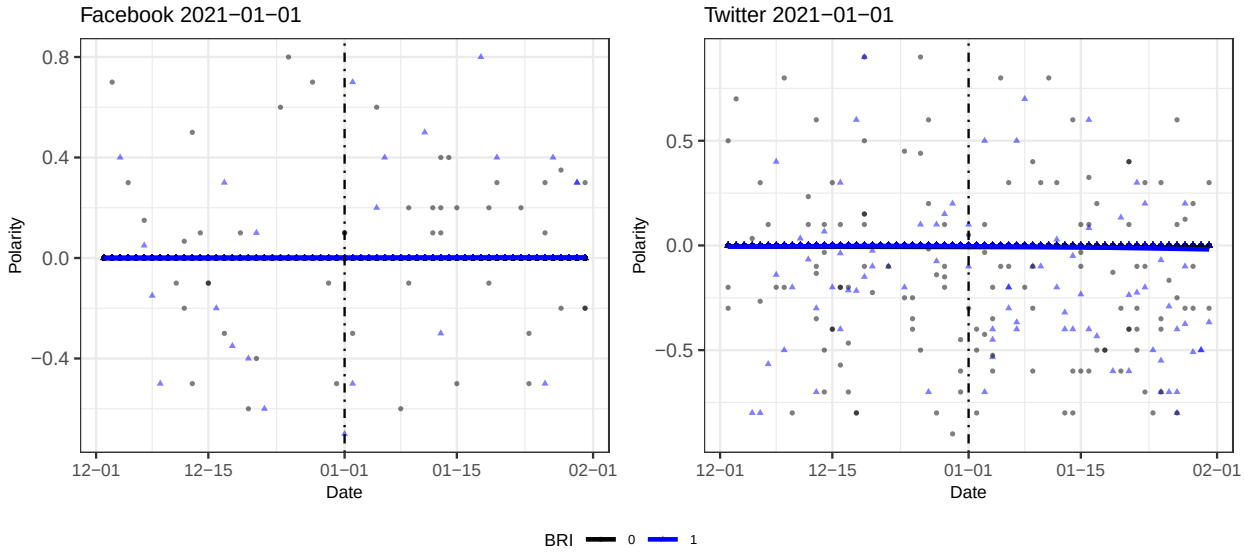
Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0.
222 constituencies and 13542 constituency−date level observation of elite sentiment within the 61−day event window in total.

Figure A-20: Elite Sentiment Toward China Grouped by Project Treatment Status: Balanced Panel (continued)

(c) Indonesia



Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0.
77 constituencies and 4697 constituency−date level observation of elite sentiment within the 61−day event window in total.



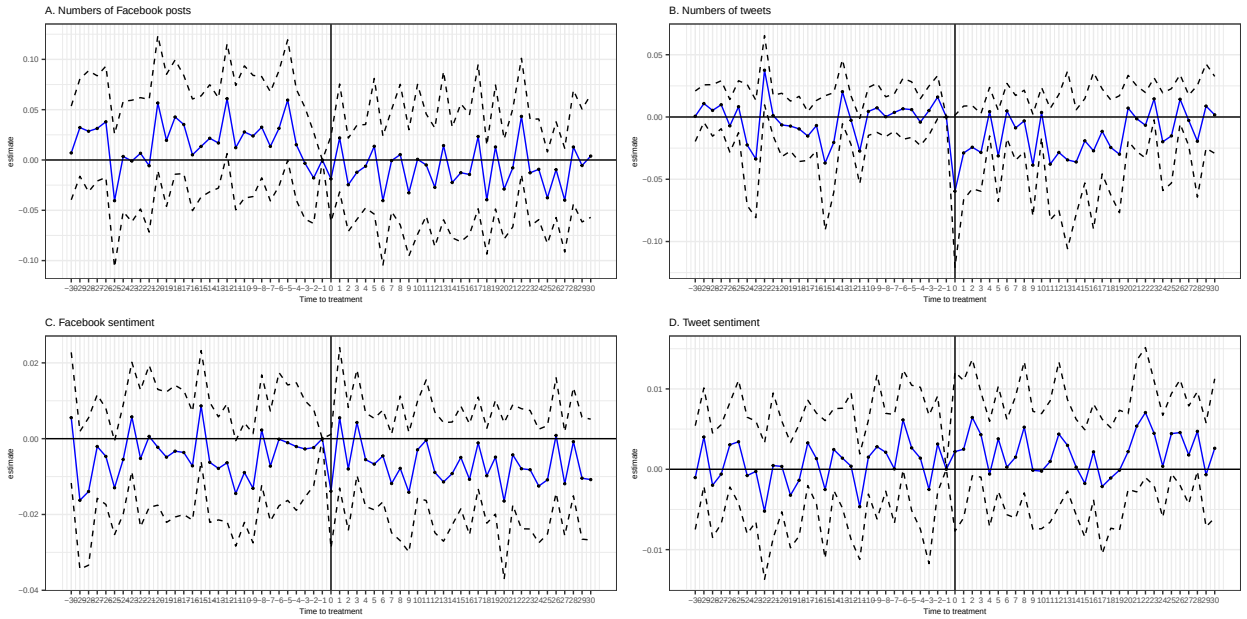
Note: Graphs show the balanced panel. For constituency−date pairs without any social media posting, the sentiment scores will be set to 0.
77 constituencies and 4697 constituency−date level observation of elite sentiment within the 61−day event window in total.

3 Regression Results

3.1 Google NLP API Sentiment

Figure A-21: Stacked Event Study on Elite Social Media Posts about China: constituency-date level

(a) The Philippines



(b) Malaysia

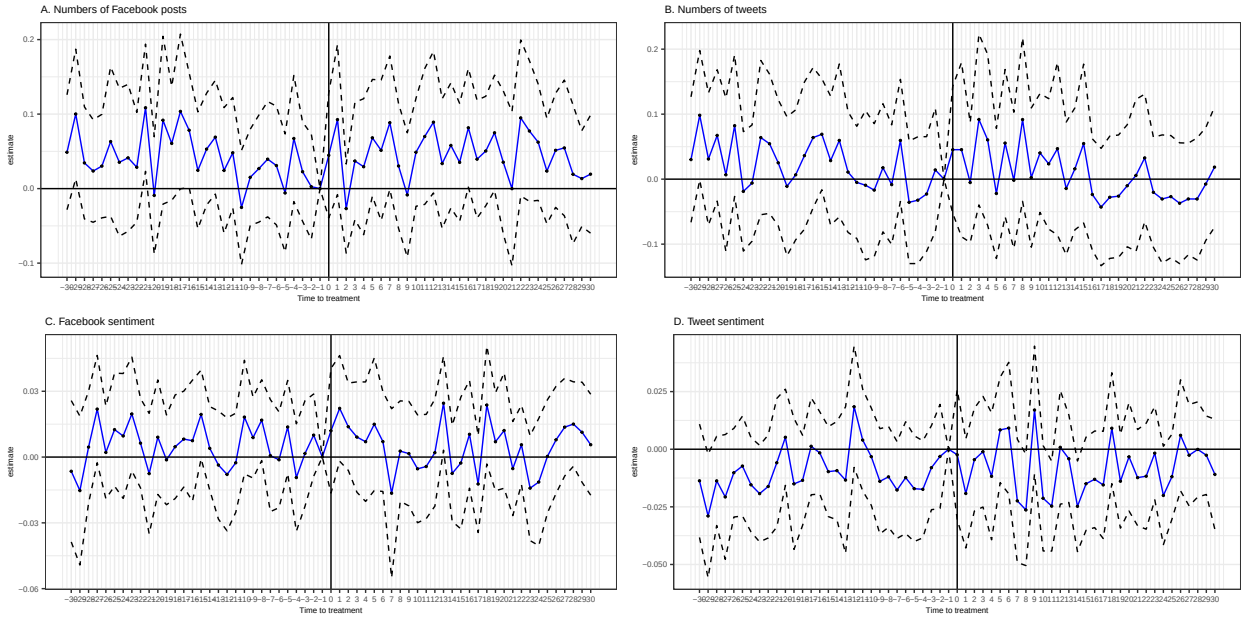
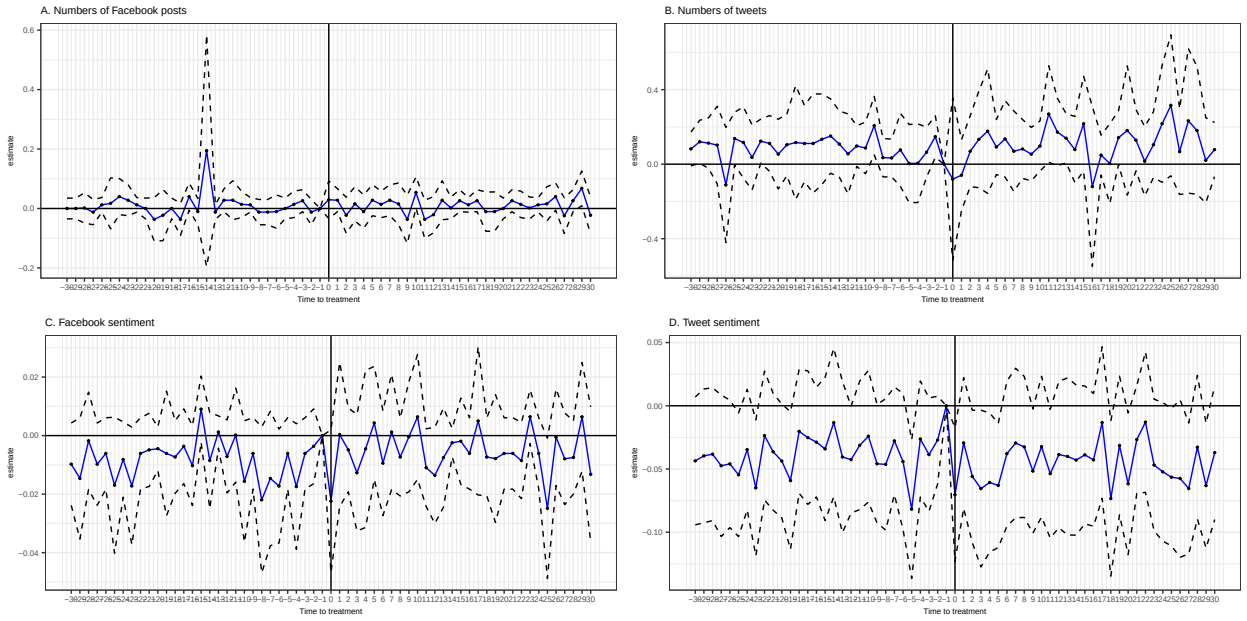


Figure A-21: Stacked Event Study on Elite Social Media Posts about China: constituency-date level (continued)

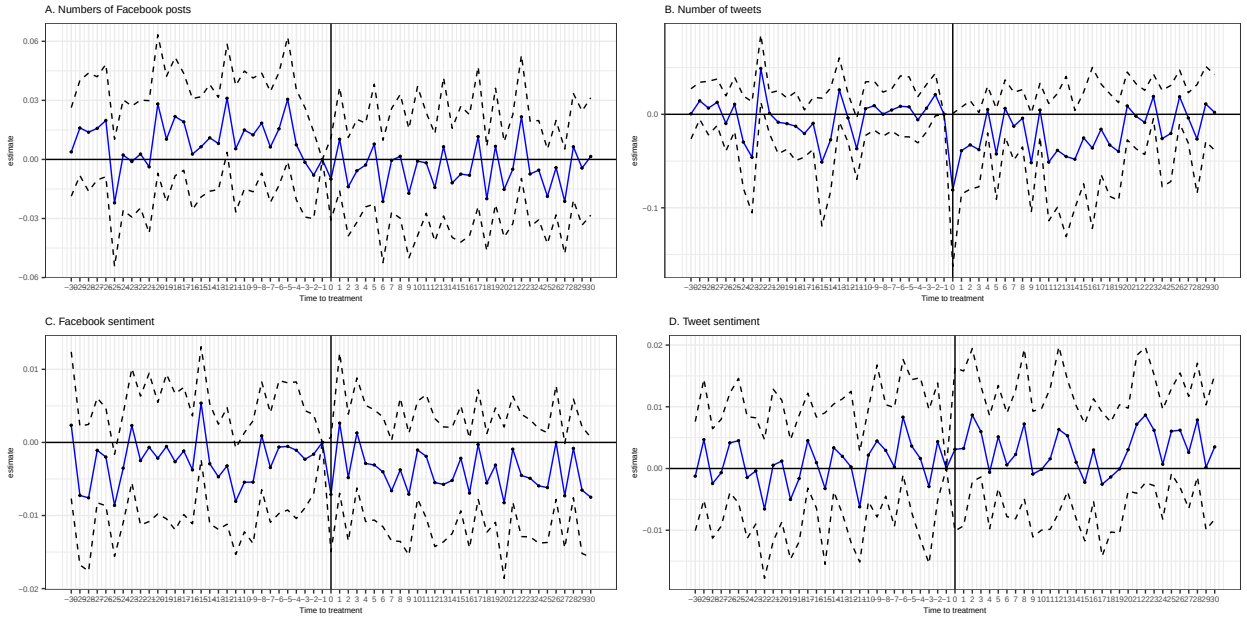
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-22: Stacked Event Study on Elite Social Media Posts about China: politician-date level

(a) The Philippines



(b) Malaysia

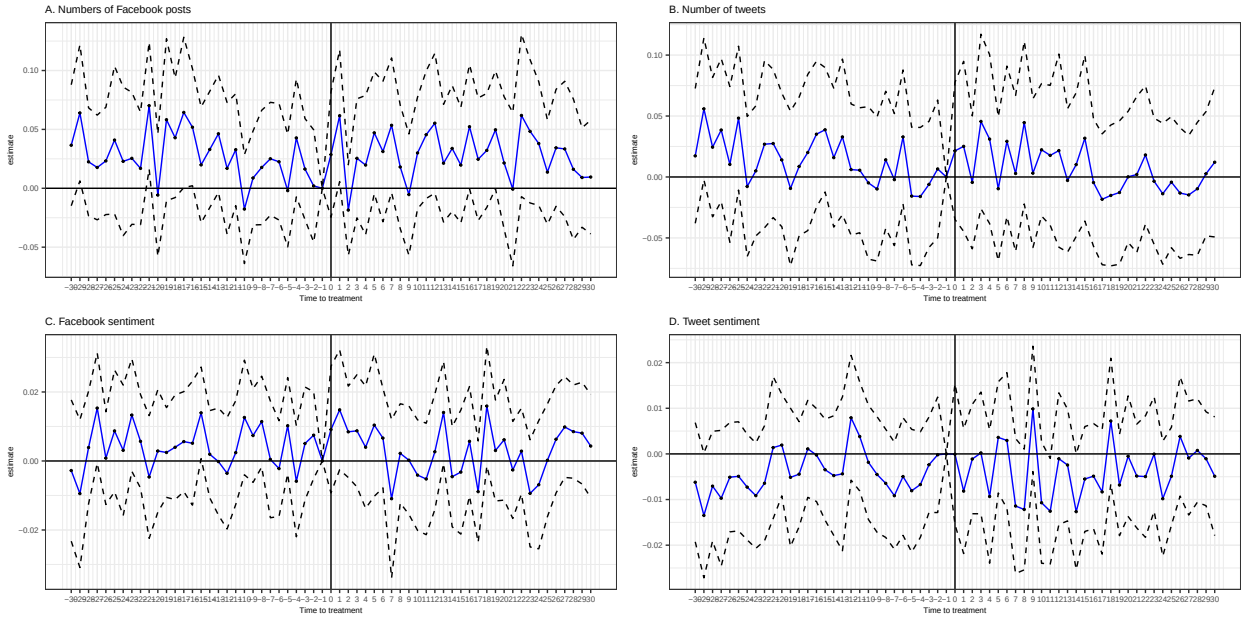
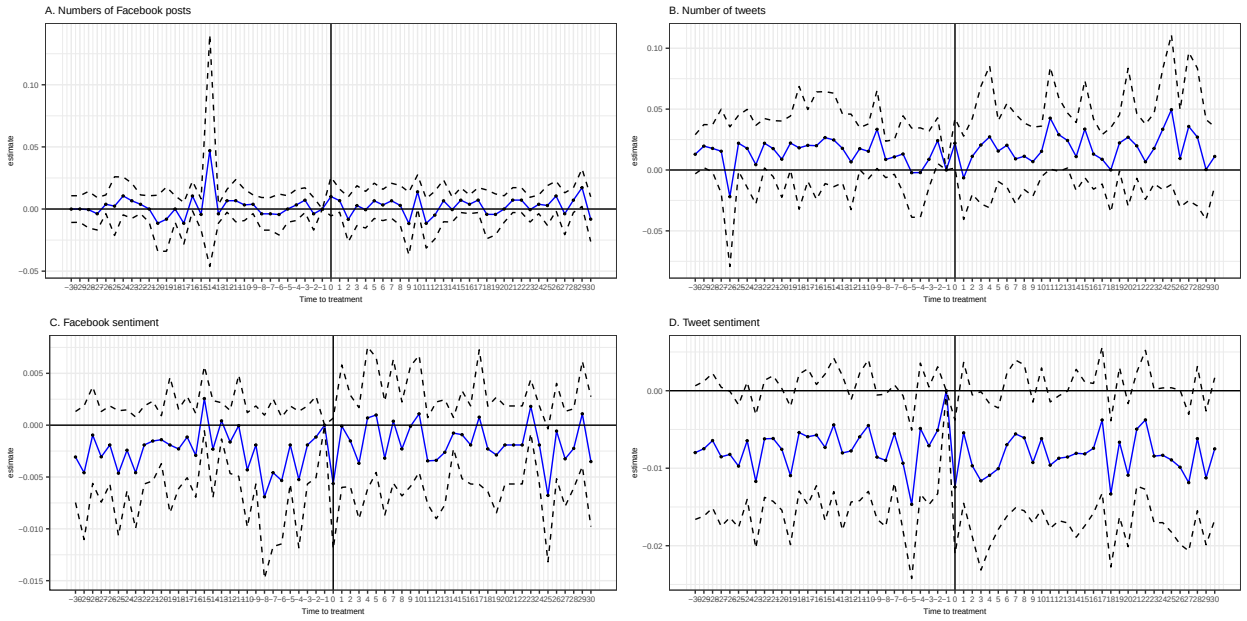


Figure A-22: Stacked Event Study on Elite Social Media Posts about China: politician-date level (continued)

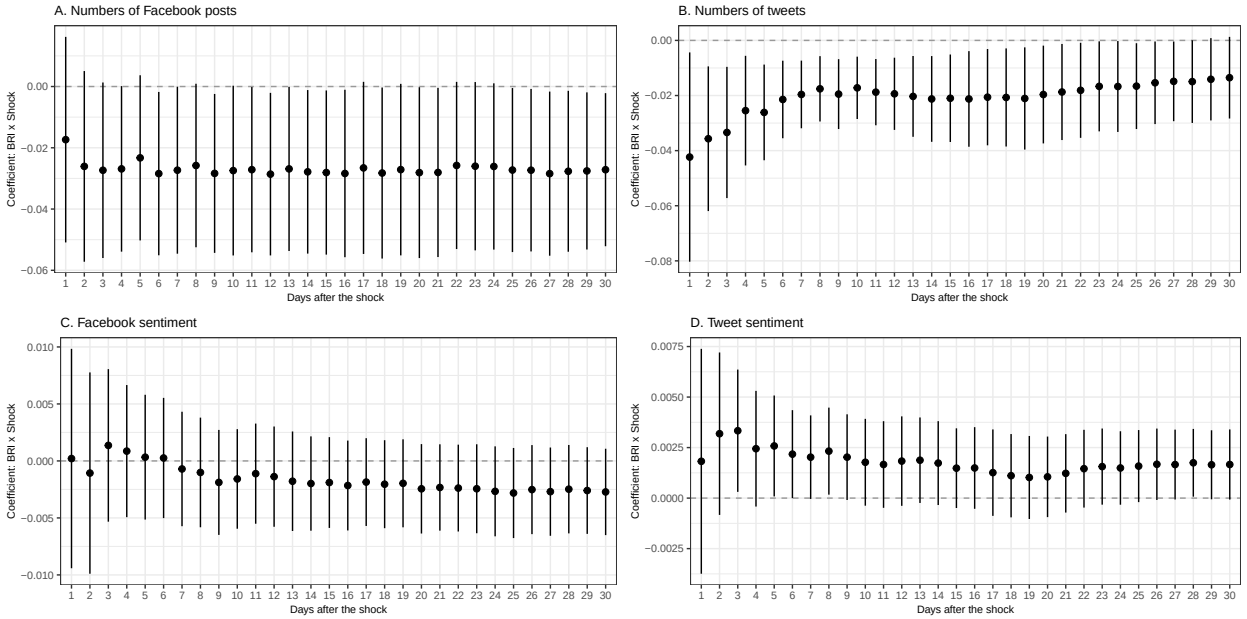
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C, -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

Figure A-23: Stacked DID on Elite Social Media Posts about China: Constituency-date level

(a) The Philippines



(b) Malaysia

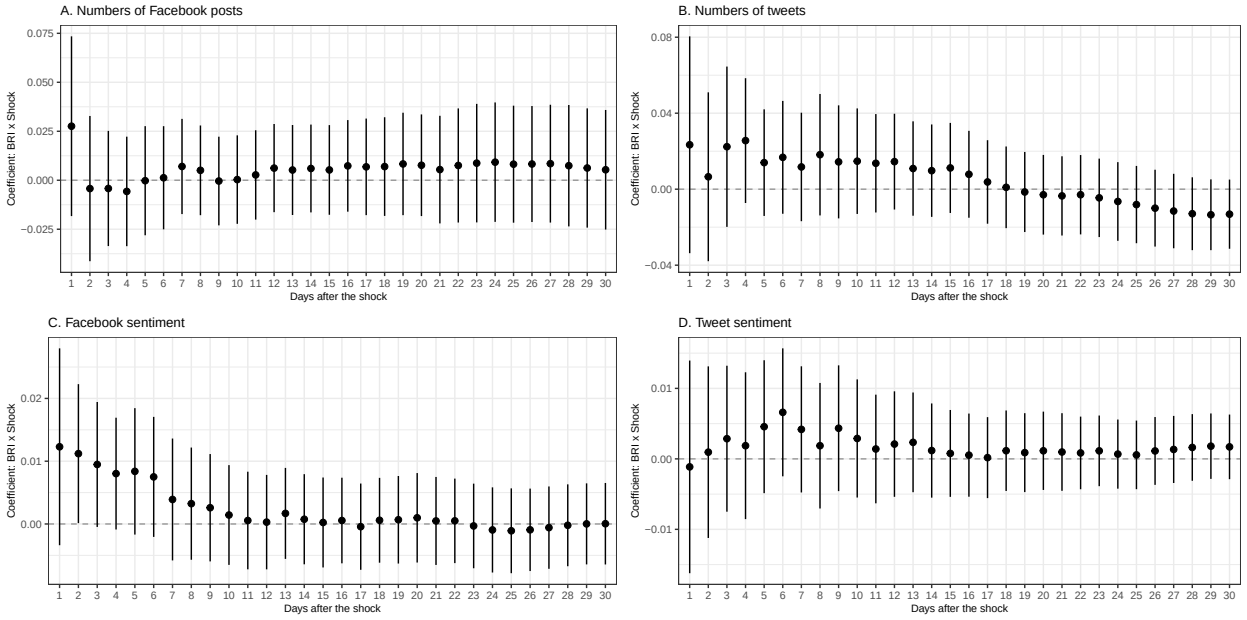
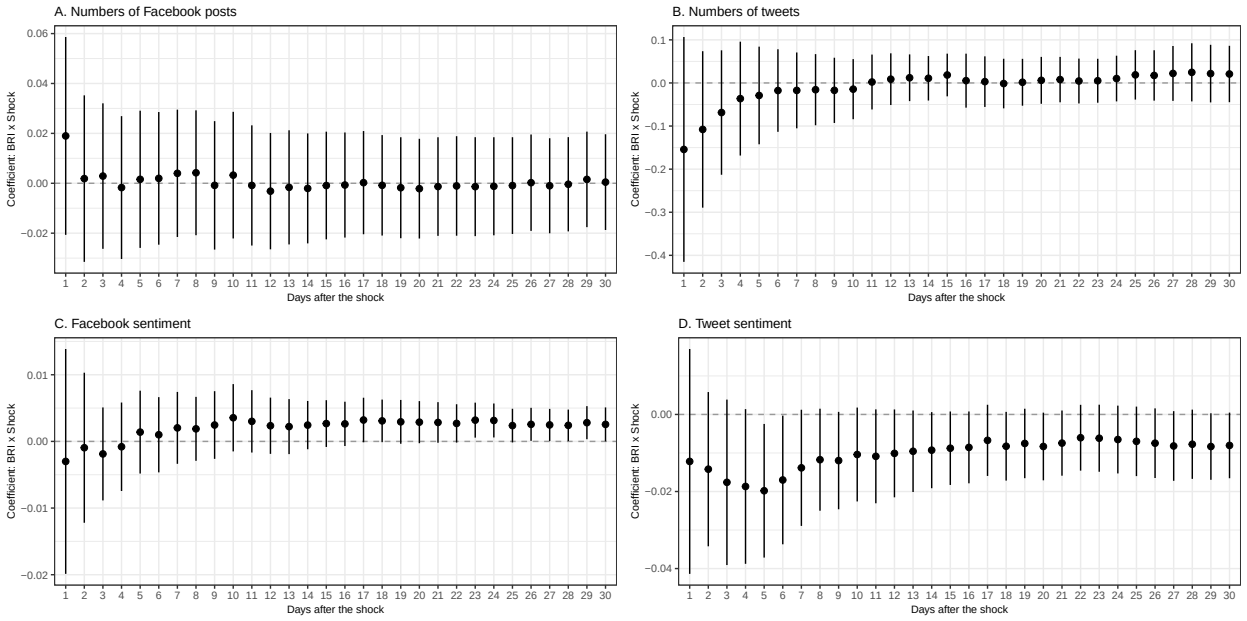


Figure A-23: Stacked DID on Elite Social Media Posts about China: Constituency-date level (continued)

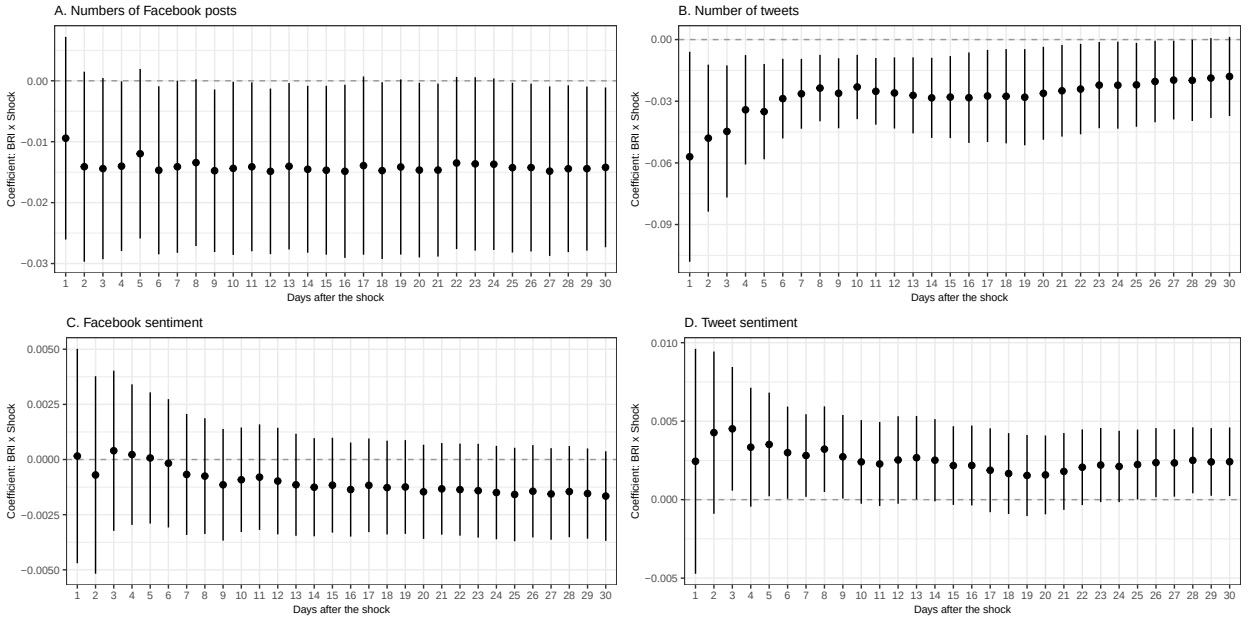
(c) Indonesia



Note: Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-24: Stacked DID on Elite Social Media Posts about China: Politician-date level

(a) The Philippines



(b) Malaysia

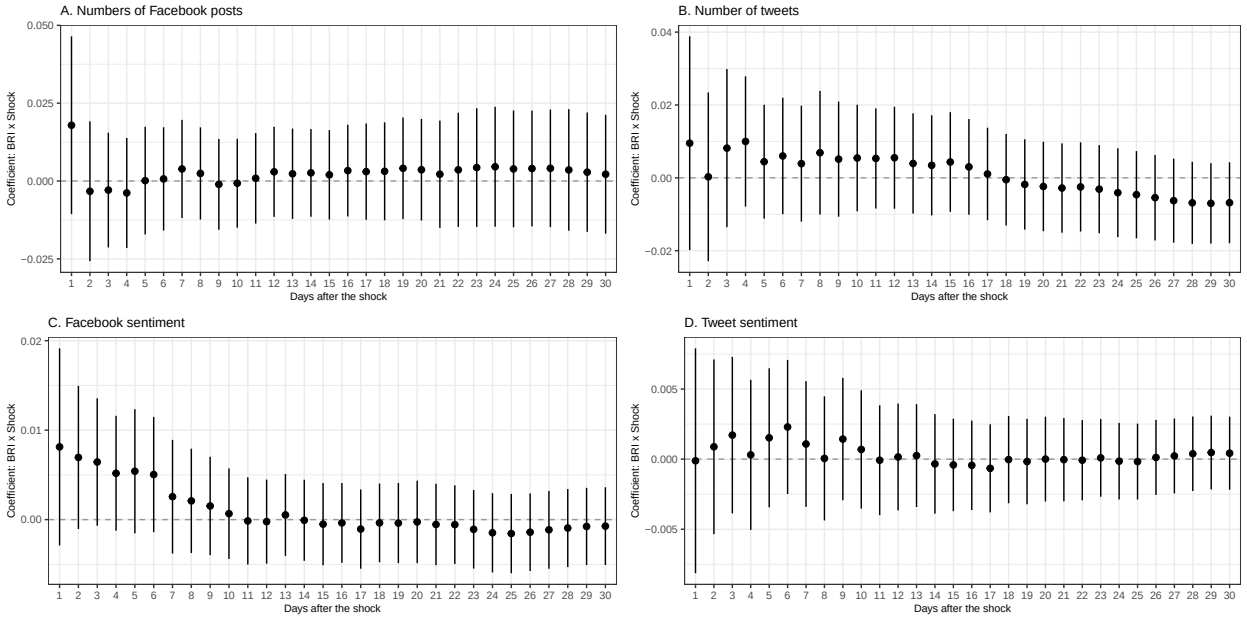
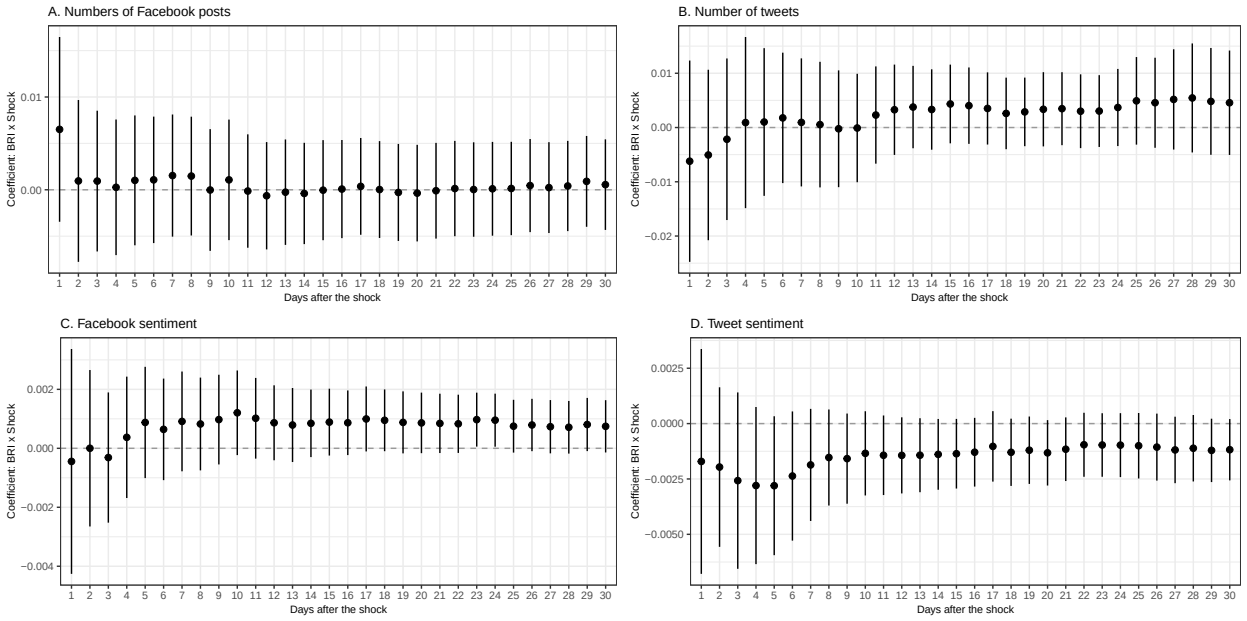


Figure A-24: Stacked DID on Elite Social Media Posts about China: Politician-date level (continued)

(c) Indonesia

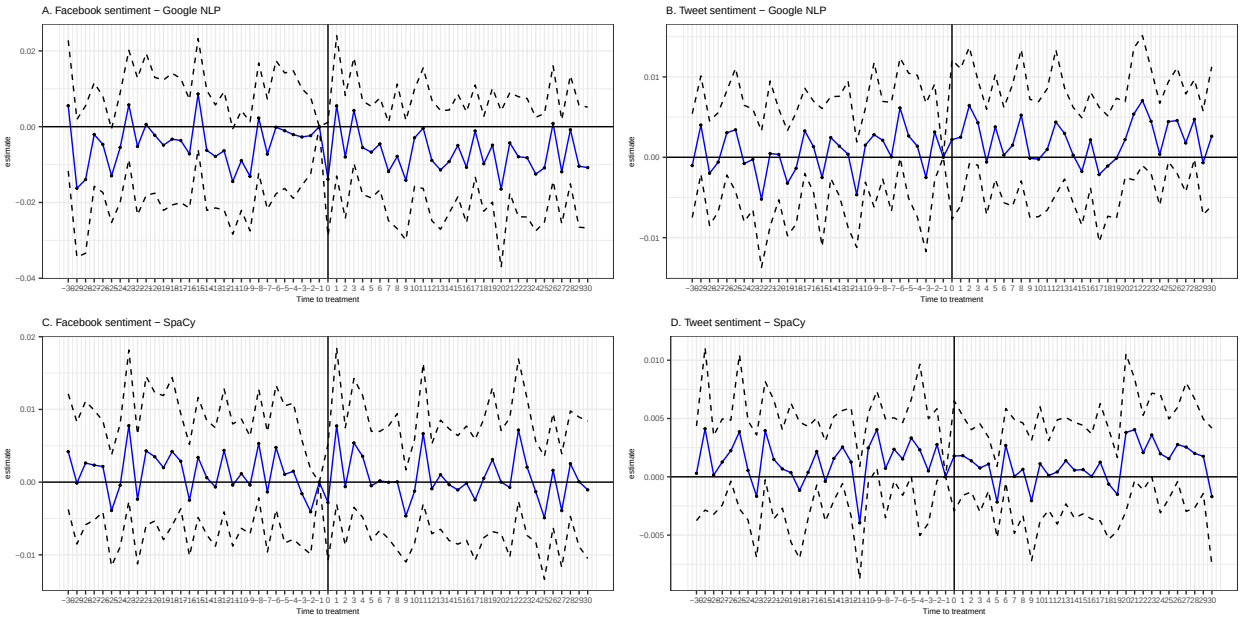


Note: Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts, B. number of tweets, C. sentiment of Facebook posts, and D. sentiment of tweets. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the politician level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

3.2 SpaCy Sentiment

Figure A-25: Stacked Event Study on Elite Social Media Posts about China (SpaCy Sentiment): constituency-date level

(a) The Philippines



(b) Malaysia

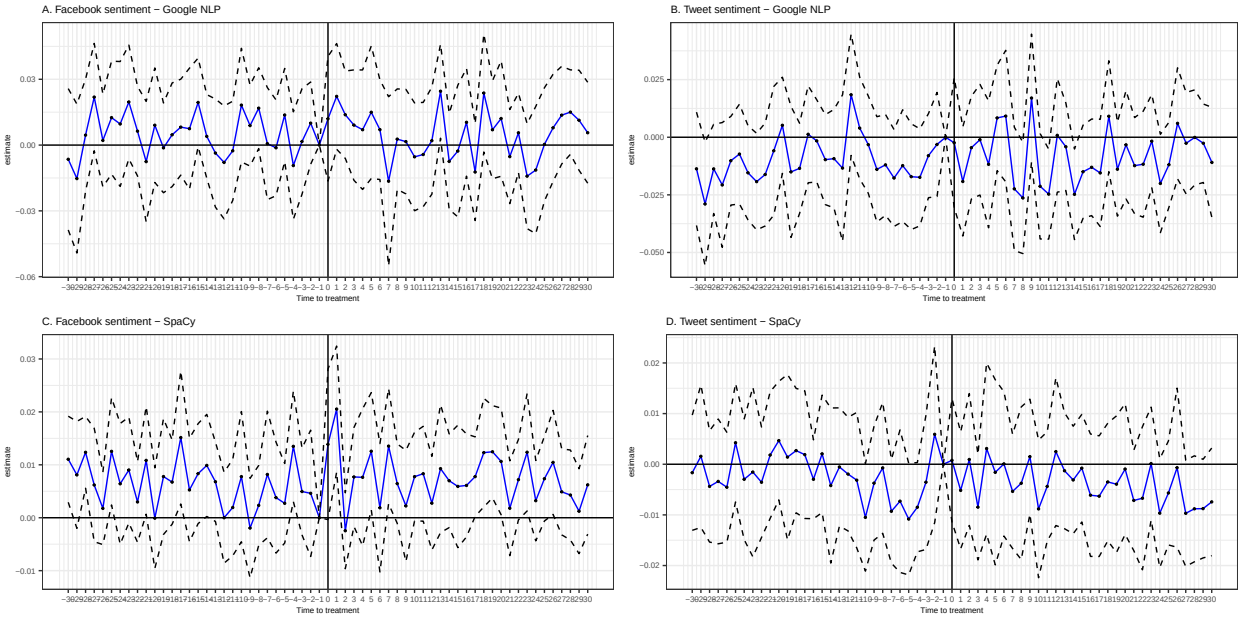
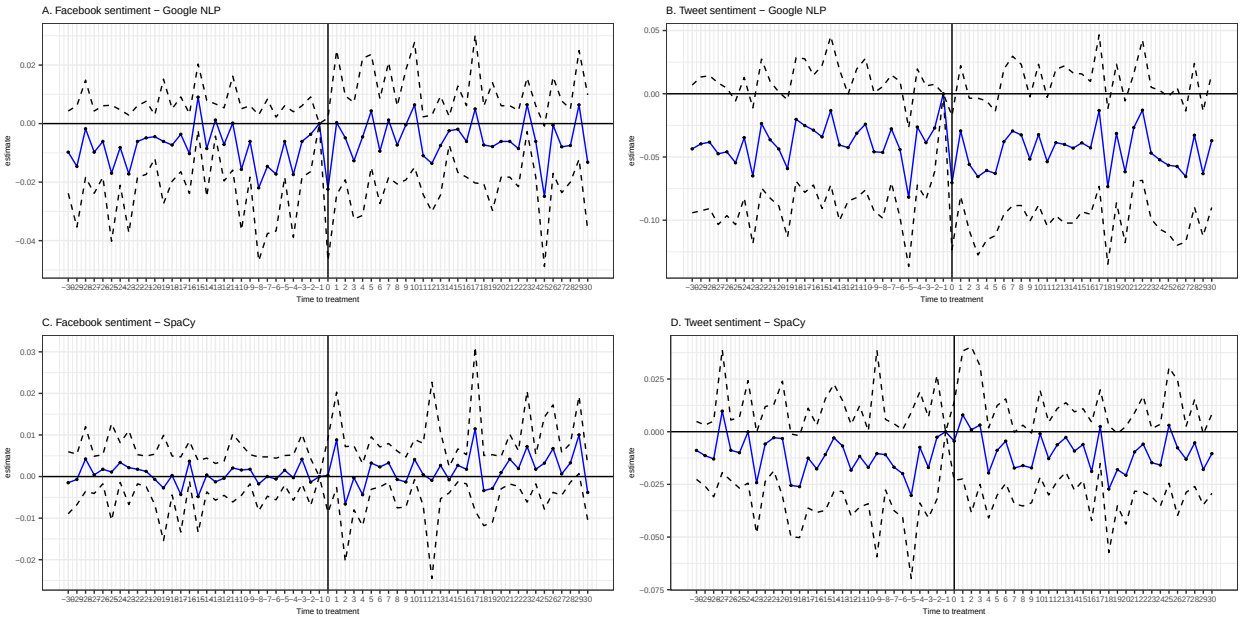


Figure A-25: Stacked Event Study on Elite Social Media Posts about China (SpaCy Sentiment): constituency-date level (continued)

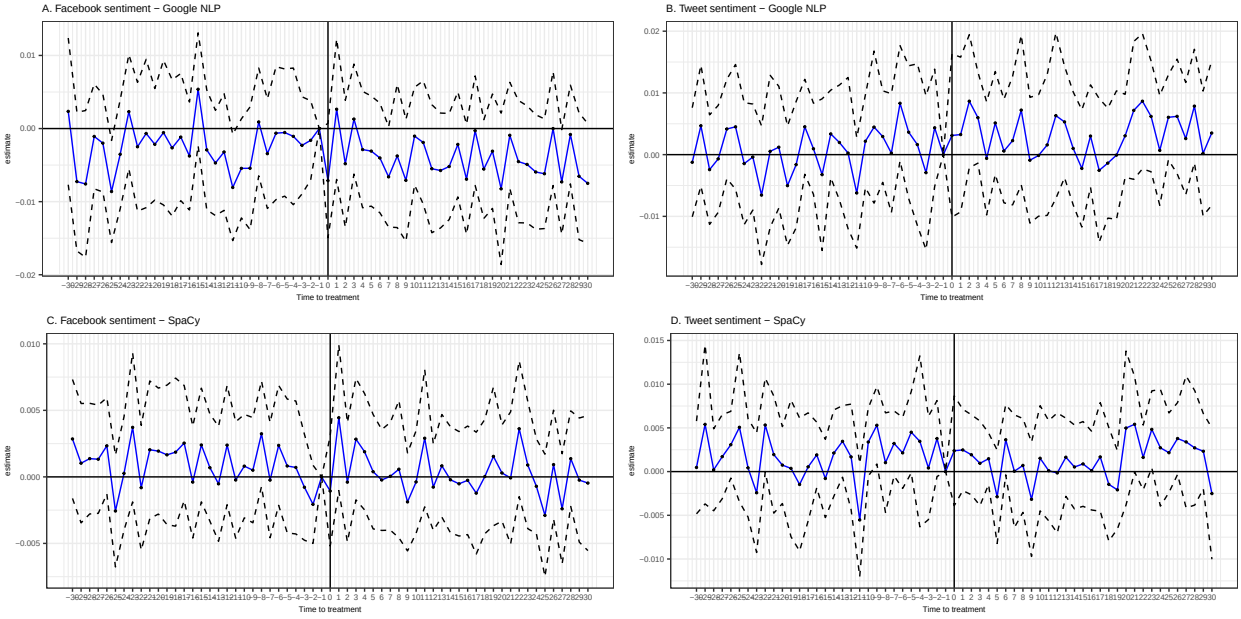
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. To facilitate the comparison of results obtained with Google NLP API, The dependent variable is A. Facebook sentiment toward China obtained with Google NLP API, B. tweet sentiment toward China obtained with Google NLP API, C. Facebook sentiment toward China using SpaCy Lexicon method, and D. tweet sentiment toward China obtained using SpaCy Lexicon method. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-26: Stacked Event Study on Elite Social Media Posts about China (SpaCy Sentiment): politician-date level

(a) The Philippines



(b) Malaysia

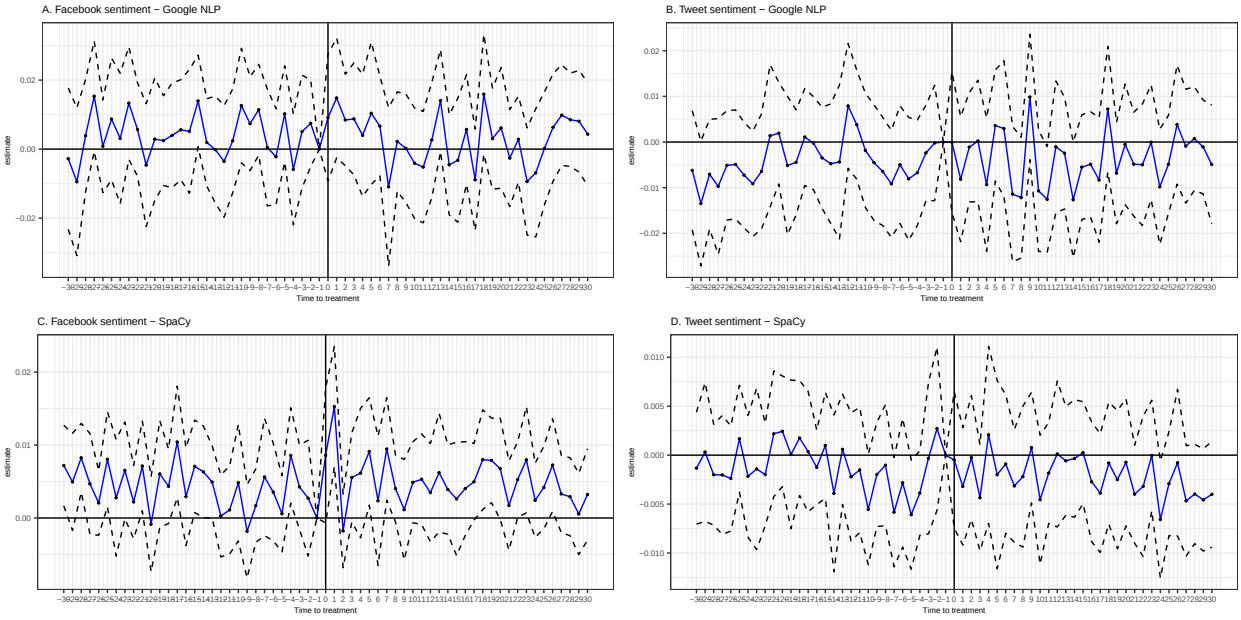
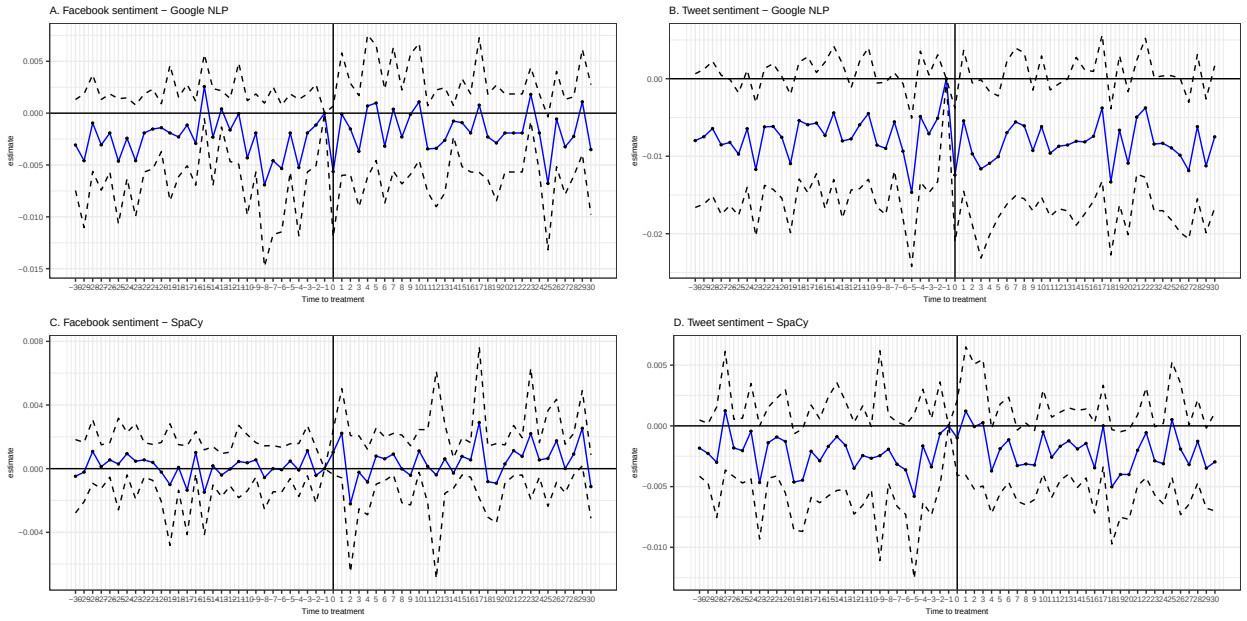


Figure A-26: Stacked Event Study on Elite Social Media Posts about China (SpaCy Sentiment): politician-date level (continued)

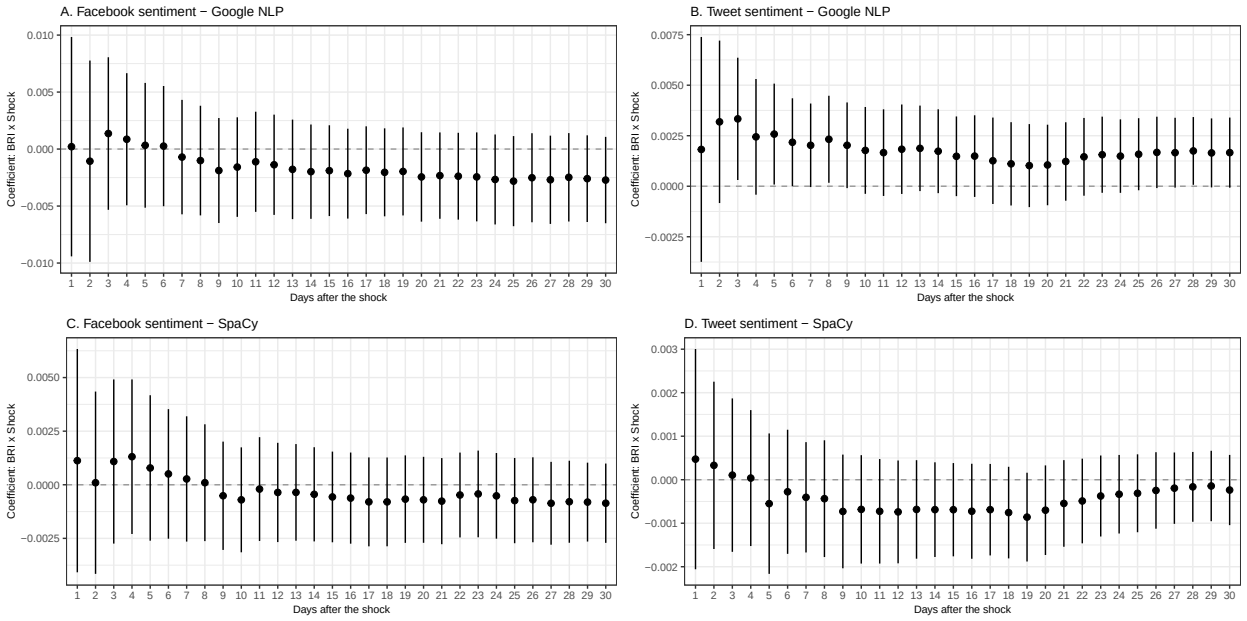
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. To facilitate the comparison of results obtained with Google NLP API, The dependent variable is A. Facebook sentiment toward China obtained with Google NLP API, B. tweet sentiment toward China obtained with Google NLP API, C. Facebook sentiment toward China using SpaCy Lexicon method, and D. tweet sentiment toward China obtained using SpaCy Lexicon method. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-27: Stacked DID on Elite Social Media Posts about China (SpaCy Sentiment): Constituency-date level

(a) The Philippines



(b) Malaysia

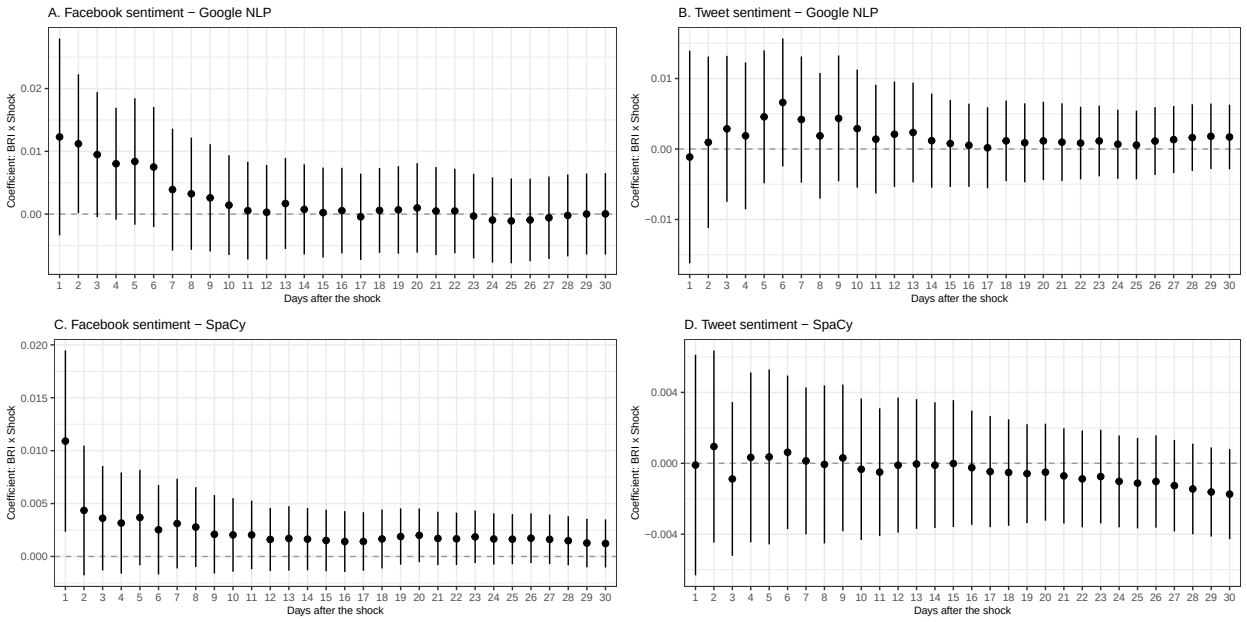
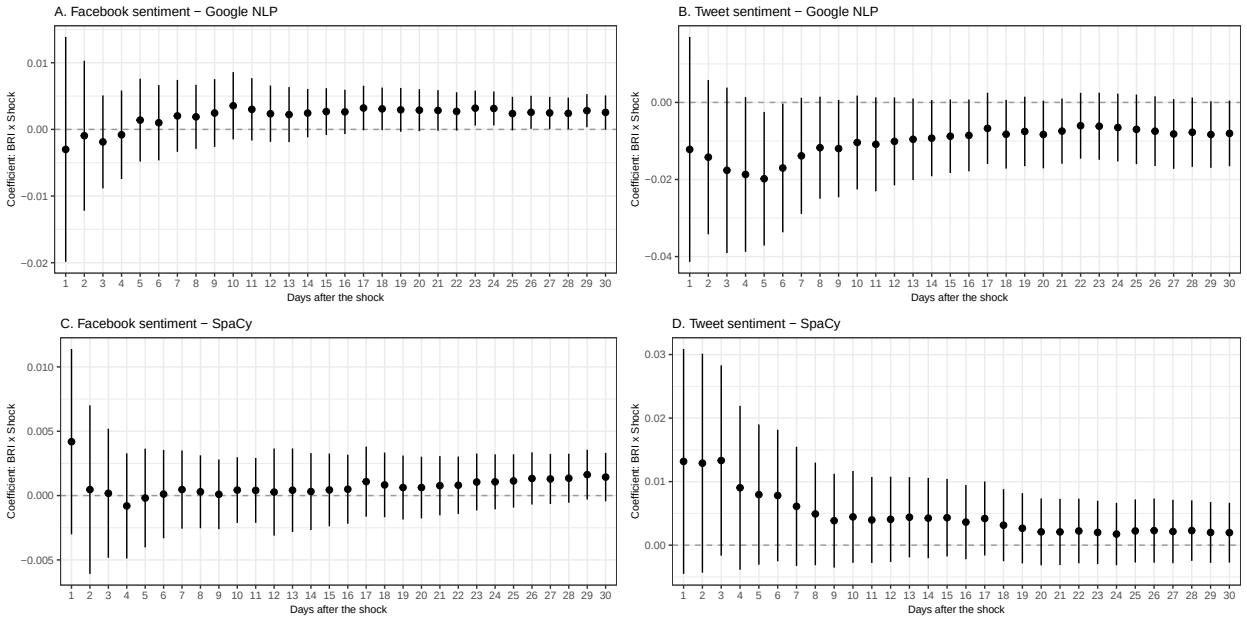


Figure A-27: Stacked DID on Elite Social Media Posts about China (SpaCy Sentiment): Constituency-date level (continued)

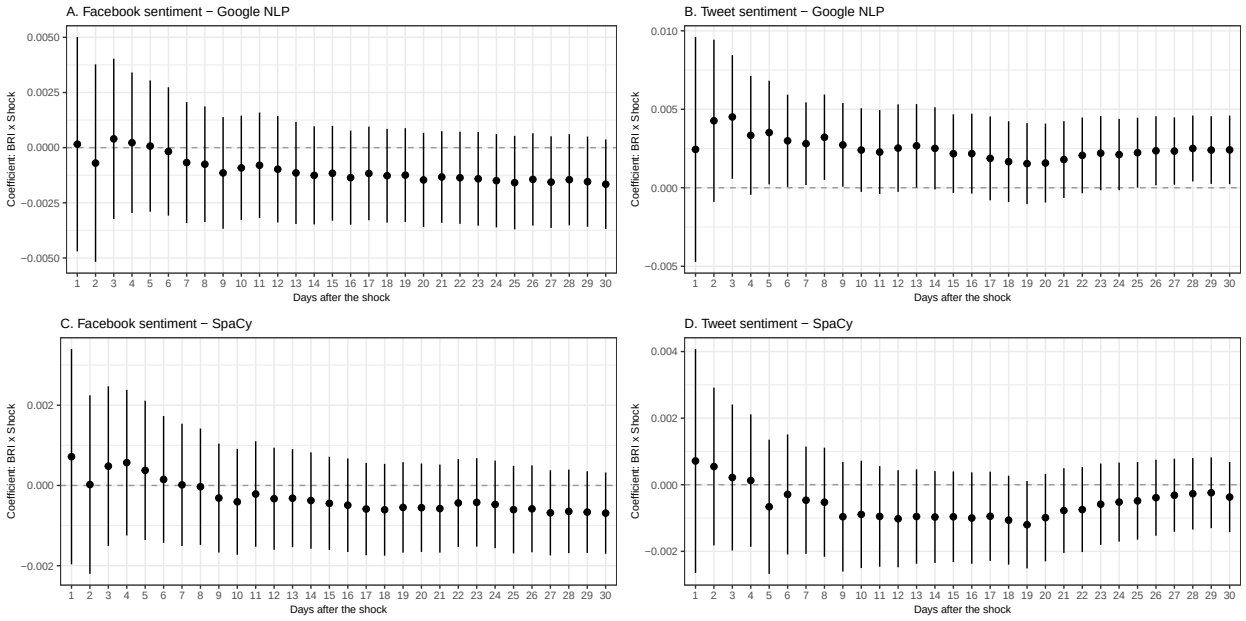
(c) Indonesia



Note: Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. Facebook sentiment toward China obtained with Google NLP API, B. tweet sentiment toward China obtained with Google NLP API, C. Facebook sentiment toward China using SpaCy Lexicon method, and D. tweet sentiment toward China obtained using SpaCy Lexicon method. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-28: Stacked DID on Elite Social Media Posts about China (SpaCy Sentiment): Politician-date level

(a) The Philippines



(b) Malaysia

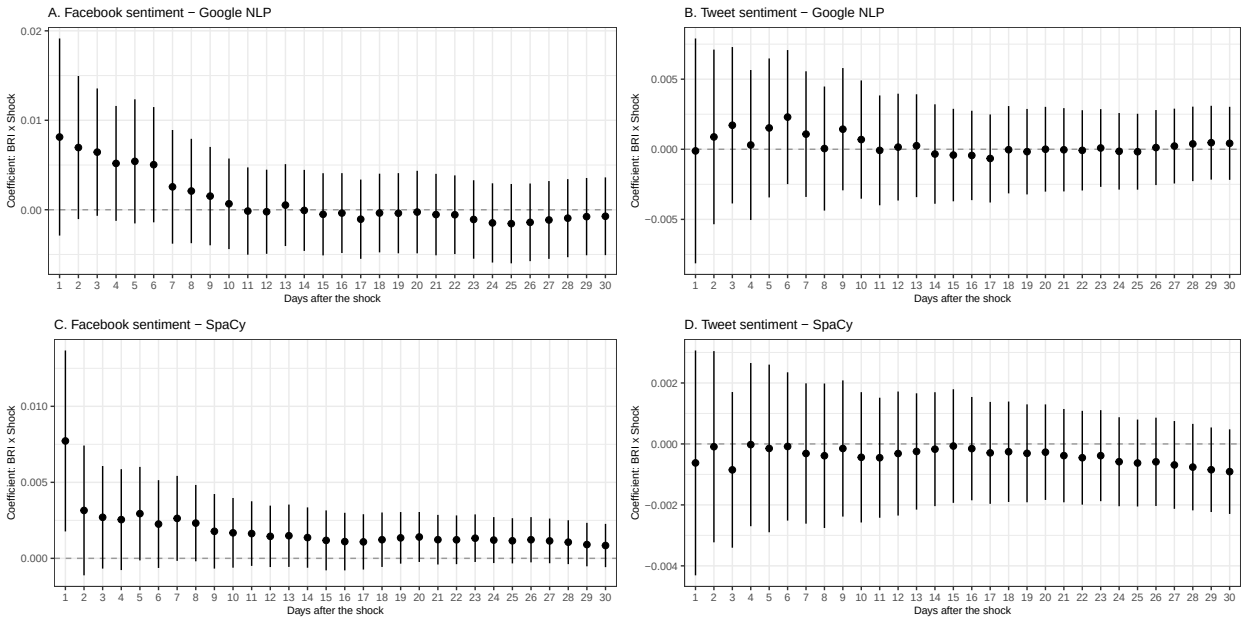
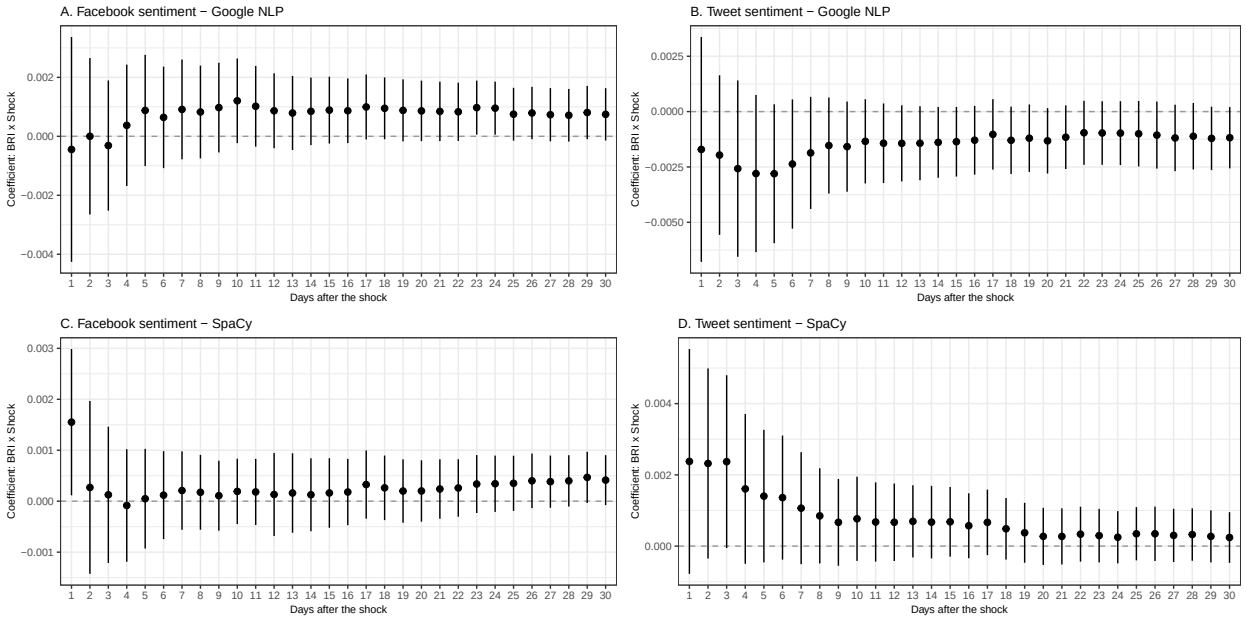


Figure A-28: Stacked DID on Elite Social Media Posts about China (SpaCy Sentiment): Politician-date level (continued)

(c) Indonesia



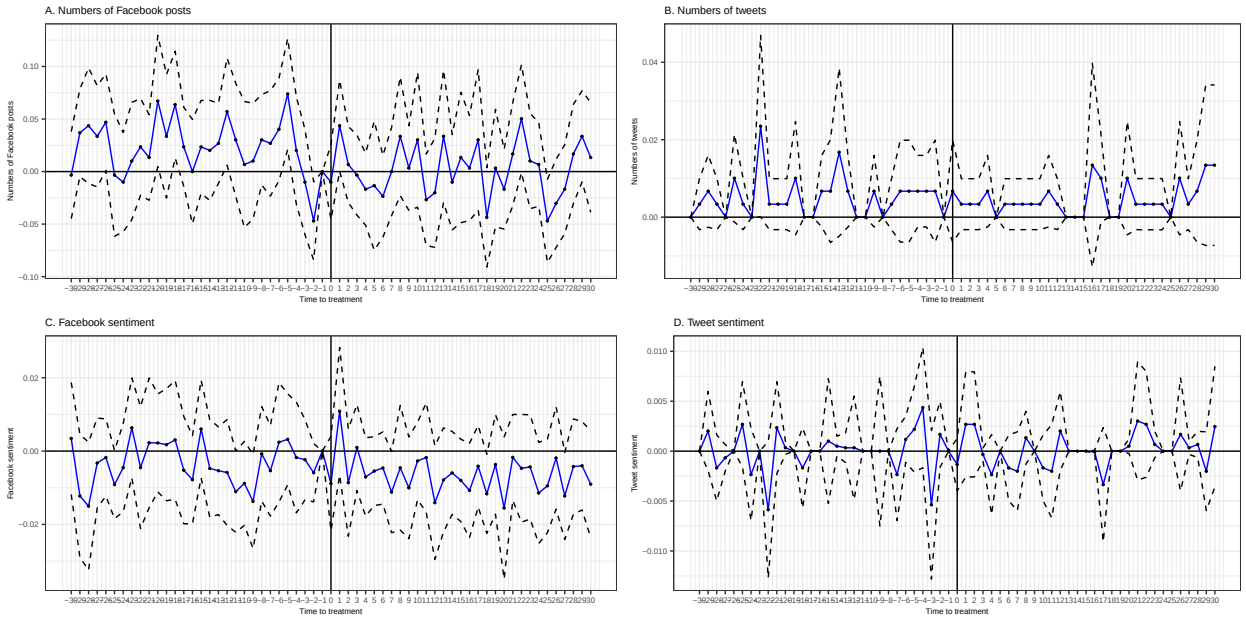
Note: Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. Facebook sentiment toward China obtained with Google NLP API, B. tweet sentiment toward China obtained with Google NLP API, C. Facebook sentiment toward China using SpaCy Lexicon method, and D. tweet sentiment toward China obtained using SpaCy Lexicon method. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the politician level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

4 Robustness Checks

4.1 Unit Fixed Effects Only

Figure A-29: Stacked Event Study: With Only Constituency Fixed Effect

(a) The Philippines



(b) Malaysia

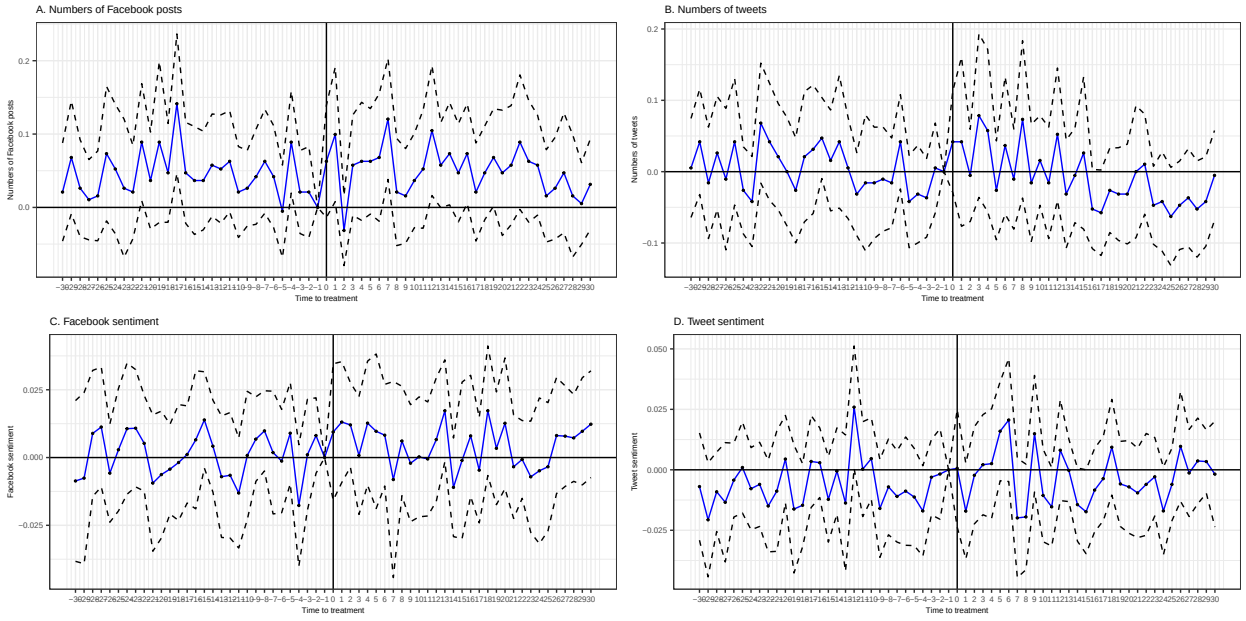
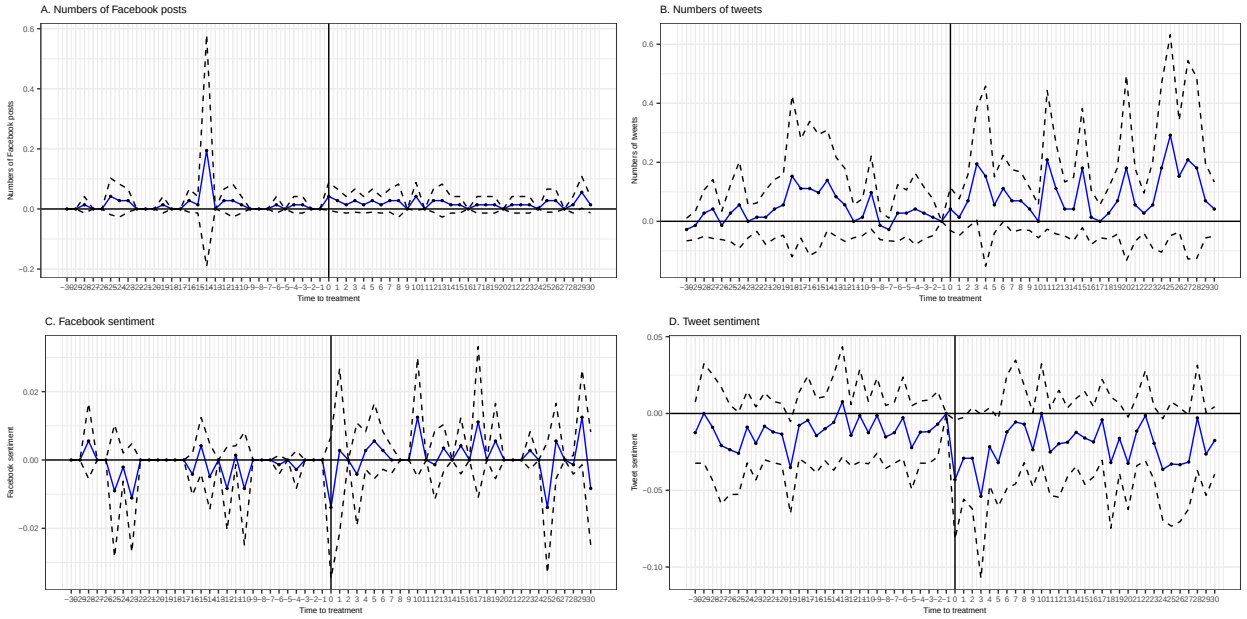


Figure A-29: Stacked Event Study: With Only Constituency Fixed Effect (continued)

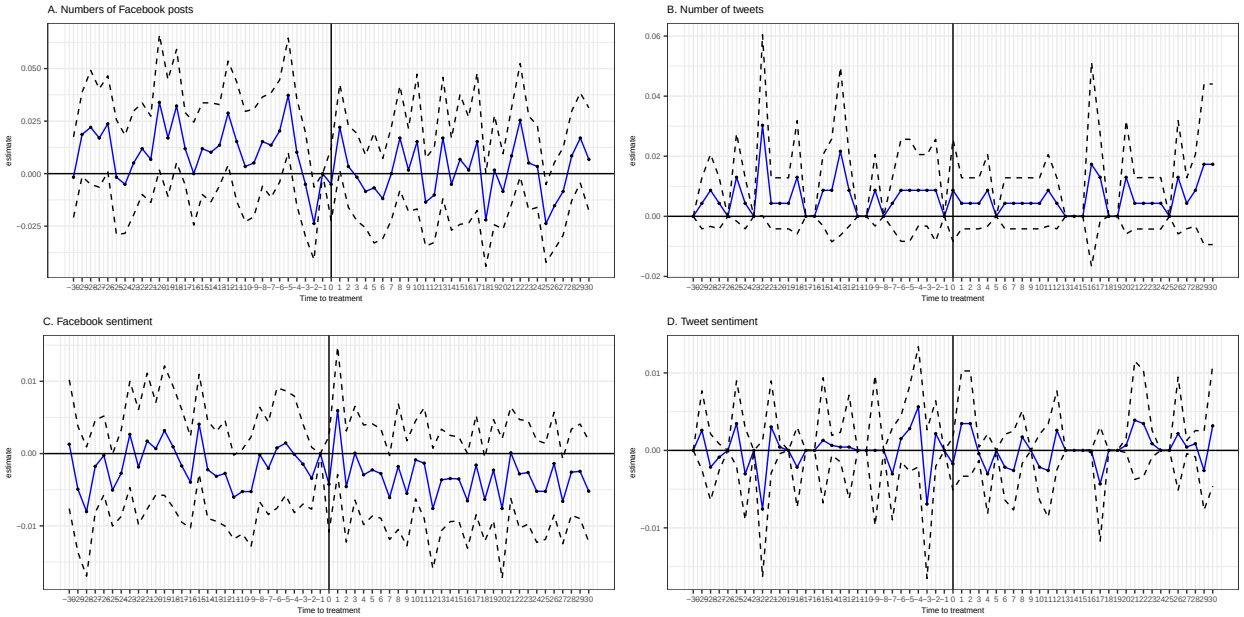
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the constituency-fixed effect.

Figure A-30: Stacked Event Study: With Only Politician Fixed Effect

(a) The Philippines



(b) Malaysia

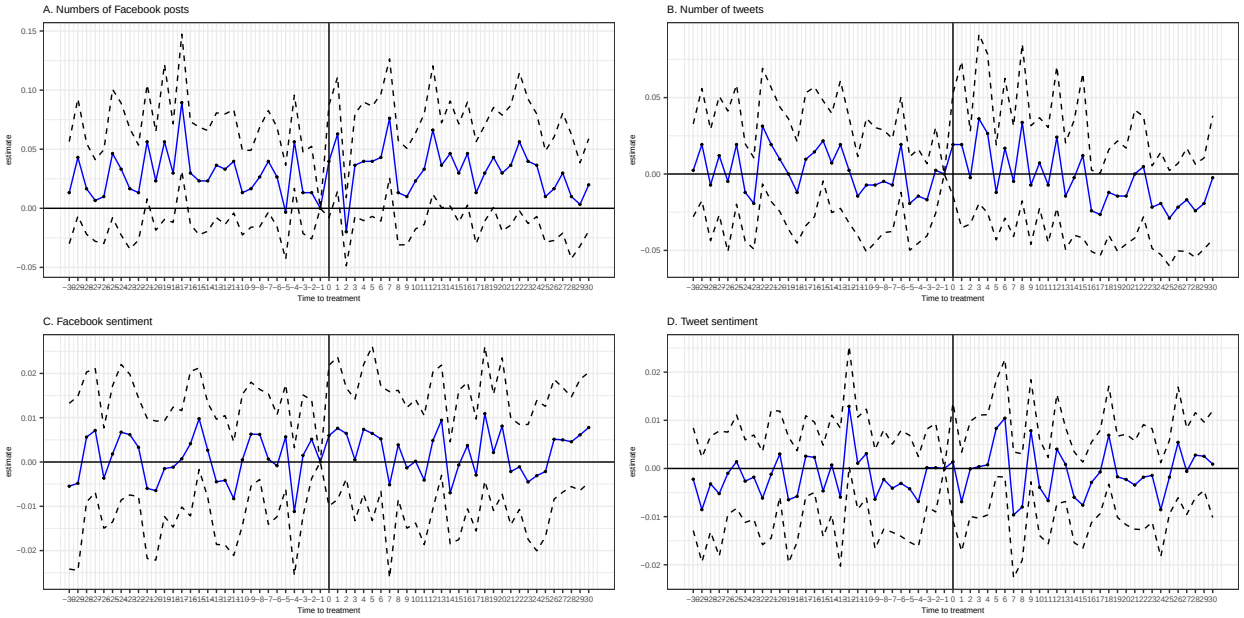
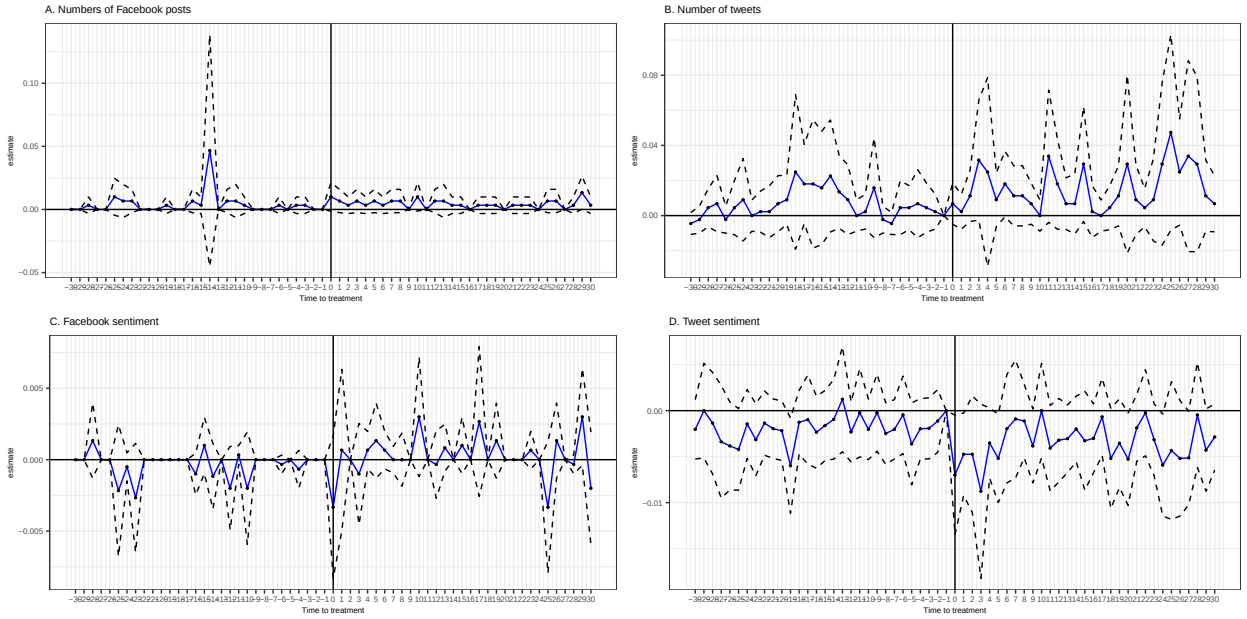


Figure A-30: Stacked Event Study: With Only Politician Fixed Effect (continued)

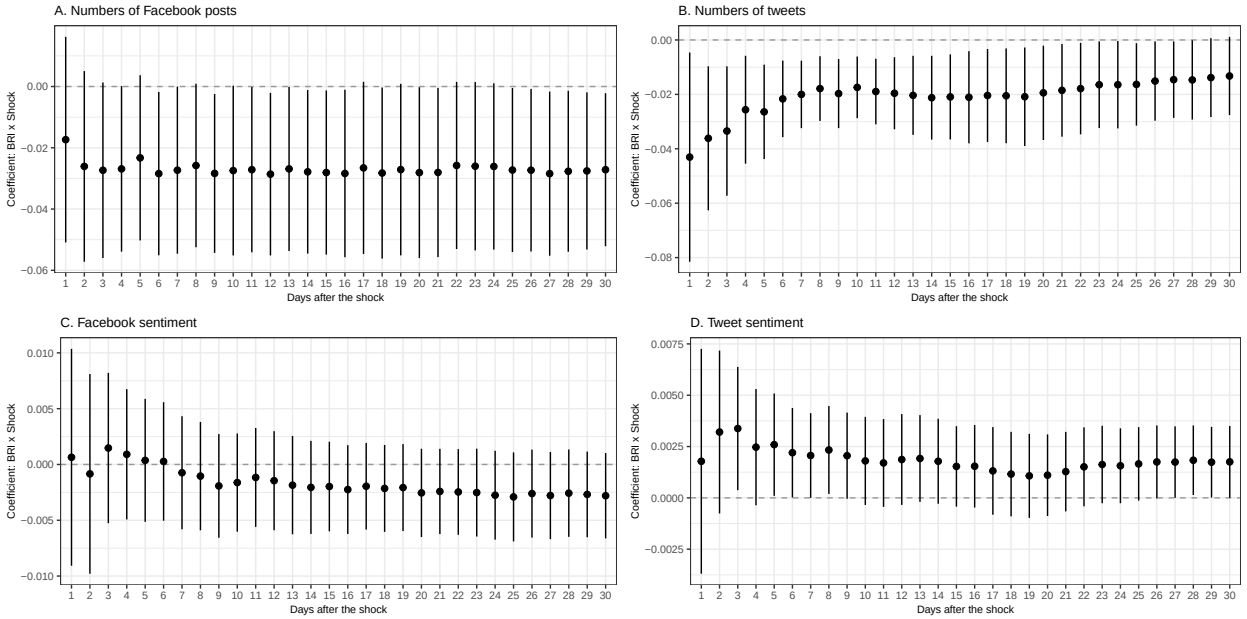
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the politician-fixed effect.

Figure A-31: Stacked Difference-in-difference: With Only Constituency Fixed Effect

(a) The Philippines



(b) Malaysia

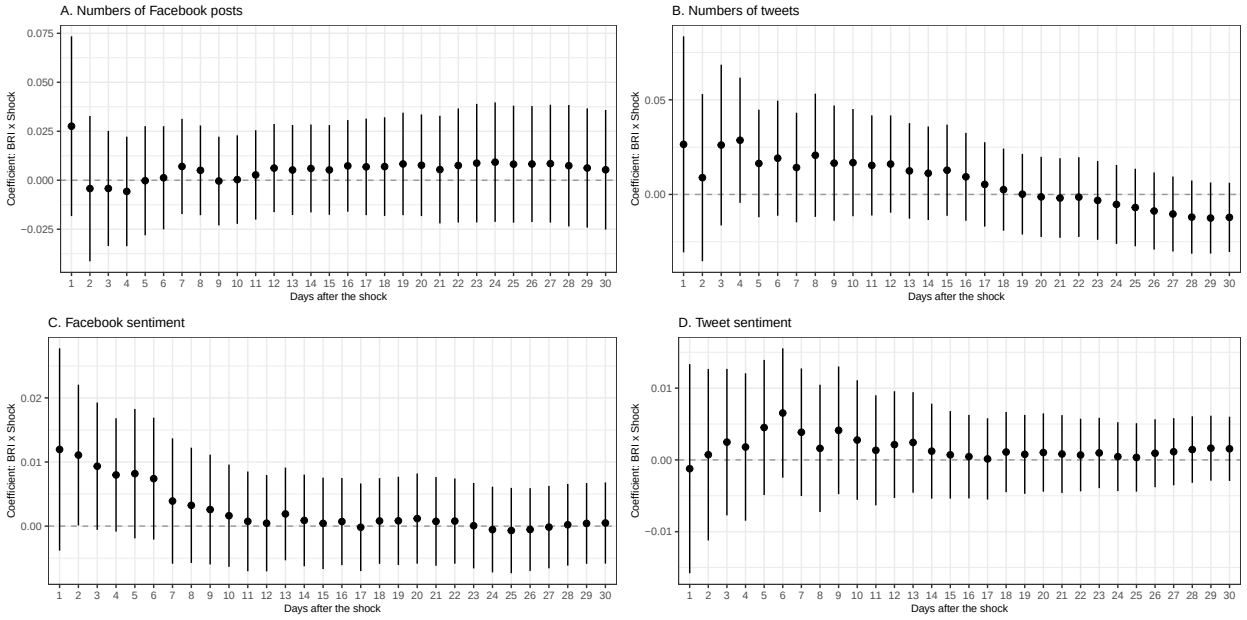
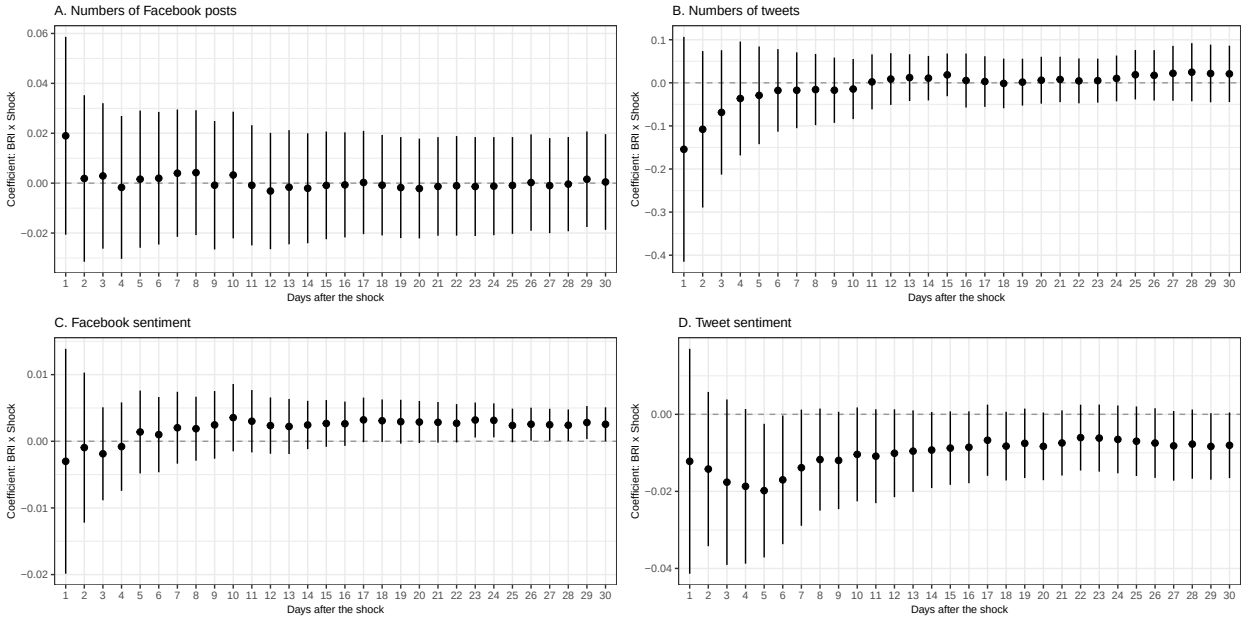


Figure A-31: Stacked Difference-in-difference: With Only Constituency Fixed Effect (continued)

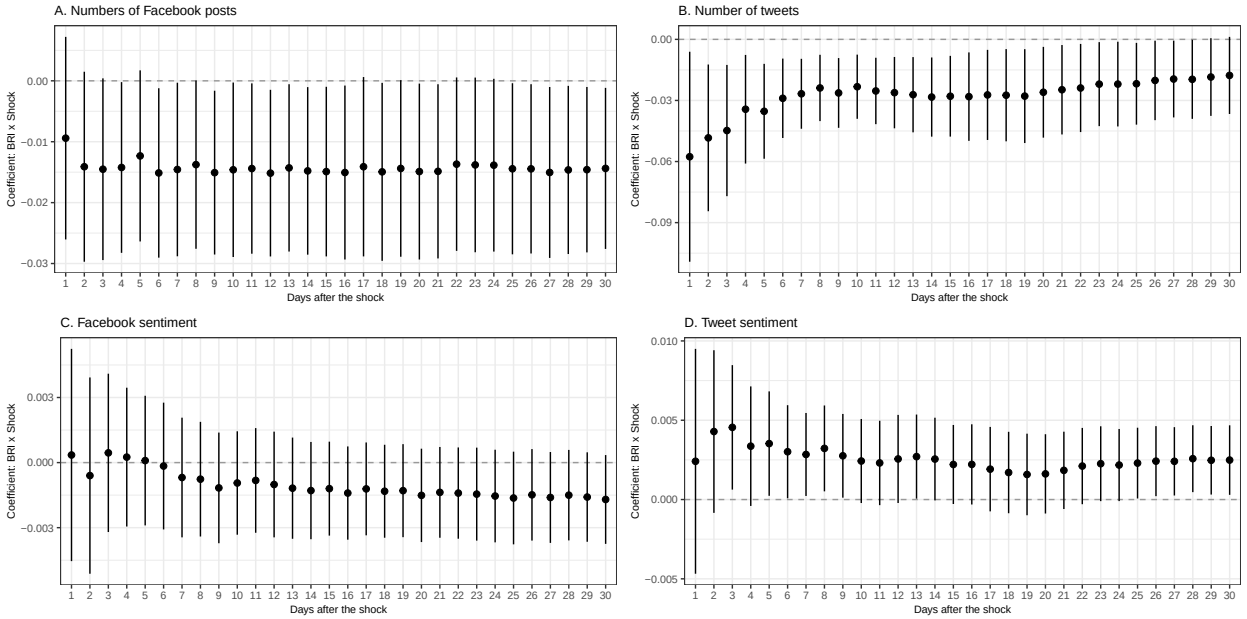
(c) Indonesia



Note: Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-32: Stacked Difference-in-difference: With Only Politician Fixed Effect

(a) The Philippines



(b) Malaysia

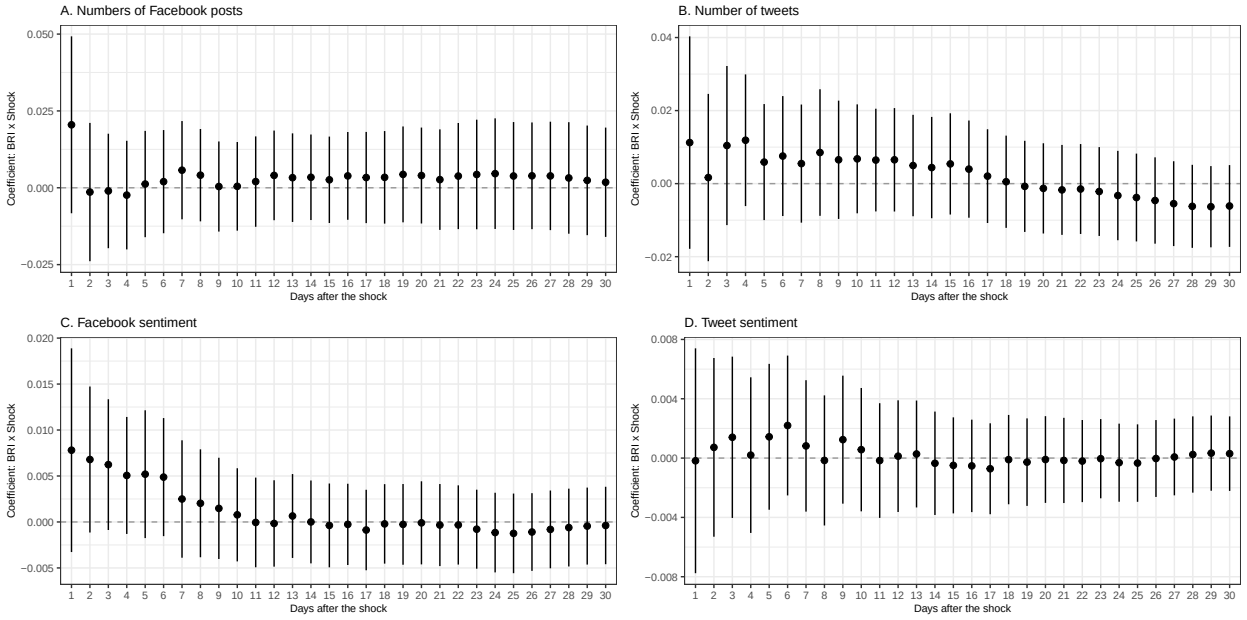
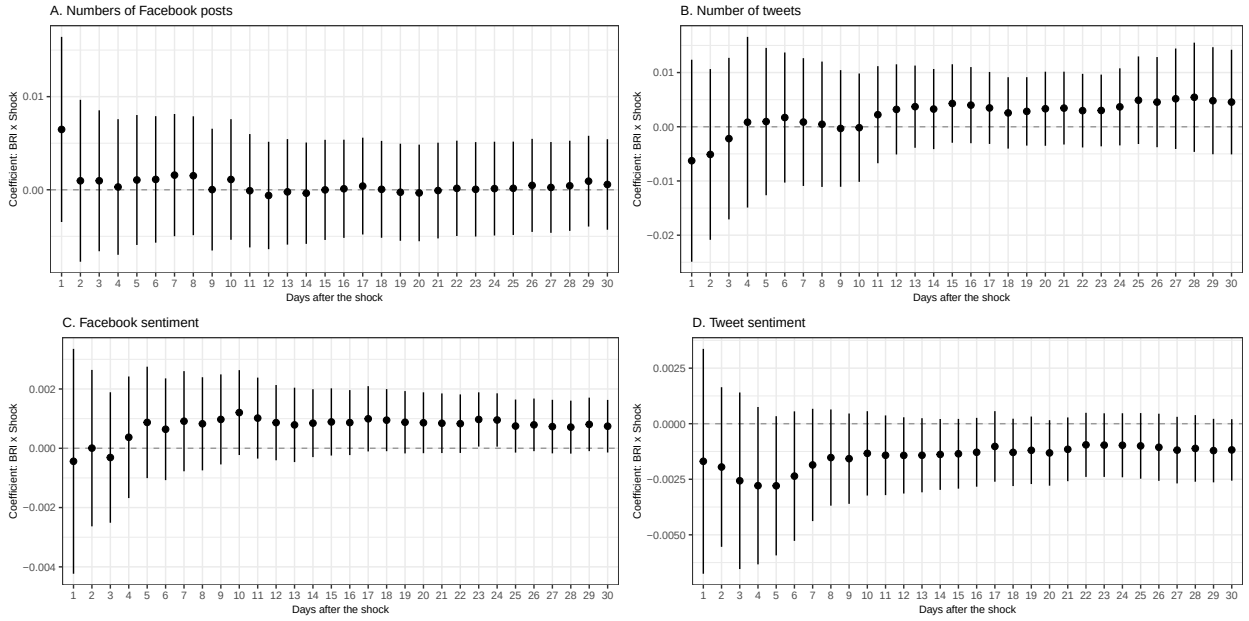


Figure A-32: Stacked Difference-in-difference: With Only Politician Fixed Effect (continued)

(c) Indonesia

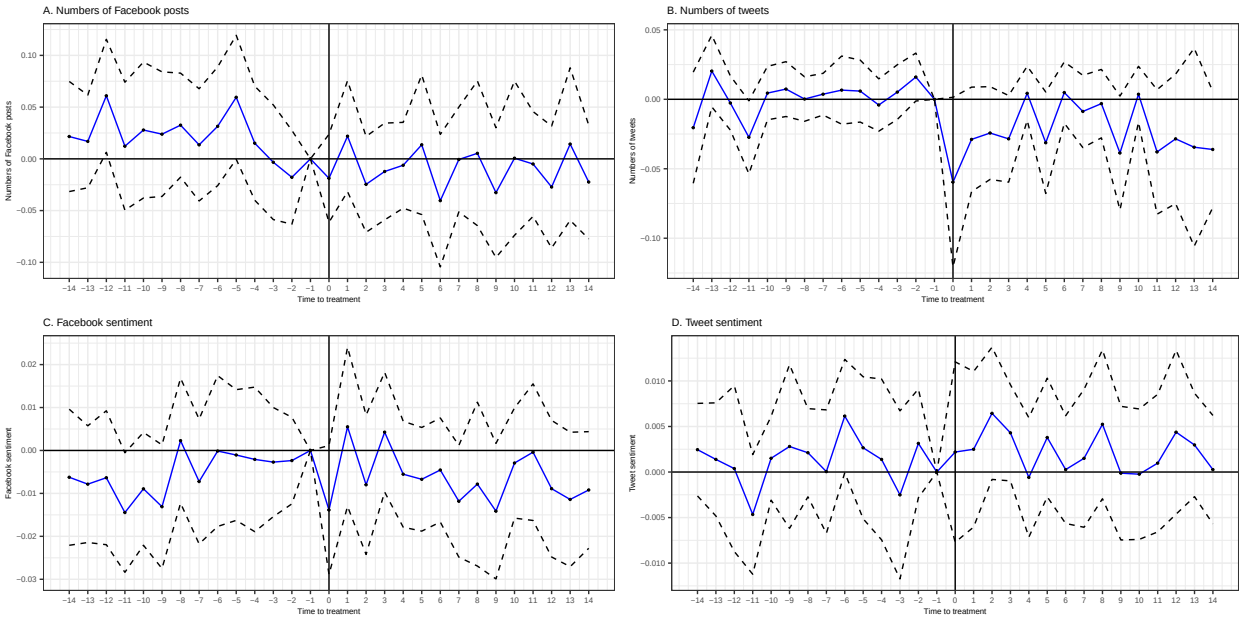


Note: Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

4.2 Event Windows

Figure A-33: Stacked Event Study: 28-day event window

(a) The Philippines



(b) Malaysia

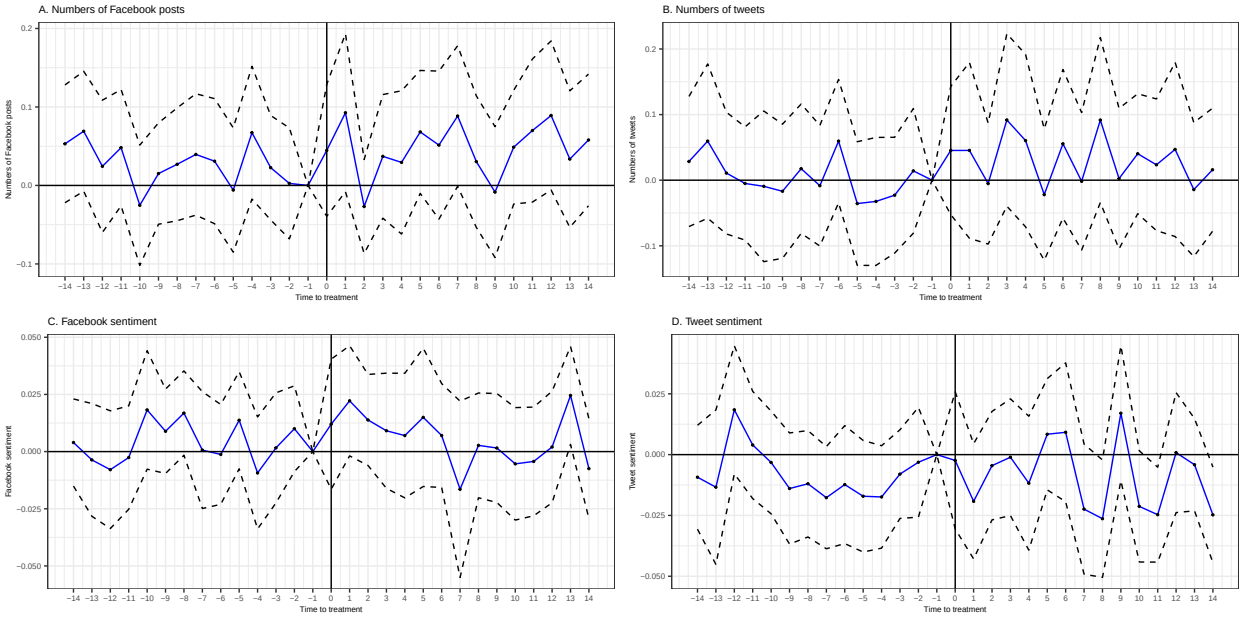
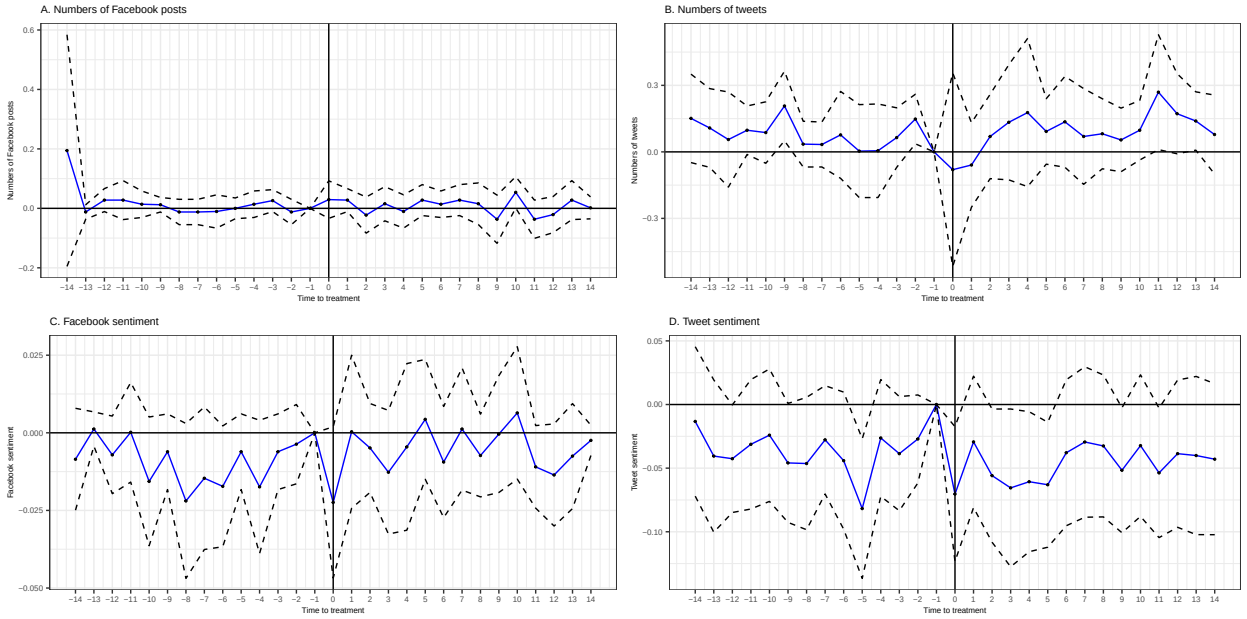


Figure A-33: Stacked Event Study: 28-day event window (continued)

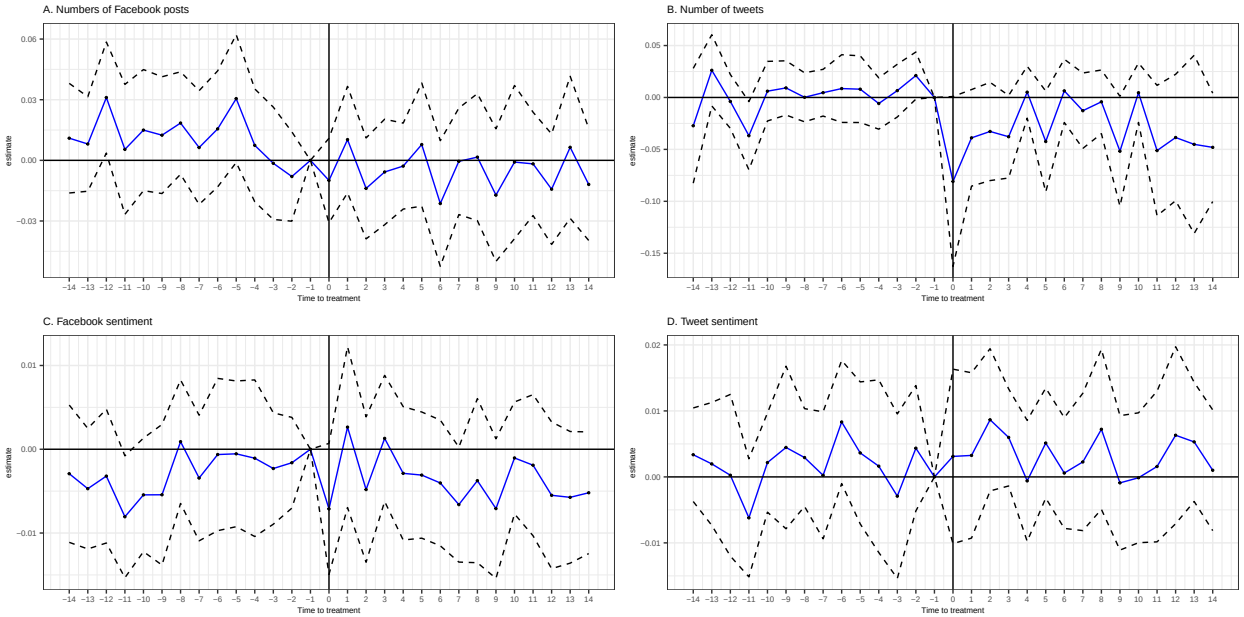
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. In this design, the pre-event window and post-window are both set to 14 days. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-34: Stacked Event Study at Politician-date Level: 28-day event window

(a) The Philippines



(b) Malaysia

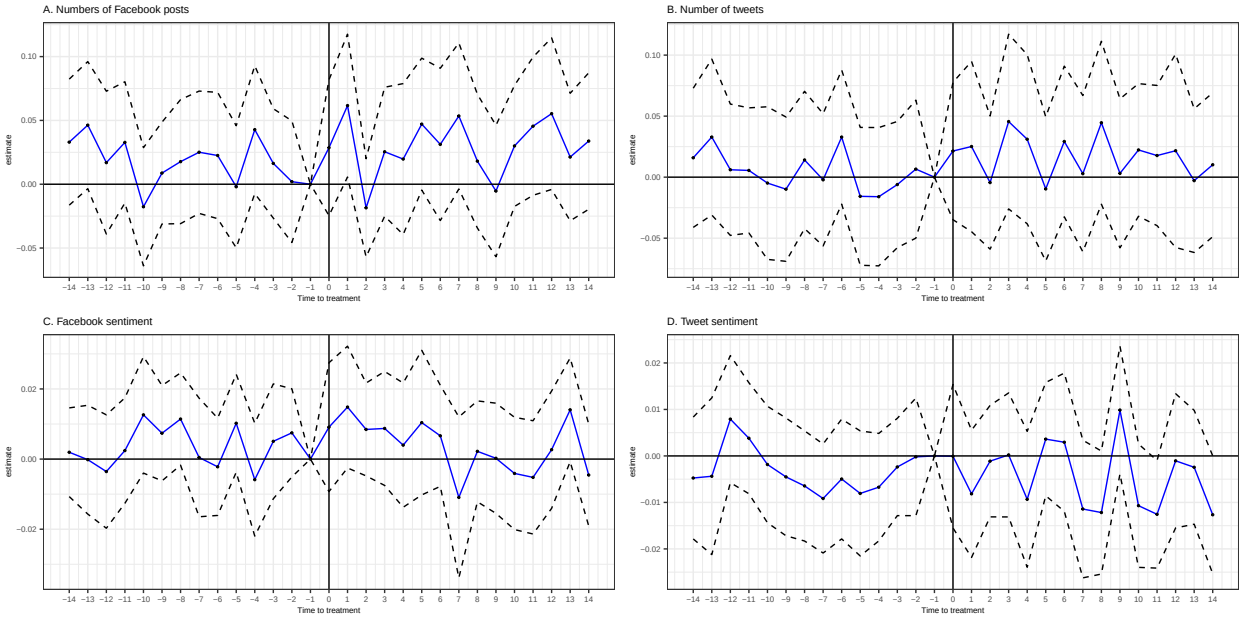
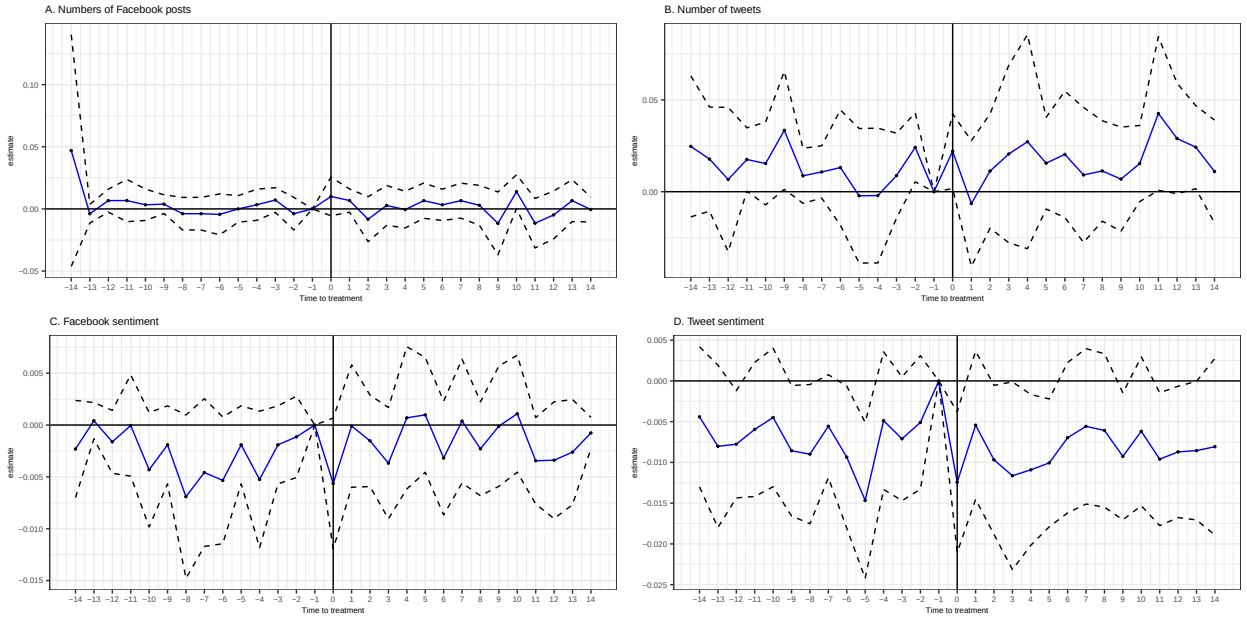


Figure A-34: Stacked Event Study at Politician-date Level: 28-day event window (continued)

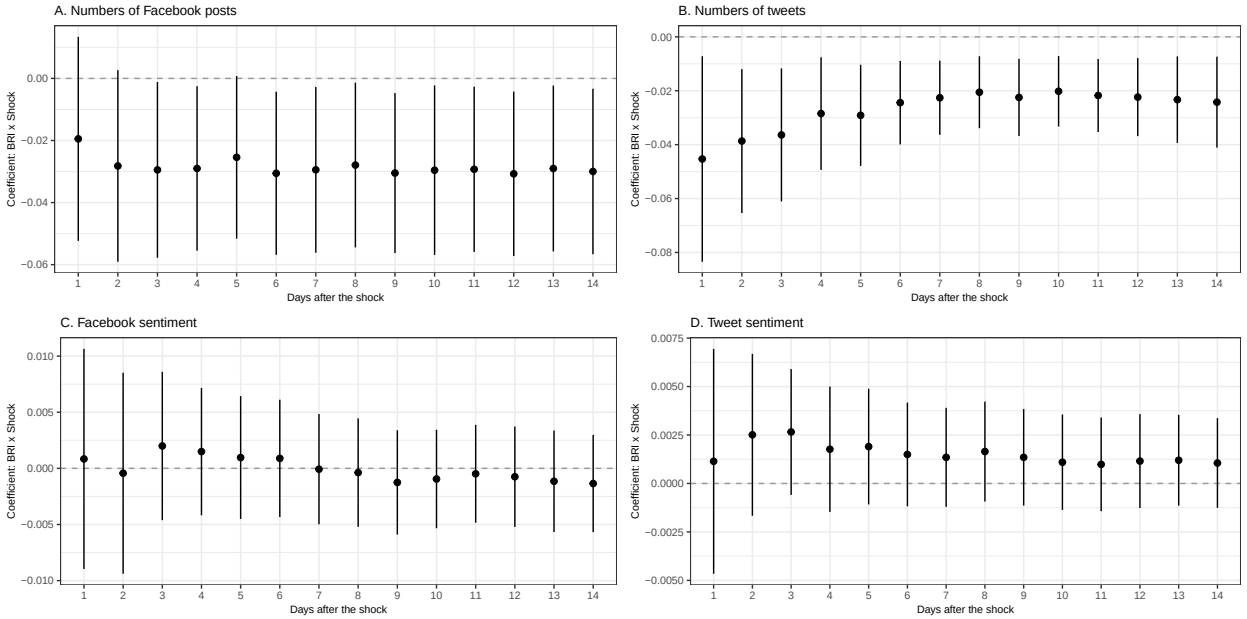
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. In this design, the pre-event window and post-window are both set to 14 days. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

Figure A-35: Stacked Difference-in-Difference: 28-day event window

(a) The Philippines



(b) Malaysia

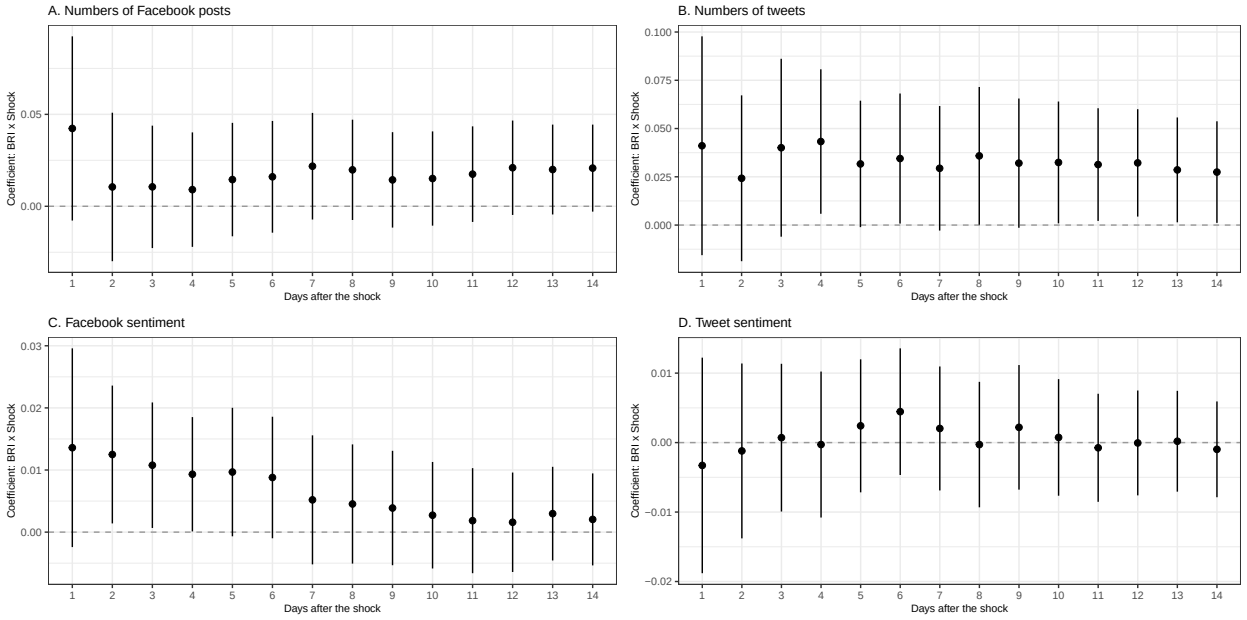
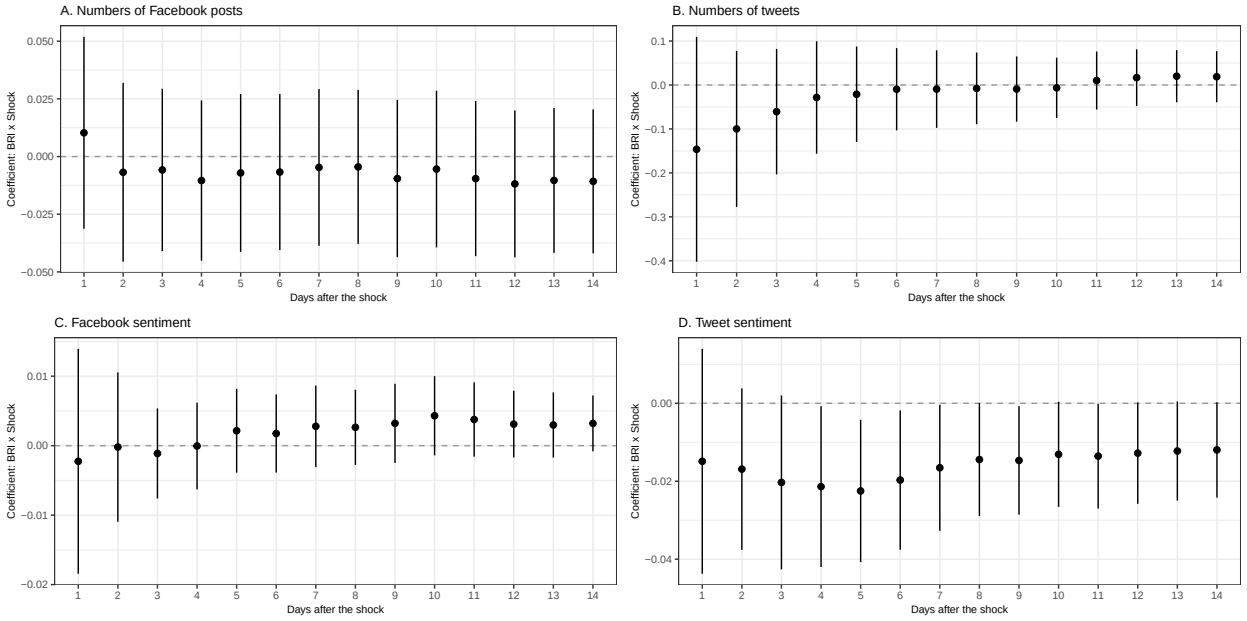


Figure A-35: Stacked Difference-in-Difference: 28-day event window (continued)

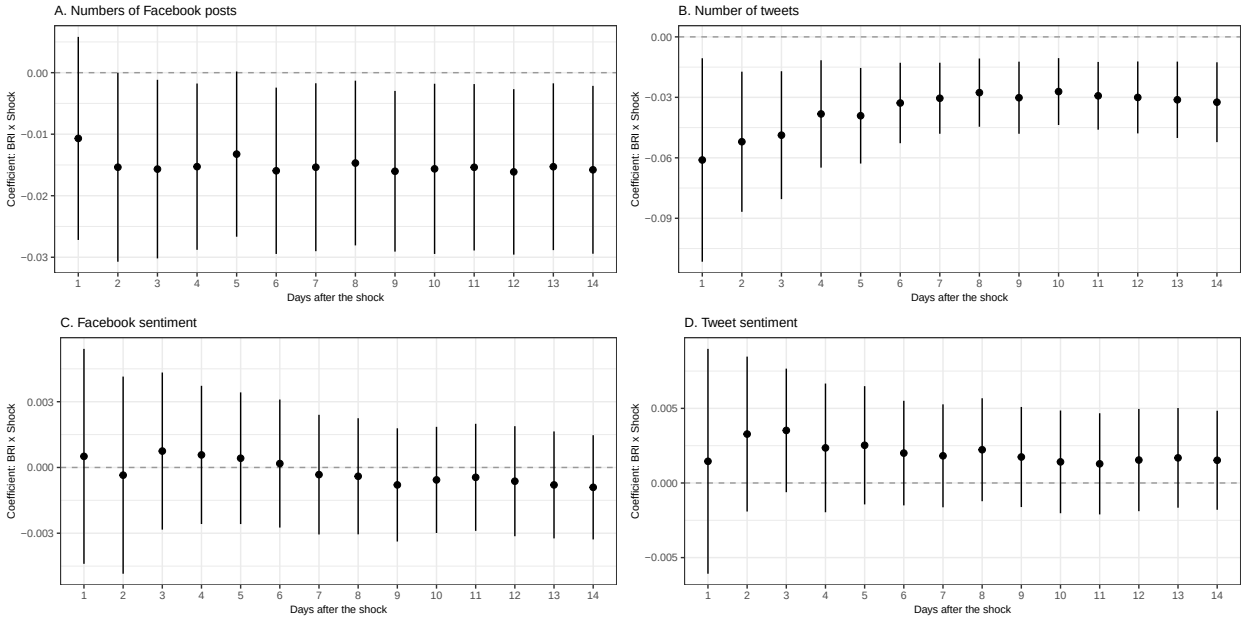
(c) Indonesia



Note: Pre-event window and post-window are both set to 14 days. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-36: Stacked Difference-in-Difference at Politician-date Level: 28-day event window

(a) The Philippines



(b) Malaysia

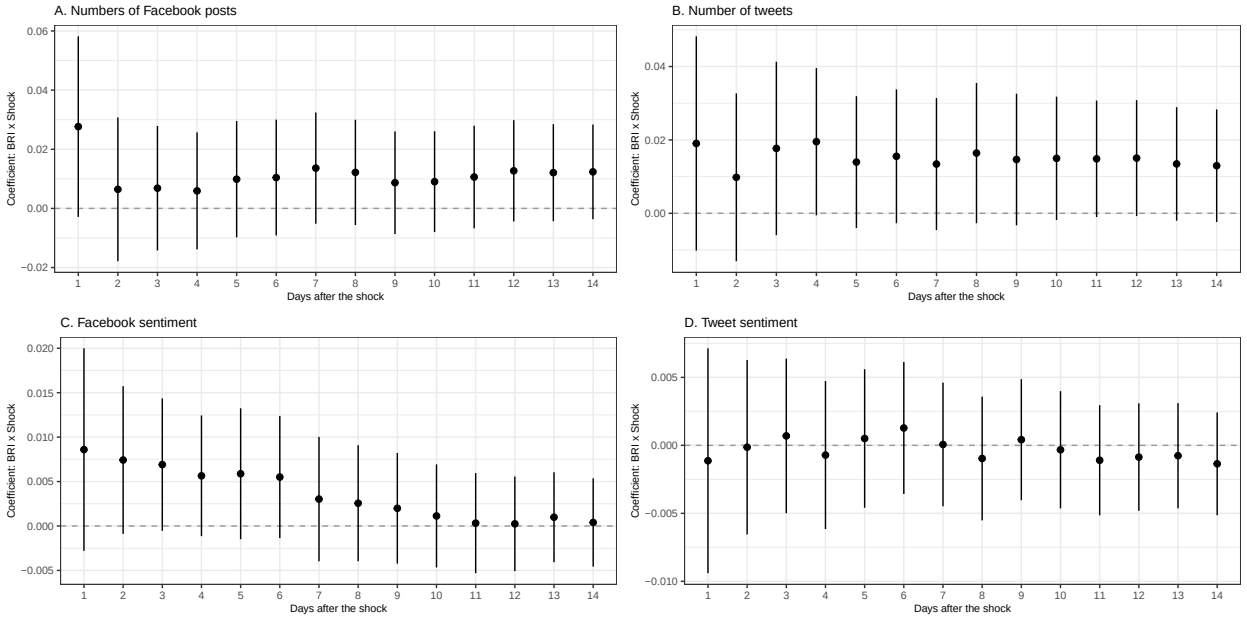
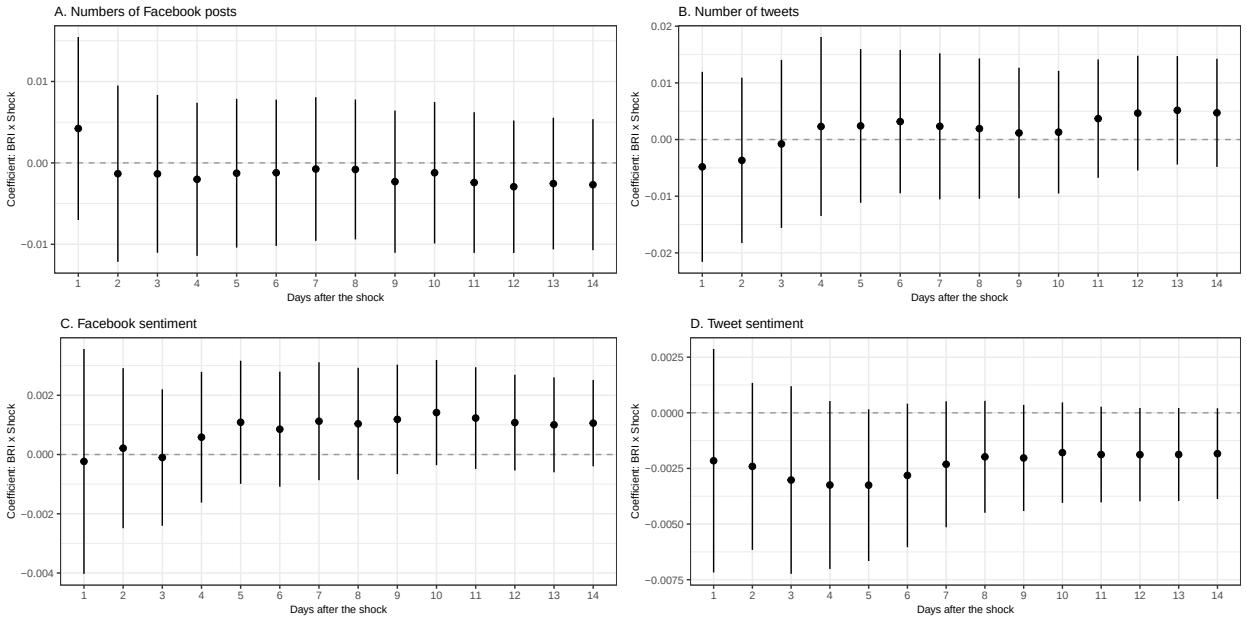


Figure A-36: Stacked Difference-in-Difference at Politician-date Level: 28-day event window (continued)

(c) Indonesia



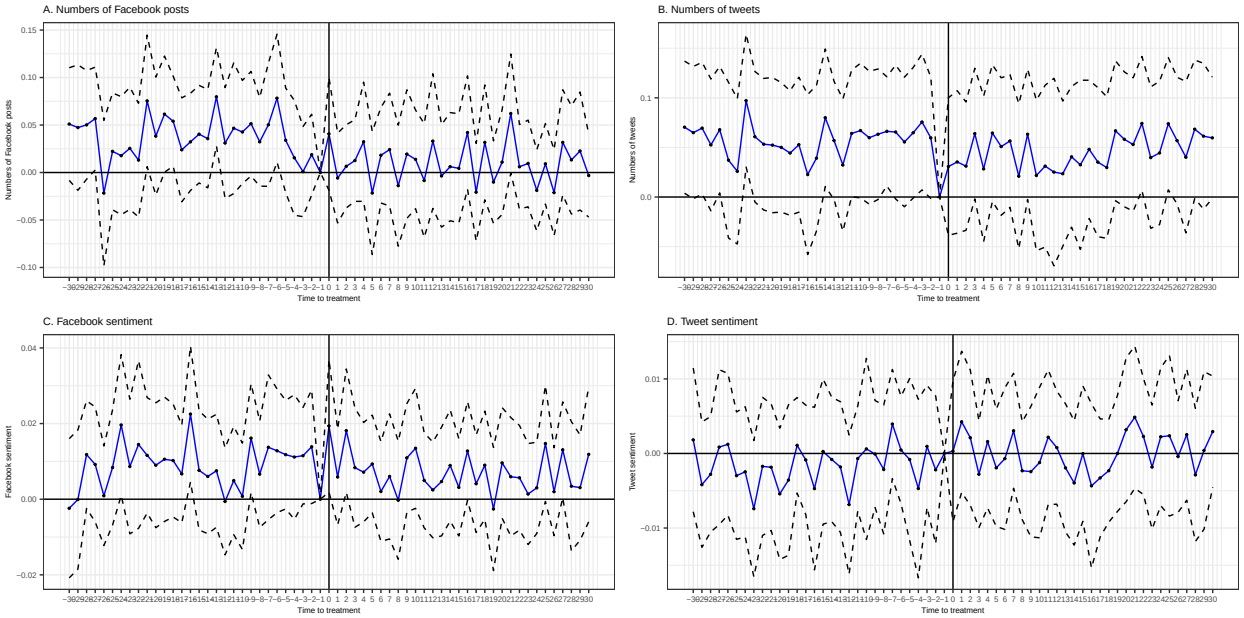
Note: Pre-event window and post-window are both set to 14 days. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

4.3 Placebo Tests

4.3.1 Shock = One day after the event date

Figure A-37: Stacked Event Study: One Day After the Event Date

(a) The Philippines



(b) Malaysia

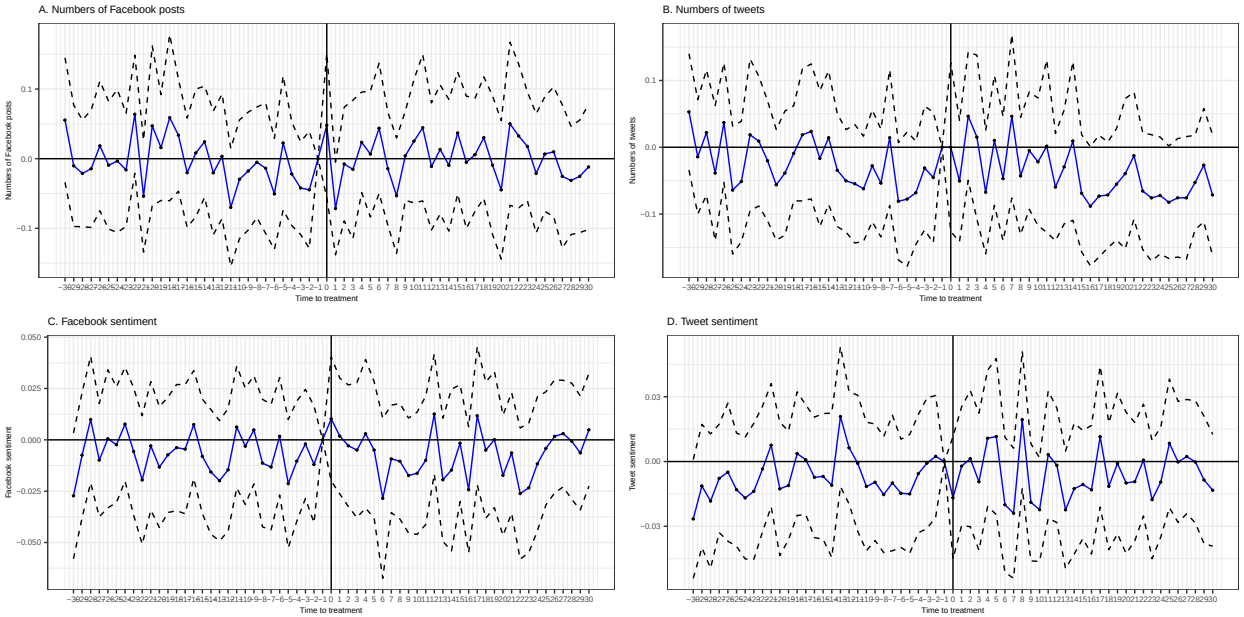
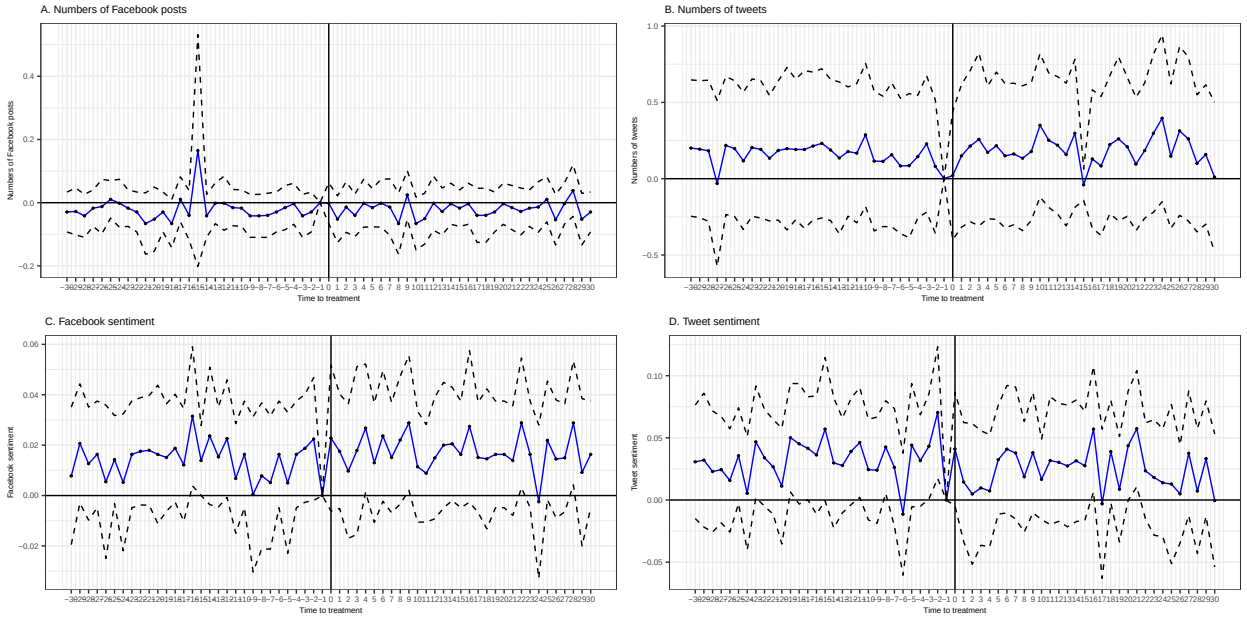


Figure A-37: Stacked Event Study: One Day After the Event Date (continued)

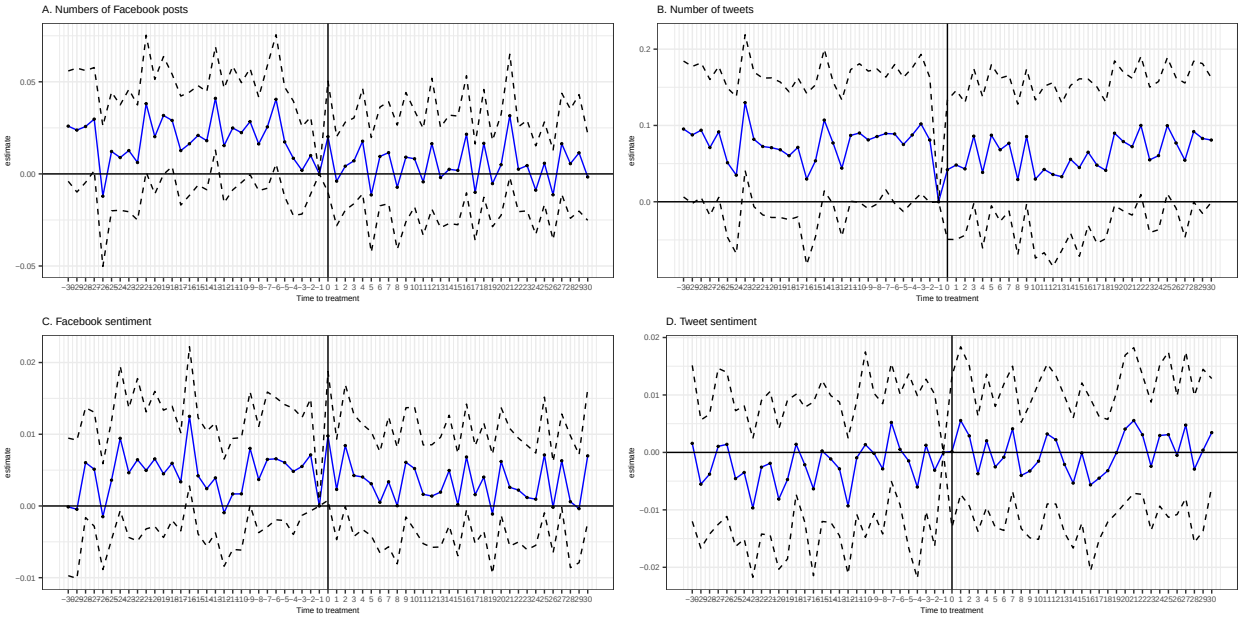
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. In this placebo test, the event date is the original date of the shock + 1. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-38: Stacked Event Study At Politician-Date Level: One Day After the Event Date

(a) The Philippines



(b) Malaysia

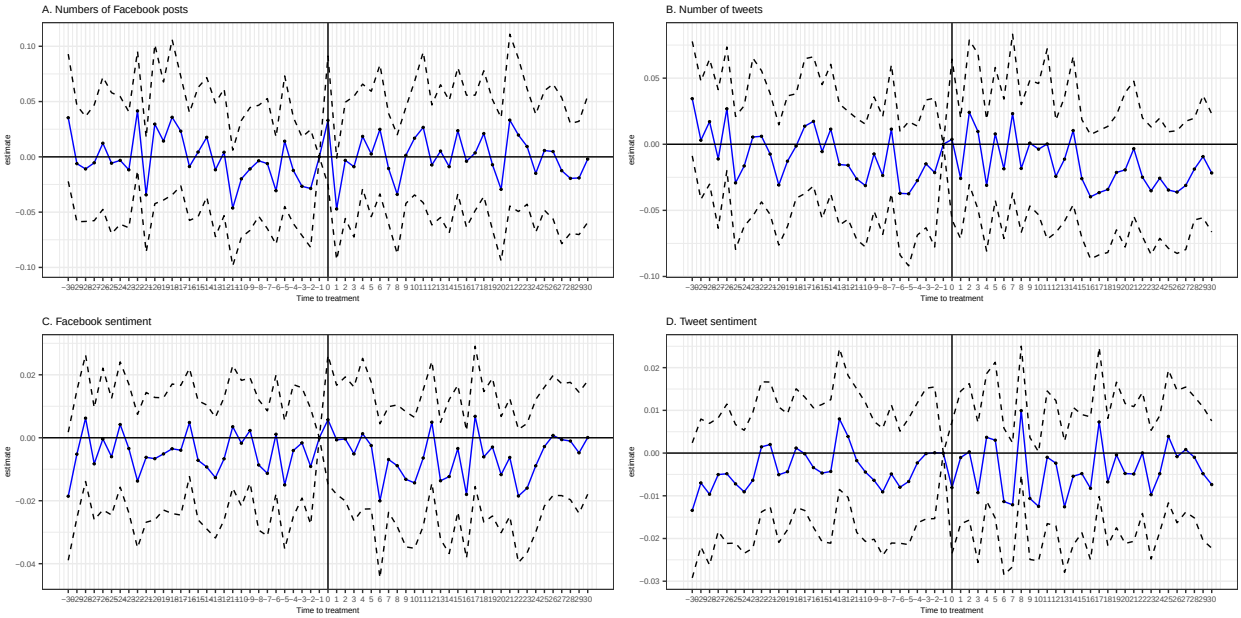
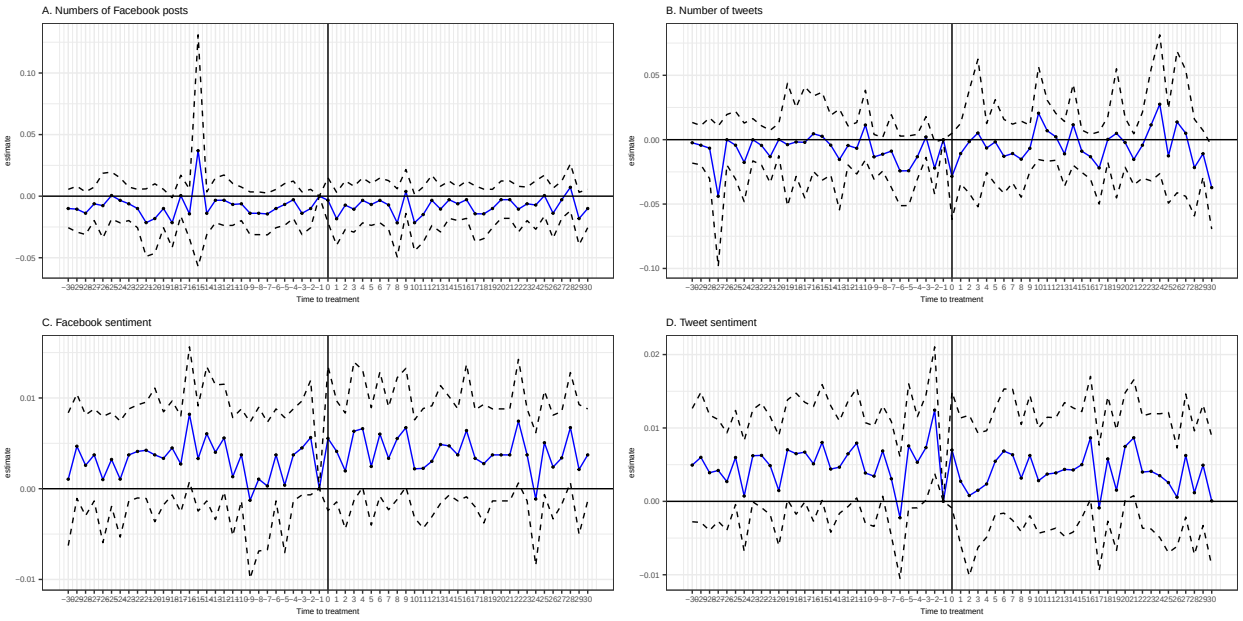


Figure A-38: Stacked Event Study At Politician-Date Level: One Day After the Event Date (continued)

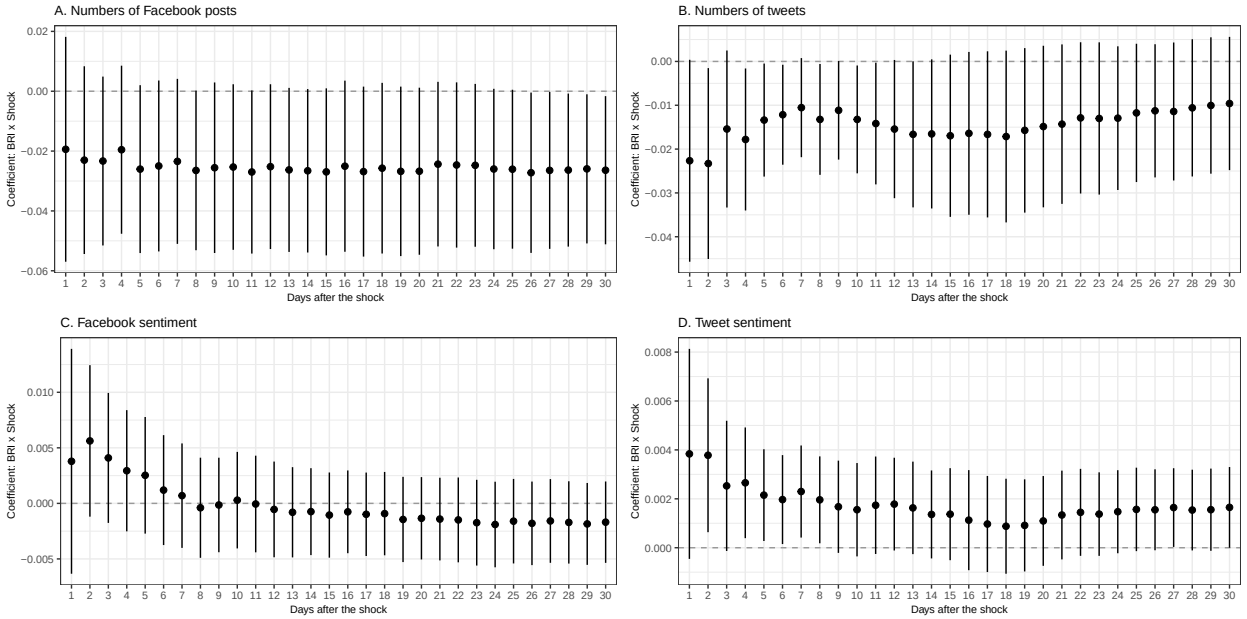
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. In this placebo test, the event date is the original date of the shock + 1. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-39: Stacked Difference-in-Difference: Placebo Event Date = one day after the shock

(a) The Philippines



(b) Malaysia

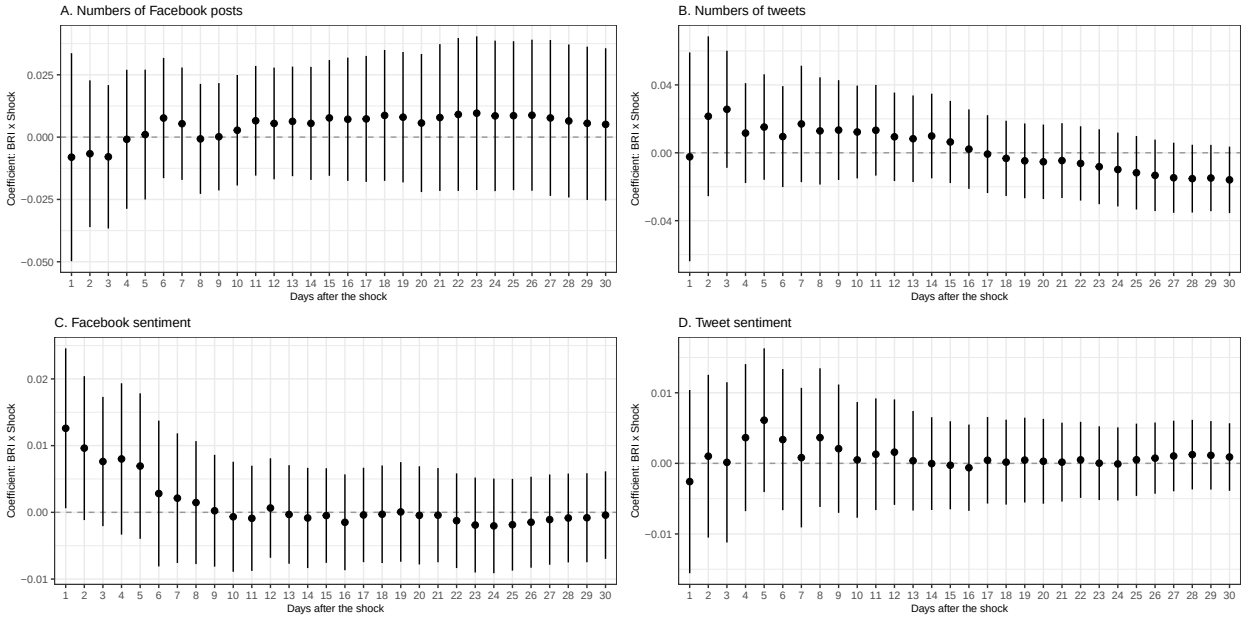
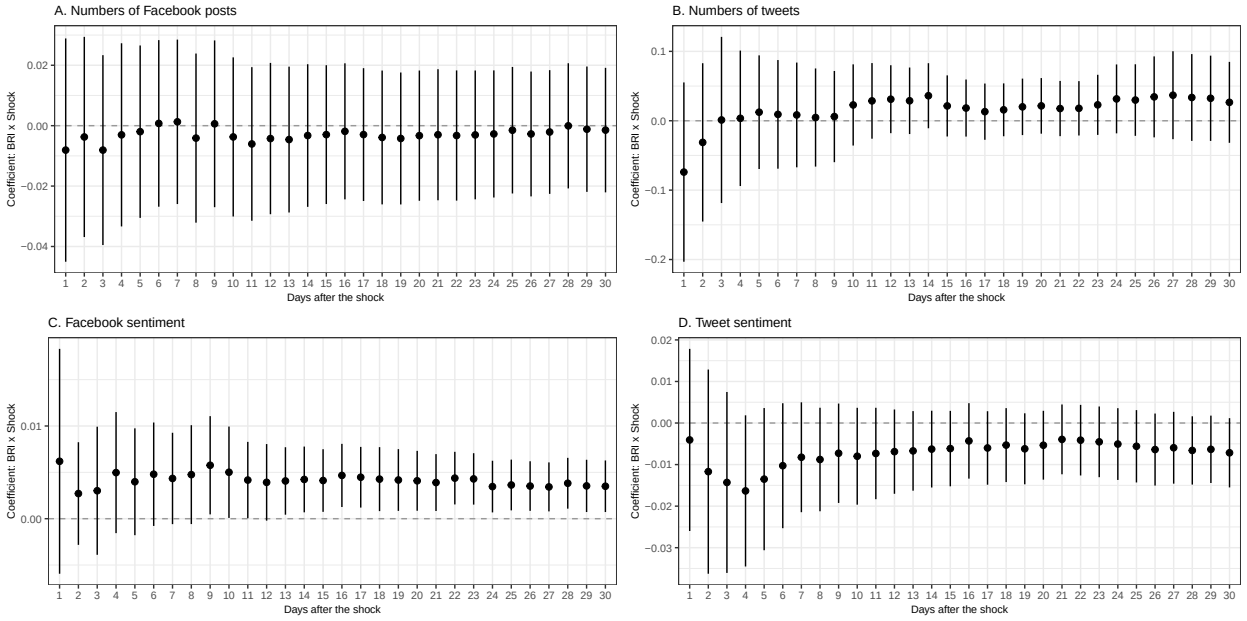


Figure A-39: Stacked Event Study: Placebo Event Date = one day after the shock (continued)

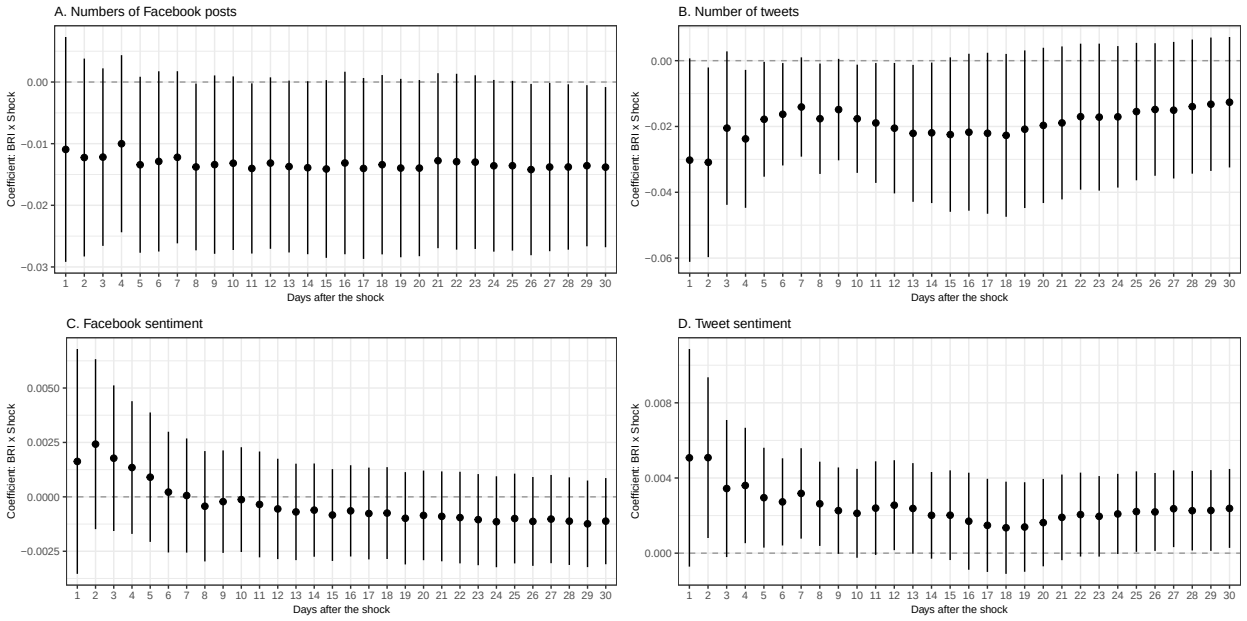
(c) Indonesia



Note: In this placebo test, the event date is the original date of the shock + 1. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-40: Stacked Difference-in-Difference at Politician-date Level: Placebo Event Date = one day after the shock

(a) The Philippines



(b) Malaysia

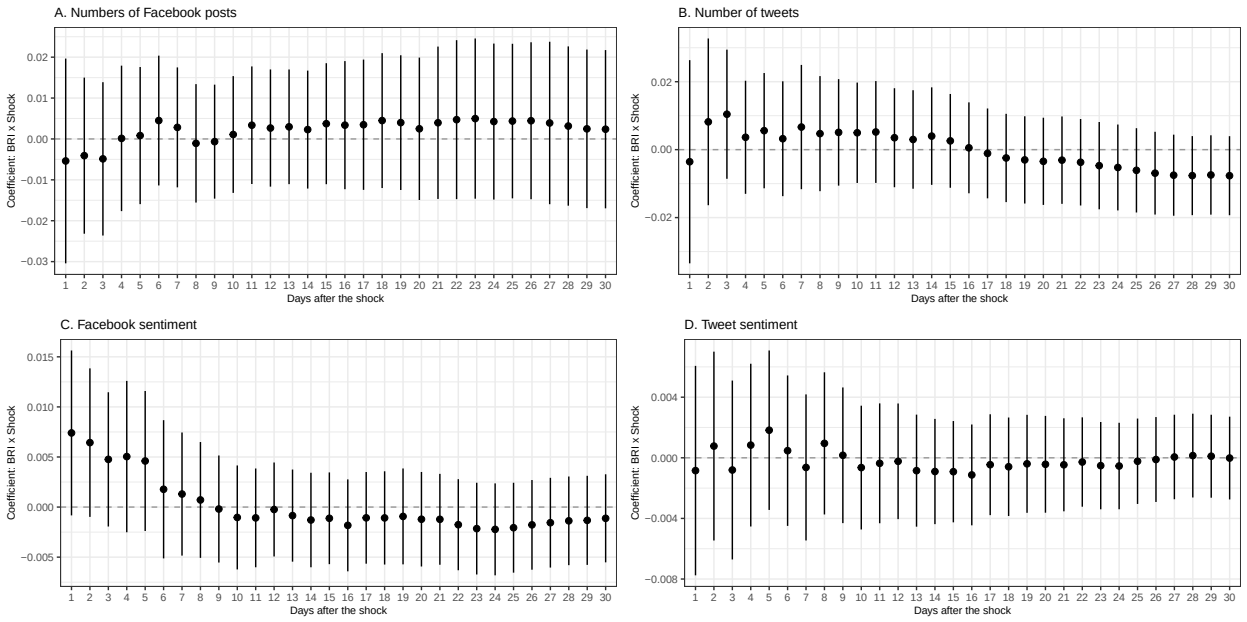
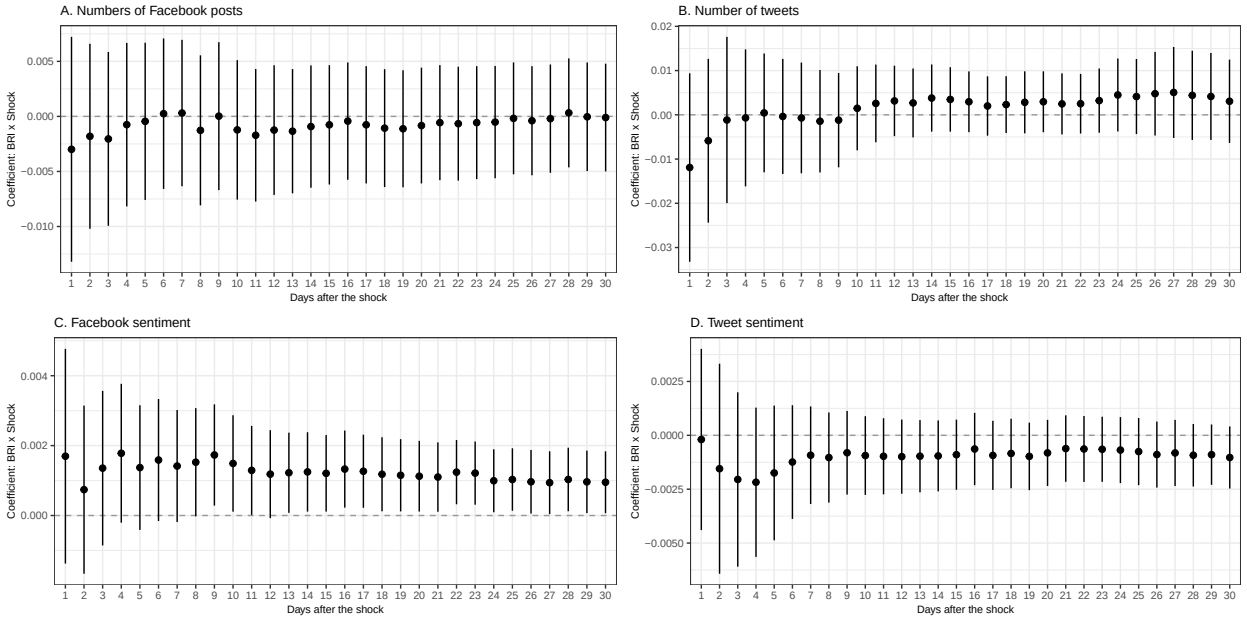


Figure A-40: Stacked Event Study: Placebo Event Date = one day after the shock (continued)

(c) Indonesia

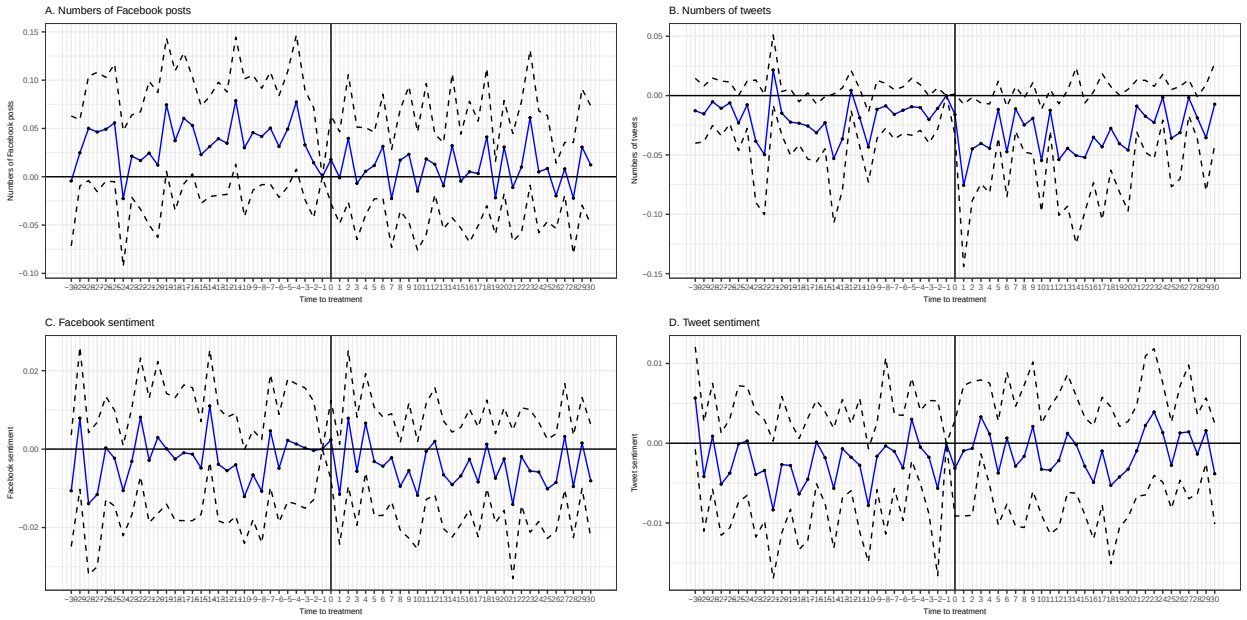


Note: In this placebo test, the event date is the original date of the shock + 1. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

4.3.2 Shock = One day before the shock

Figure A-41: Stacked Event Study: Placebo Event Date = one day before the shock

(a) The Philippines



(b) Malaysia

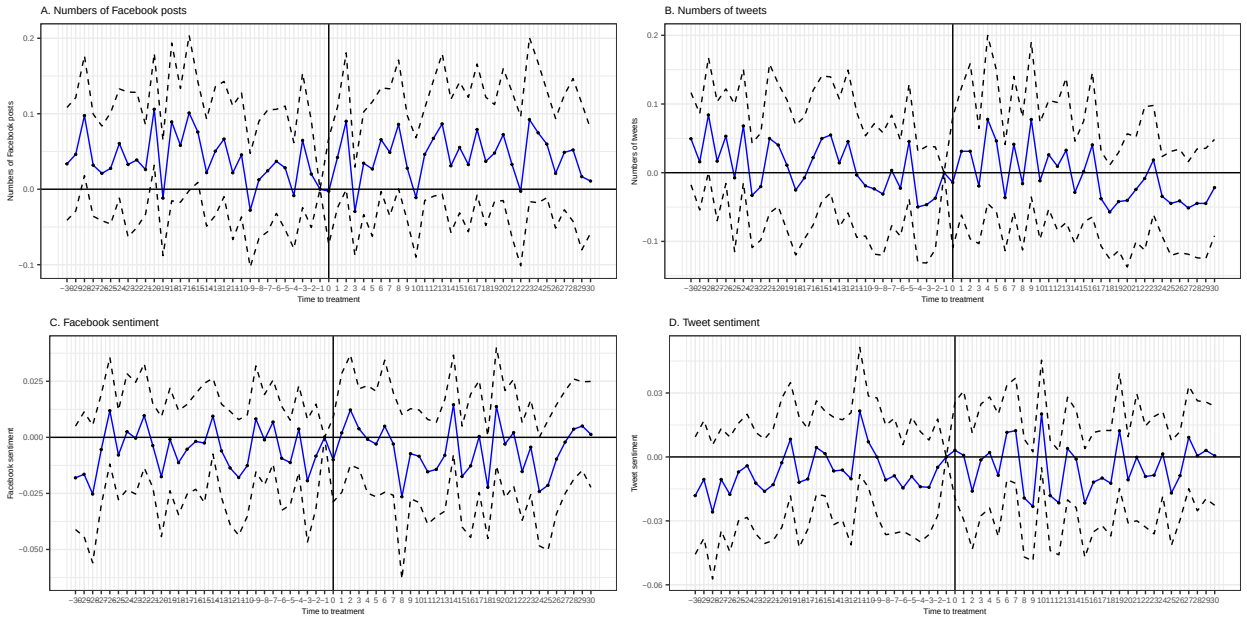
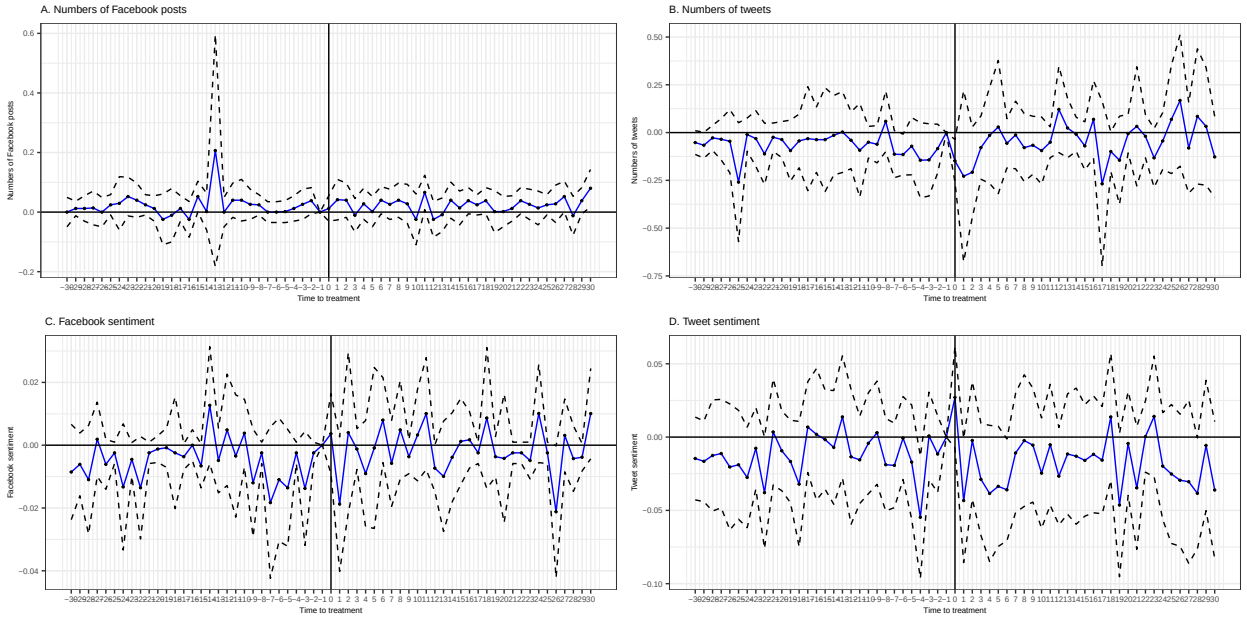


Figure A-41: Stacked Event Study: Placebo Event Date = one day before the shock (continued)

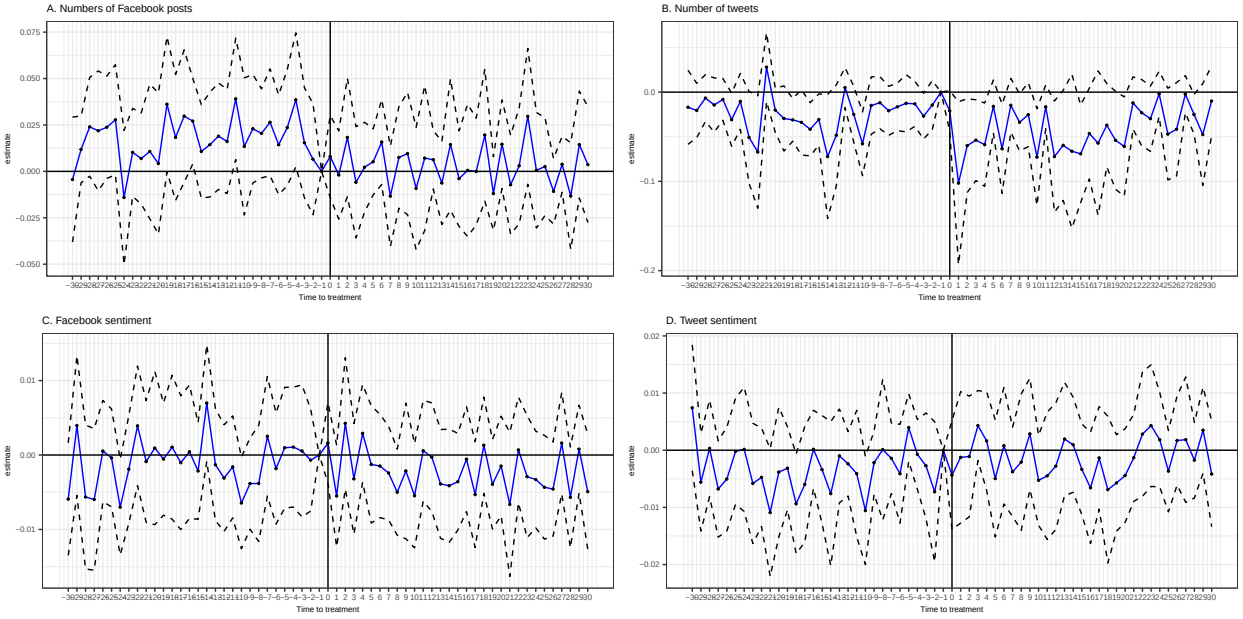
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. In this placebo test, the event date is the original date of the shock - 1. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-42: Stacked Event Study at Politician-date Level: Placebo Event Date = one day before the shock

(a) The Philippines



(b) Malaysia

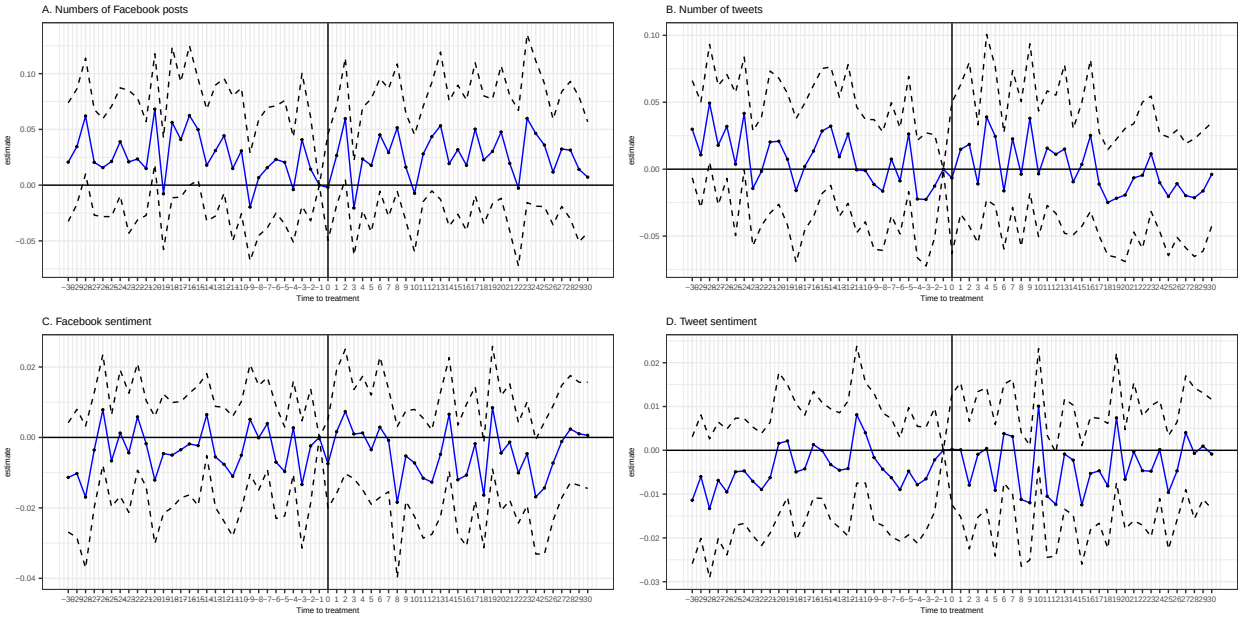
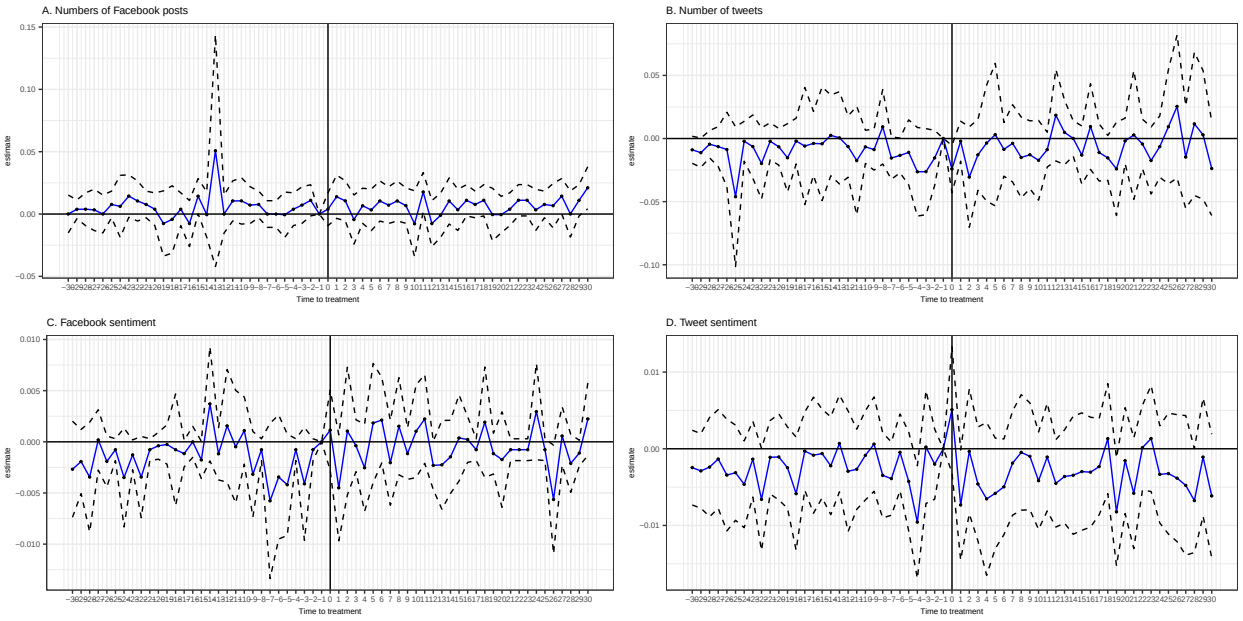


Figure A-42: Stacked Event Study at Politician-date Level: Placebo Event Date = one day before the shock (continued)

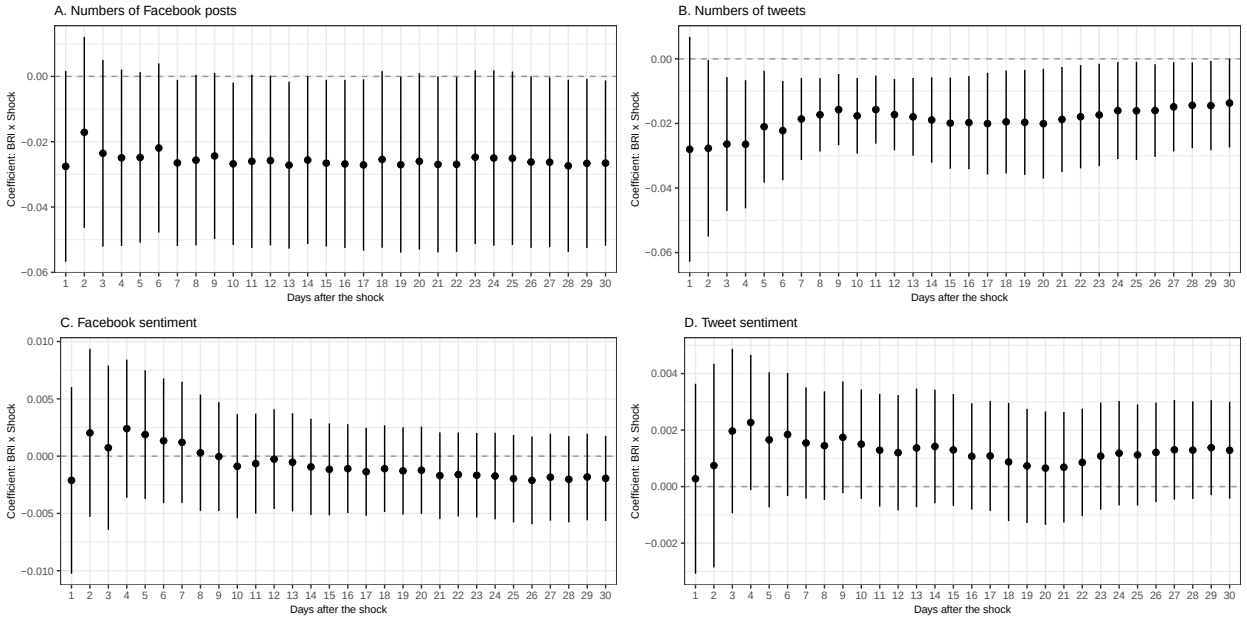
(c) Indonesia



Note: The figures show the stacked event study of the South China Sea shocks. In this placebo test, the event date is the original date of the shock - 1. The dependent variable is A. the number of Facebook posts mentioning China, B. the number of tweets mentioning China, C. the Facebook sentiment toward China, and D. the tweet sentiment toward China. The solid blue line represents the coefficients on the indicator of South China Sea disputes interacted with BRI receiving status, multiplied by the length of the window. The dashed black line represents the 95% confidence intervals using robust standard errors clustered by constituency. The horizontal axis labels denote the number of days before or after the disputes, which are included as part of the dummy variable for the event. For example, in panel A, 5 refers to the number of Facebook posts between the event date and 5 days after the event date; in panel C -5 refers to the sentiment of Facebook posts between 5 days prior to the event date and the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

Figure A-43: Stacked Difference-in-Difference: Placebo Event Date = one day before the shock

(a) The Philippines



(b) Malaysia

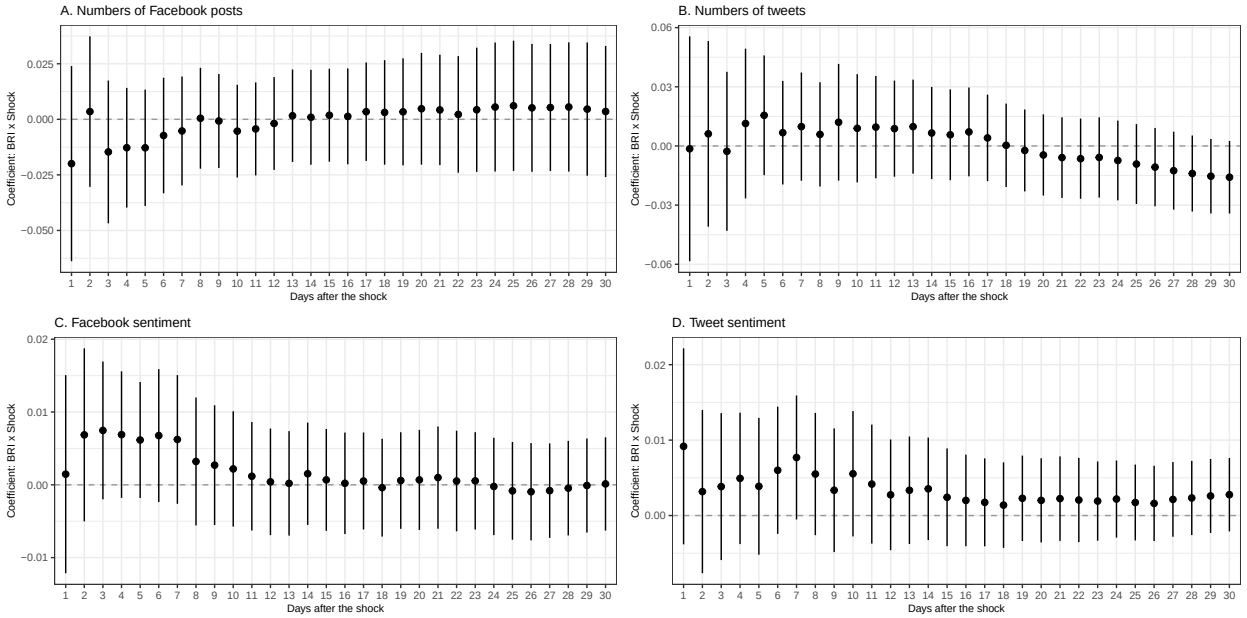
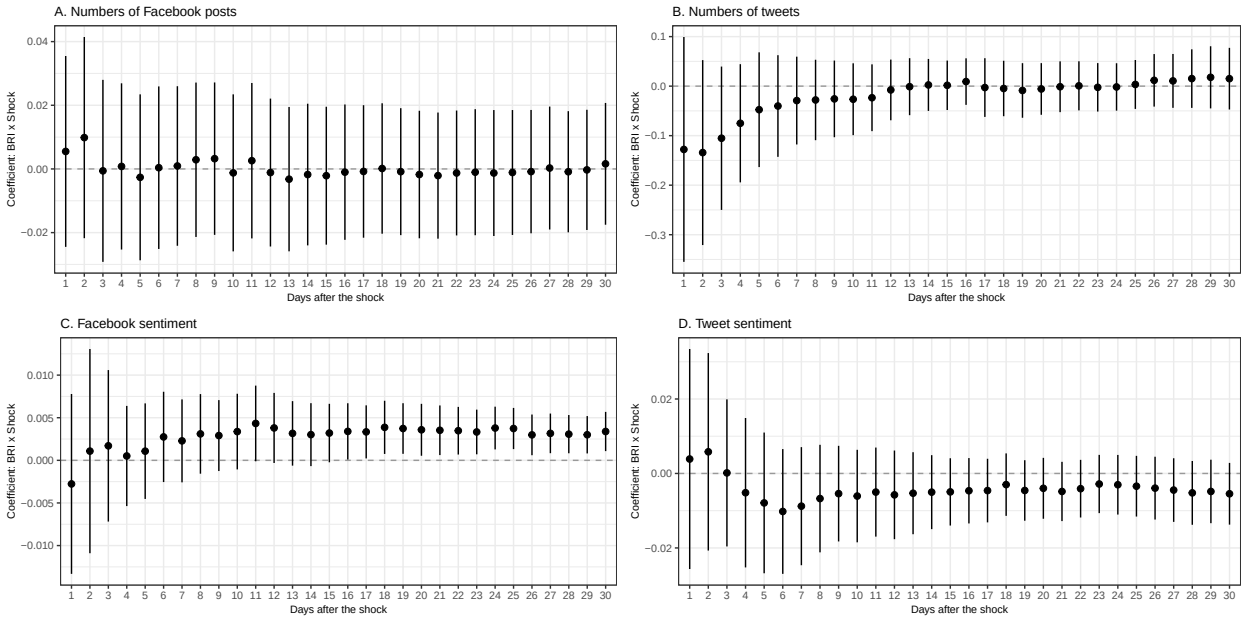


Figure A-43: Stacked Difference-in-Difference: Placebo Event Date = one day before the shock (continued)

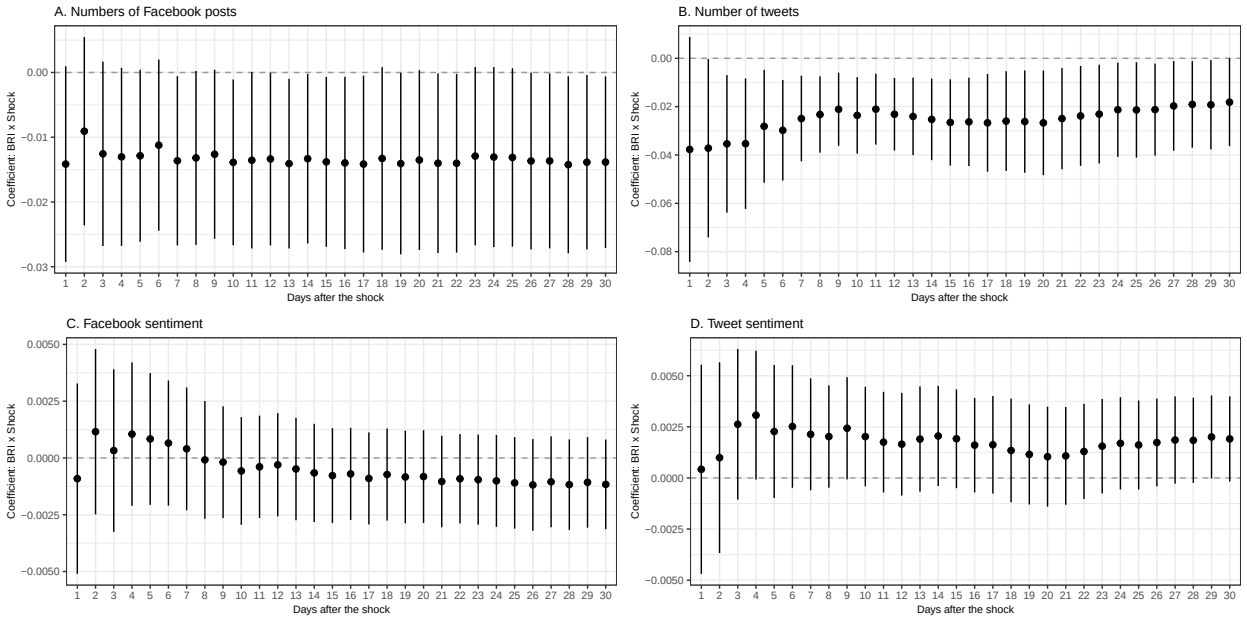
(c) Indonesia



Note: In this placebo test, the event date is the original date of the shock - 1. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

Figure A-44: Stacked Difference-in-Difference at Politician-date Level: Placebo Event Date = one day before the shock

(a) The Philippines



(b) Malaysia

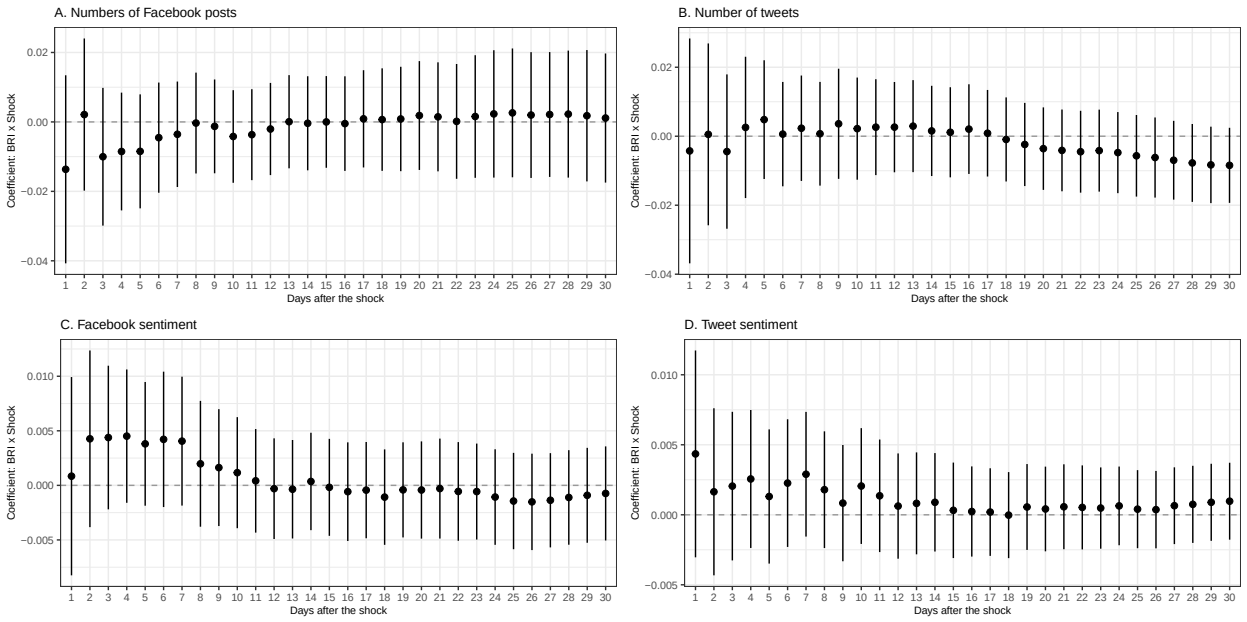
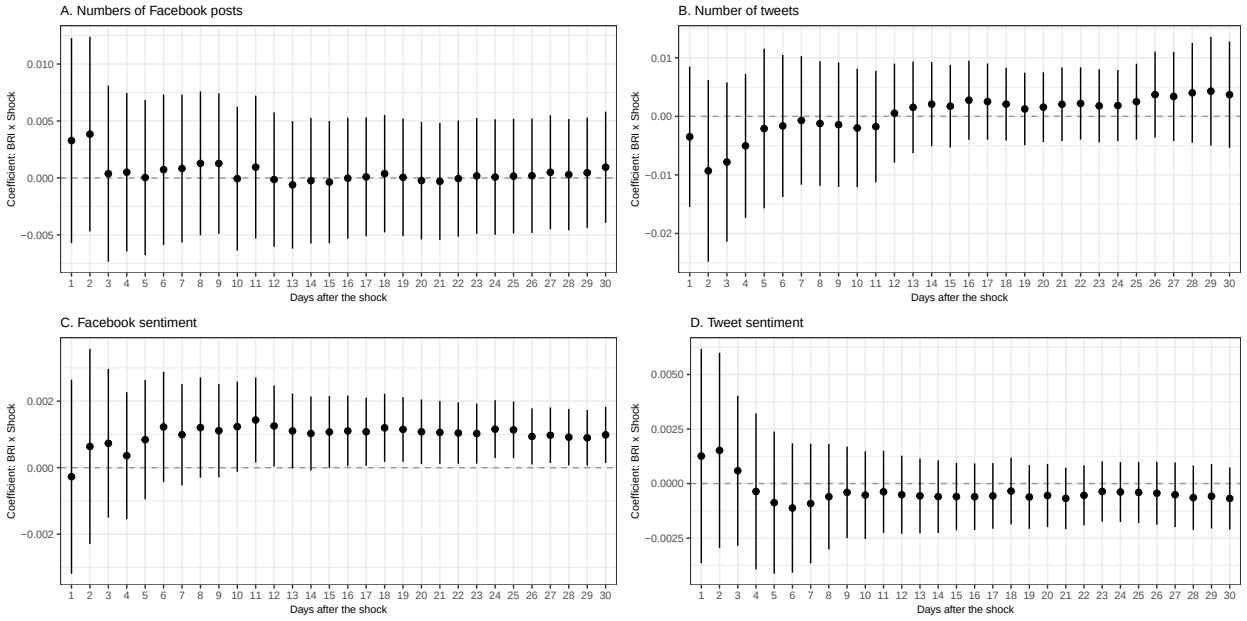


Figure A-44: Stacked Event Study: Placebo Event Date = one day before the shock (continued)

(c) Indonesia

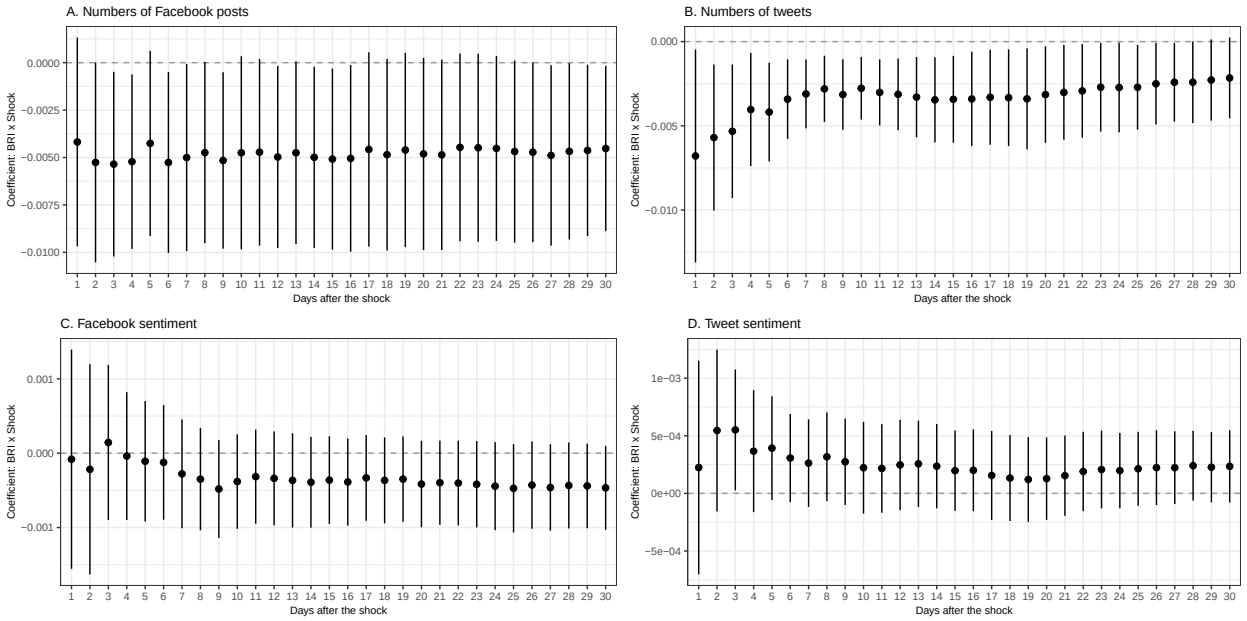


Note: In this placebo test, the event date is the original date of the shock - 1. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

4.4 Alternative Measurement of Chinese Investment

Figure A-45: Stacked Difference-in-Difference at Constituency-date Level: Logged Cumulative Value of Chinese Projects

(a) The Philippines



(b) Malaysia

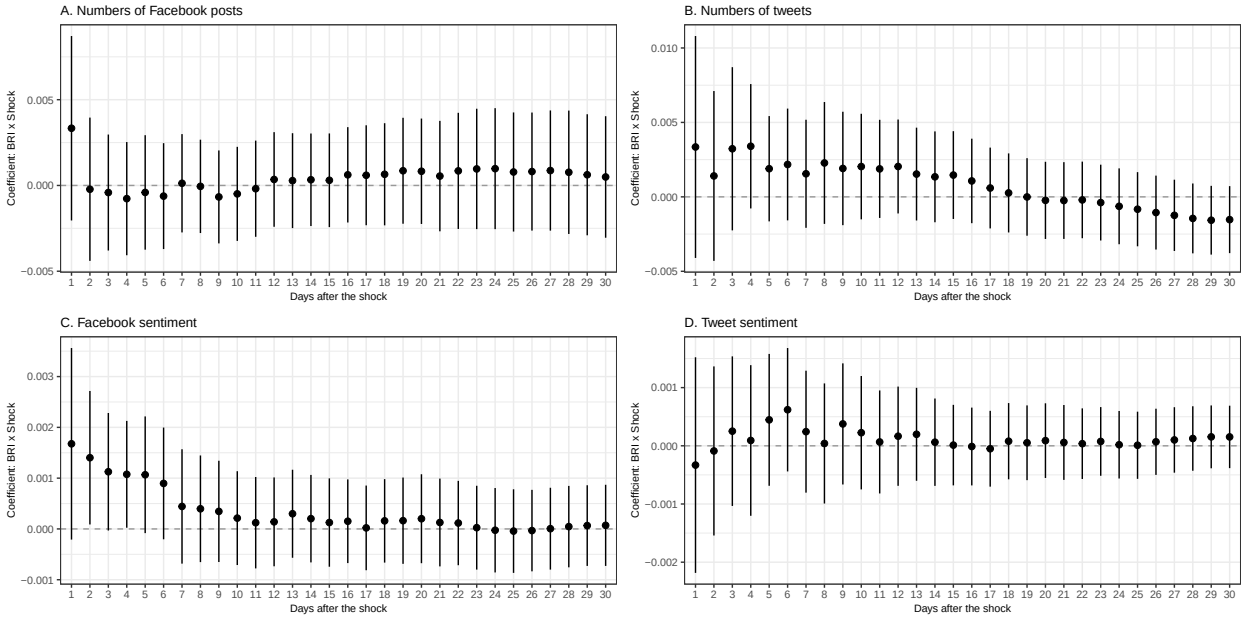
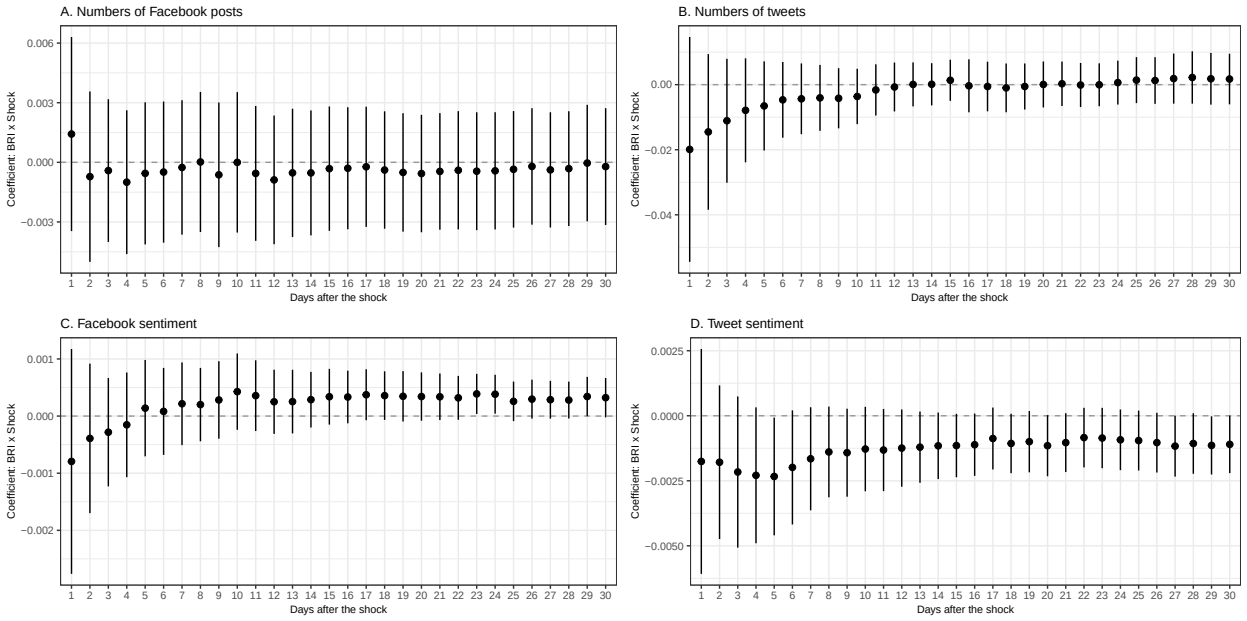


Figure A-45: Stacked Difference-in-Difference at Constituency-date Level: Logged Cumulative Value of Chinese Projects (continued)

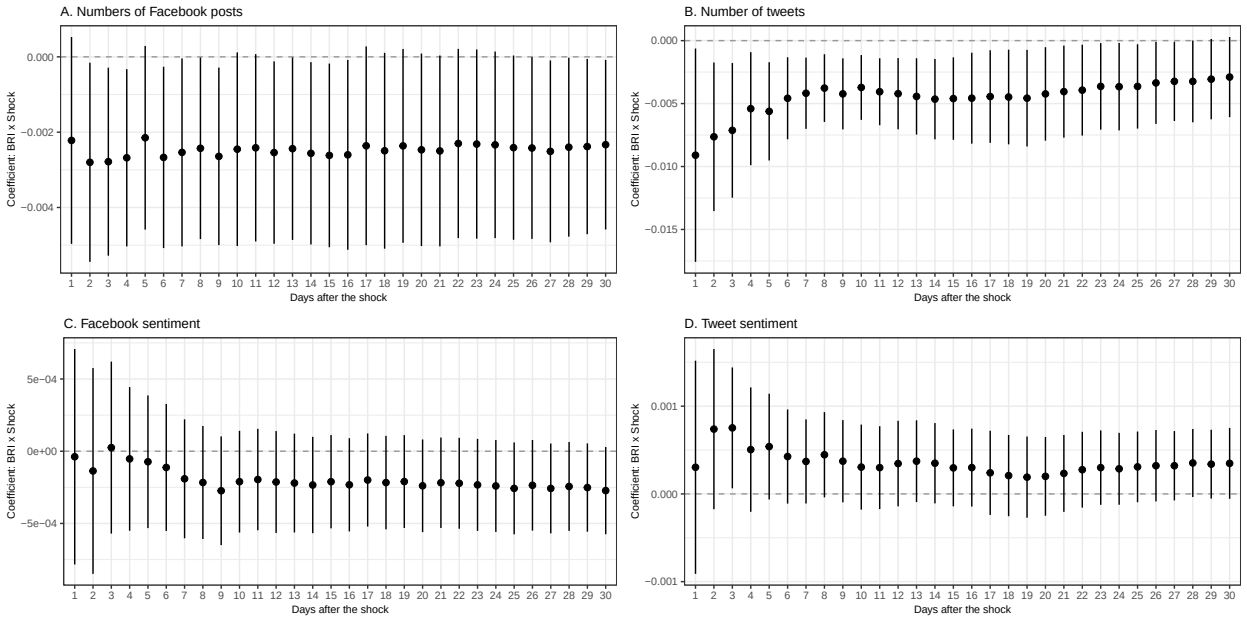
(c) Indonesia



Note: The treatment variable is log(cumulative value of BRI projects in US dollars + 1) at the constituency-date level. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-46: Stacked Difference-in-Difference at Politician-date Level: Logged Cumulative Value of Chinese Projects

(a) The Philippines



(b) Malaysia

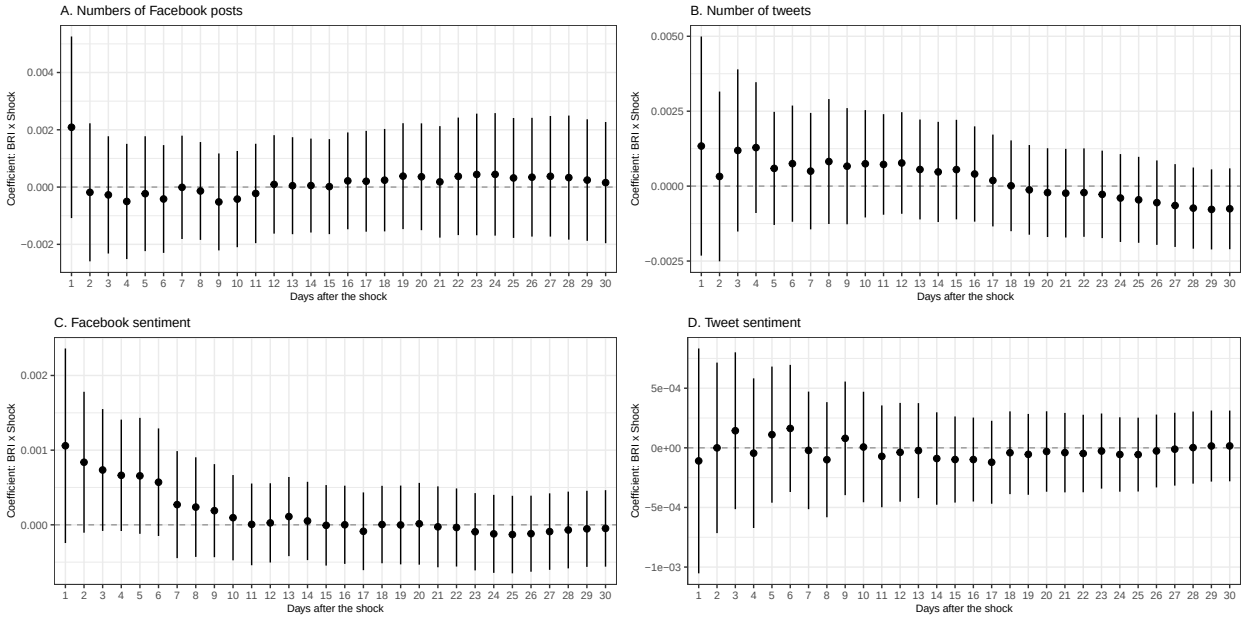
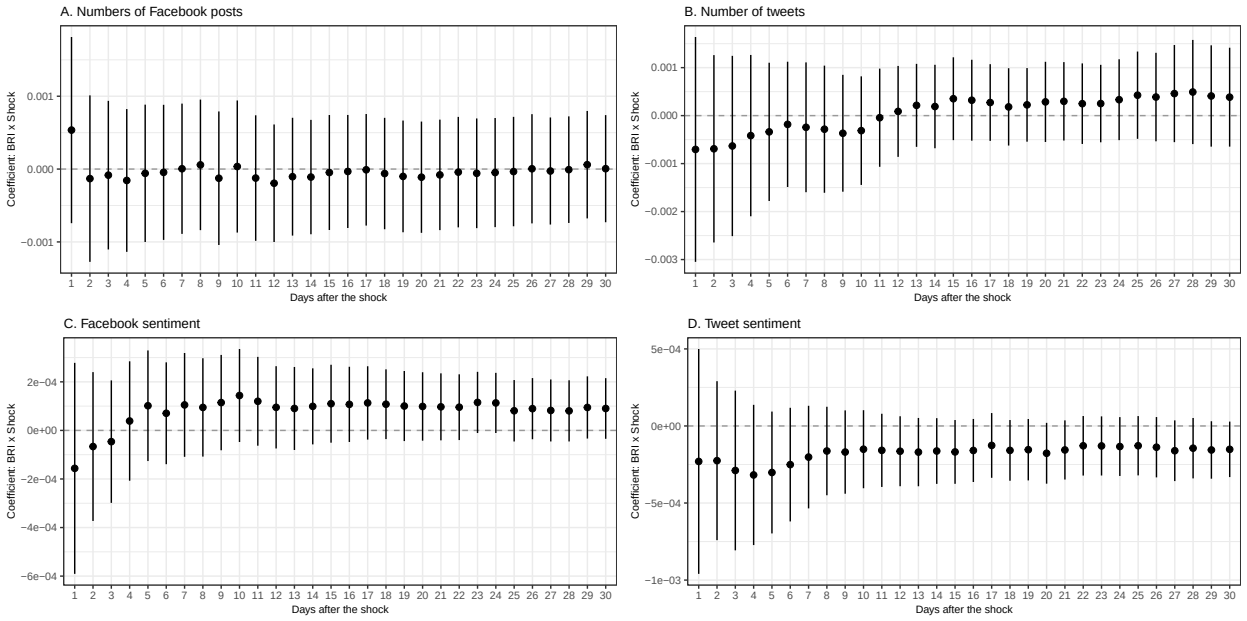


Figure A-46: Stacked Difference-in-Difference at Politician-date Level: Logged Cumulative Value of Chinese Projects (continued)

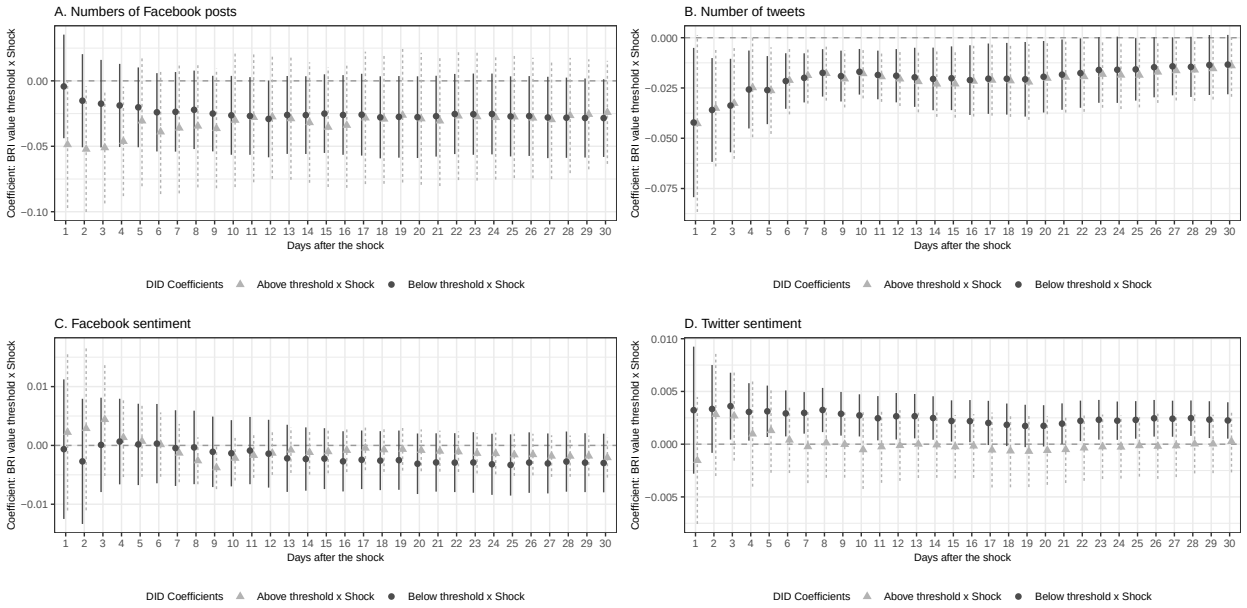
(c) Indonesia



Note: The treatment variable is log(cumulative value of BRI projects in US dollars + 1) matched to politician-date level. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

Figure A-47: Stacked Difference-in-Difference at Constituency-date Level: BRI Value Threshold of Chinese Projects

(a) The Philippines



(b) Malaysia

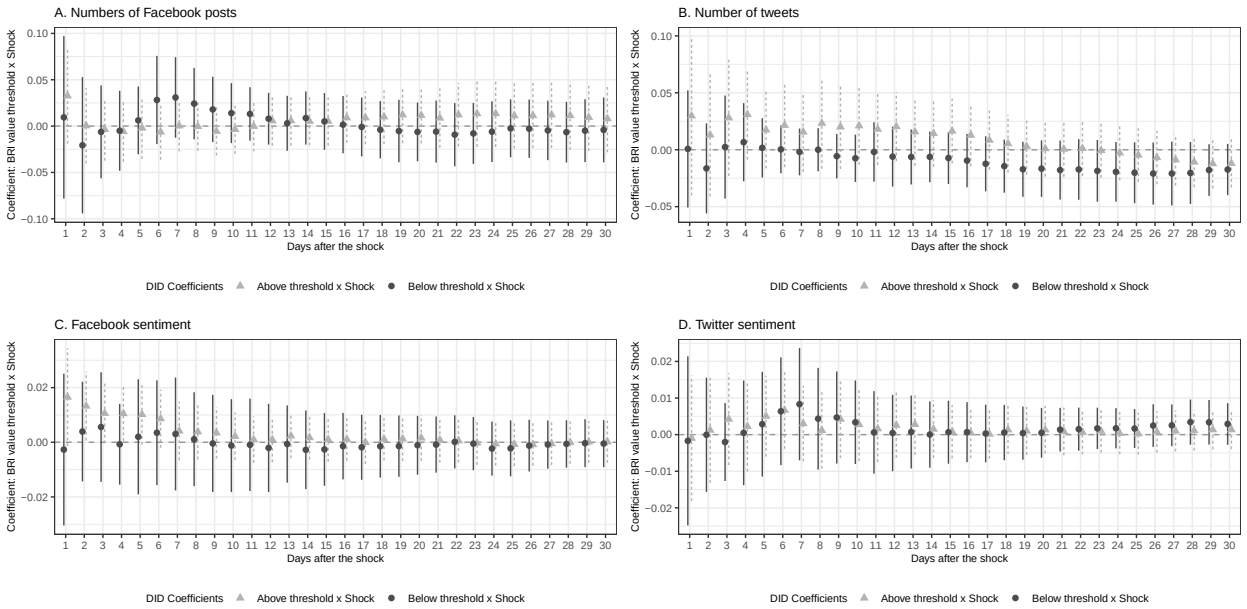
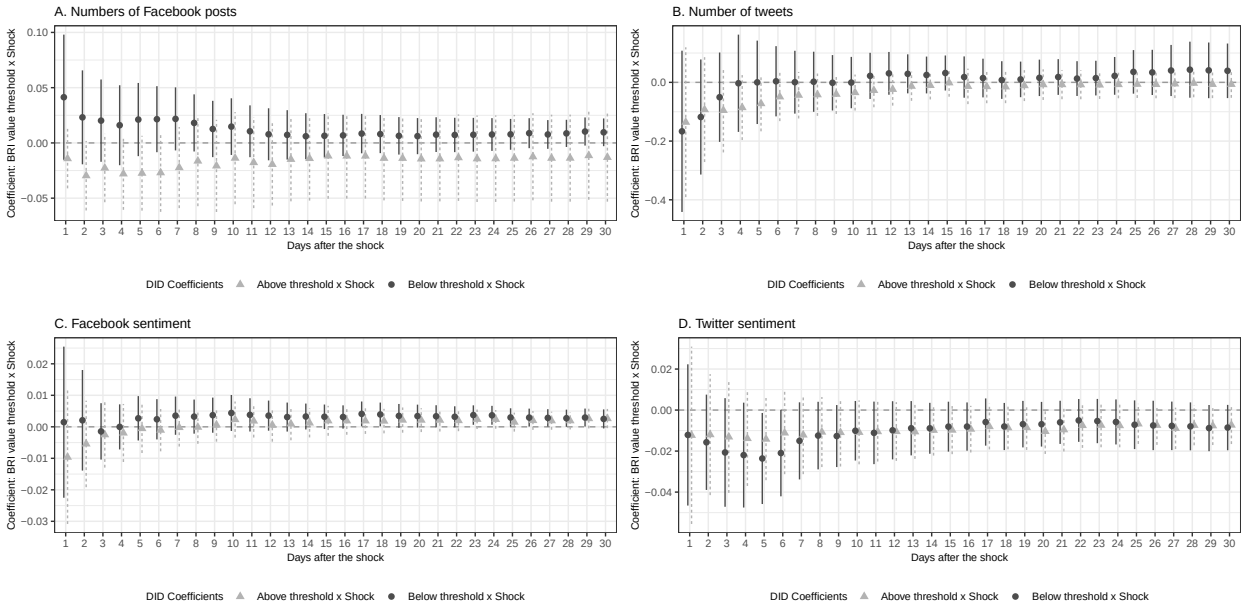


Figure A-47: Stacked Difference-in-Difference at Constituency-date Level: BRI Value Threshold of Chinese Projects (continued)

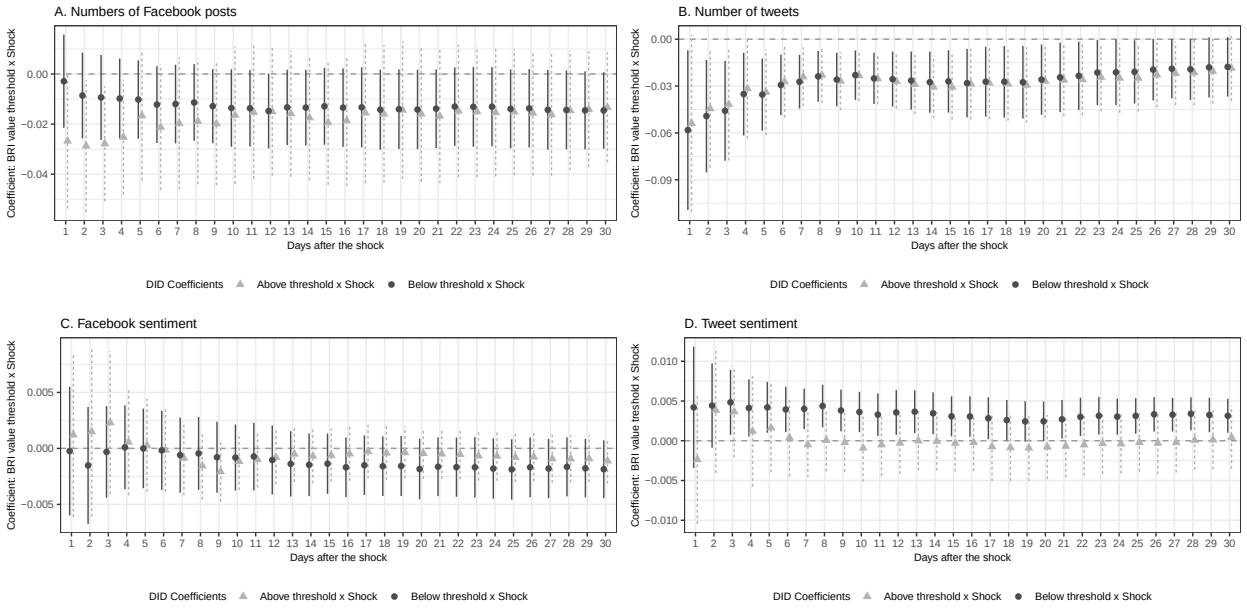
(c) Indonesia



Note: The treatment variable is a binary variable of BRI project value, which equals 1 if the cumulative value of BRI projects in the constituency is larger than the average value of BRI projects in the country and equals 0 if smaller than the average value of BRI projects in the country. The control group in the graphs (dropped out of the model) is constituencies without BRI projects. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

Figure A-48: Stacked Difference-in-Difference at Politician-date Level: BRI Value Threshold of Chinese Projects

(a) The Philippines



(b) Malaysia

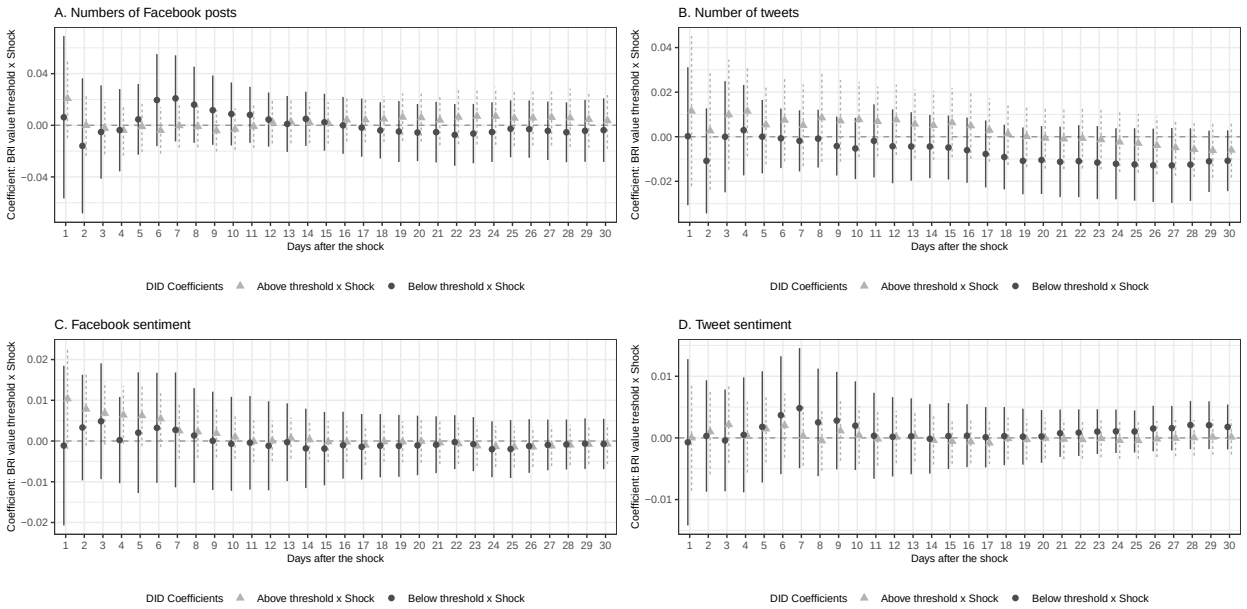
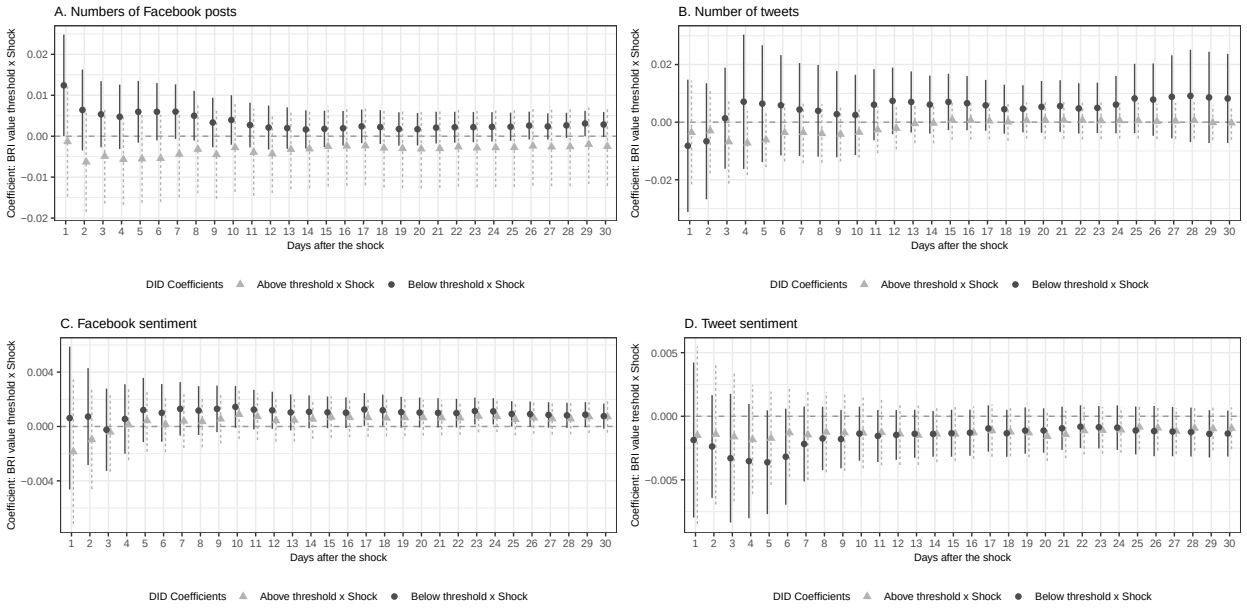


Figure A-48: Stacked Difference-in-Difference at Politician-date Level: BRI Value Threshold of Chinese Projects (continued)

(c) Indonesia



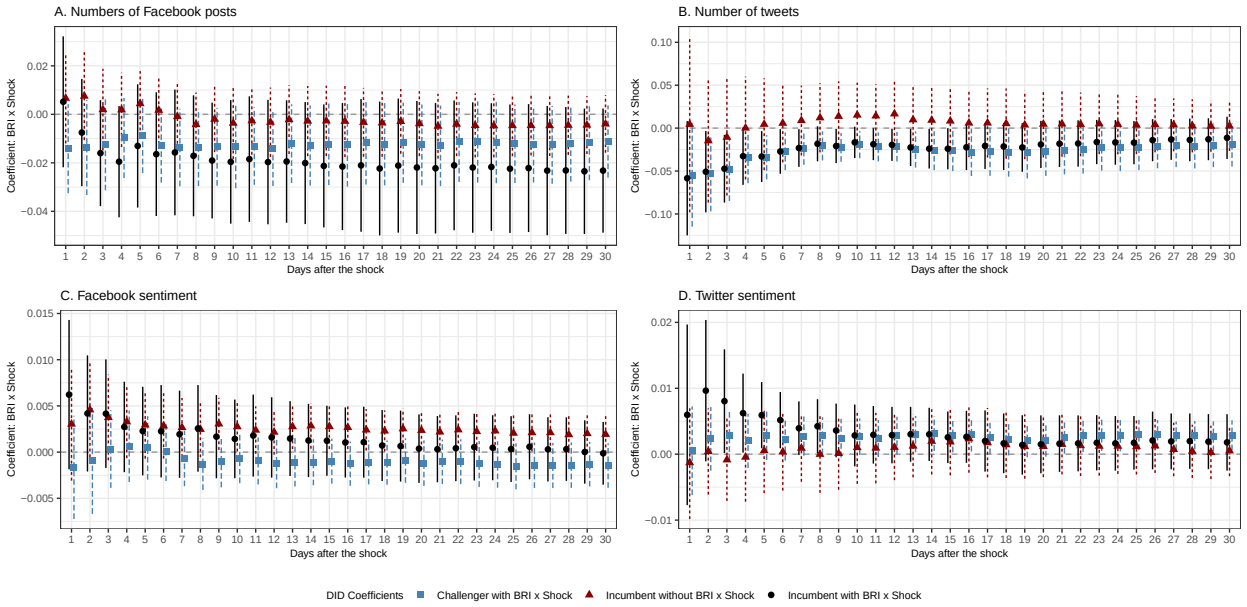
Note: The treatment variable is a binary variable of BRI project value, which equals 1 if the cumulative value of BRI projects in politicians' constituencies is larger than the average value of BRI projects in the country and equals 0 if smaller than the average value of BRI projects in the country. The control group in the graphs (dropped out of the model) is politicians of constituencies without BRI projects. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

5 Heterogeneous Effects

5.1 Incumbent vs. Challengers

Figure A-49: Incumbent vs. Challengers: Politician-date level

(a) The Philippines



(b) Malaysia

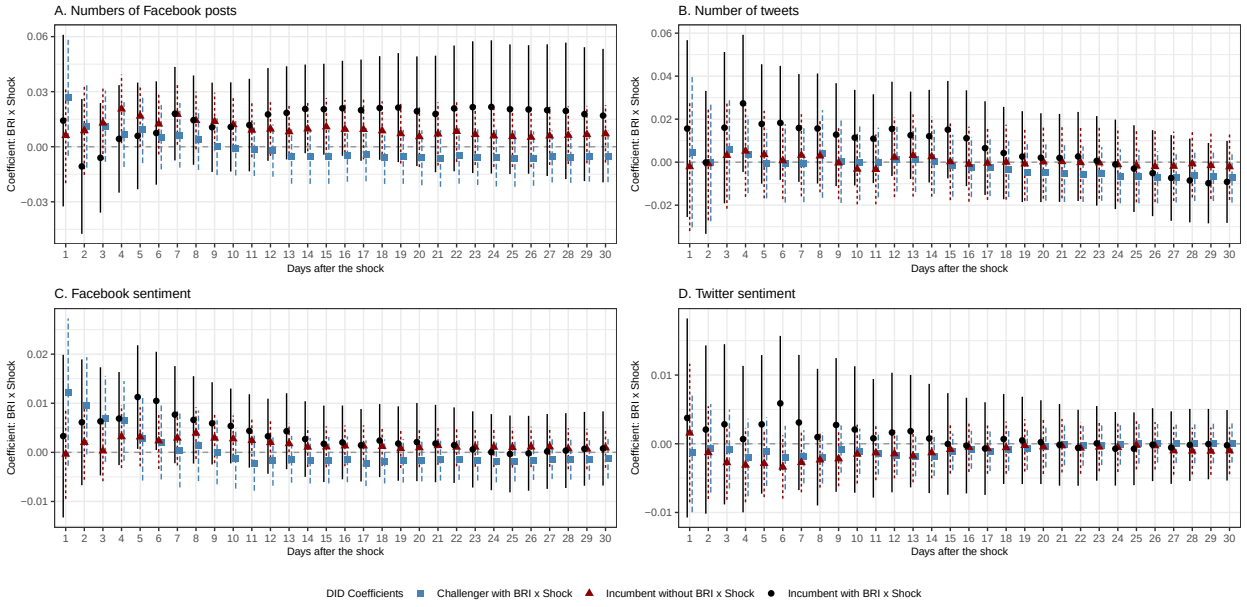
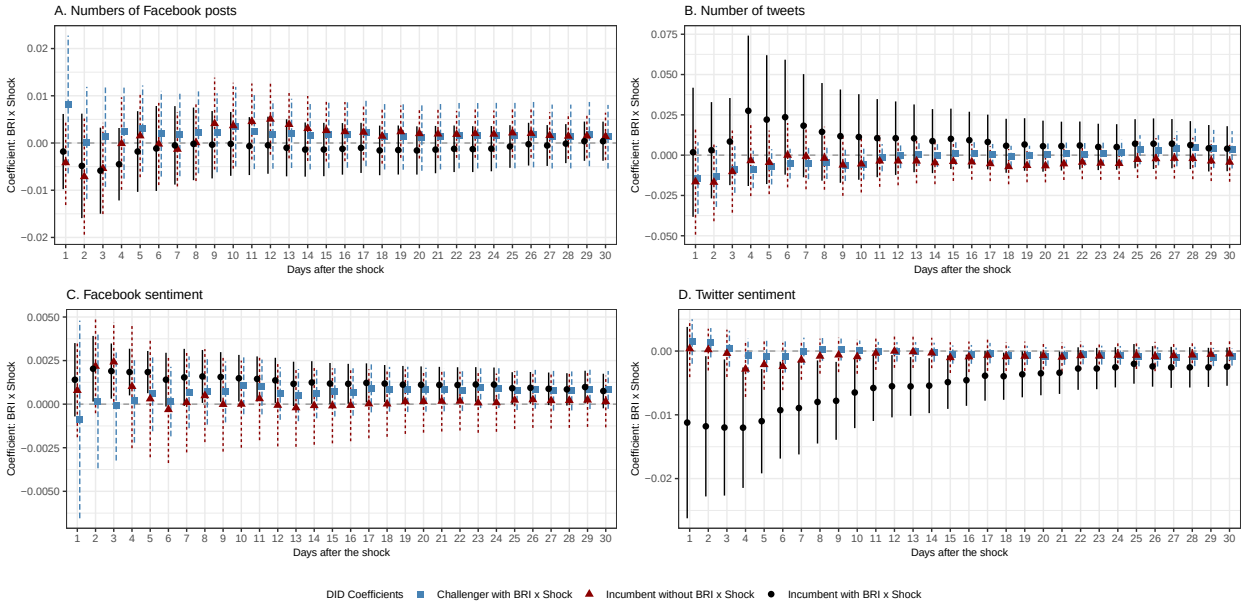


Figure A-49: Incumbent vs. Challengers: Politician-date level (continued)

(c) Indonesia

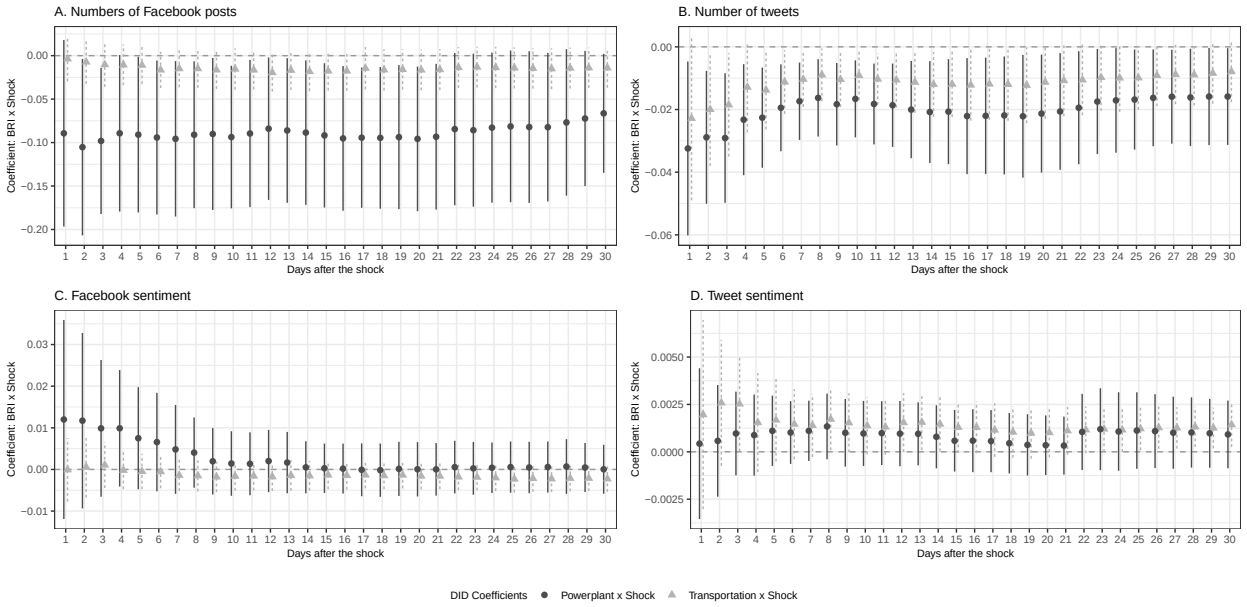


Note: Graphs illustrate the heterogeneity effect of politicians' incumbency status while posting on social media platforms. The control group (dropped out of the model) is politicians who lost elections in constituencies without BRI projects. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the politician-fixed effect and the time-fixed effect.

5.2 Public Transportation vs. Powerplants

Figure A-50: Public Transportation vs. Powerplants: Constituency-date level

(a) The Philippines



(b) Malaysia

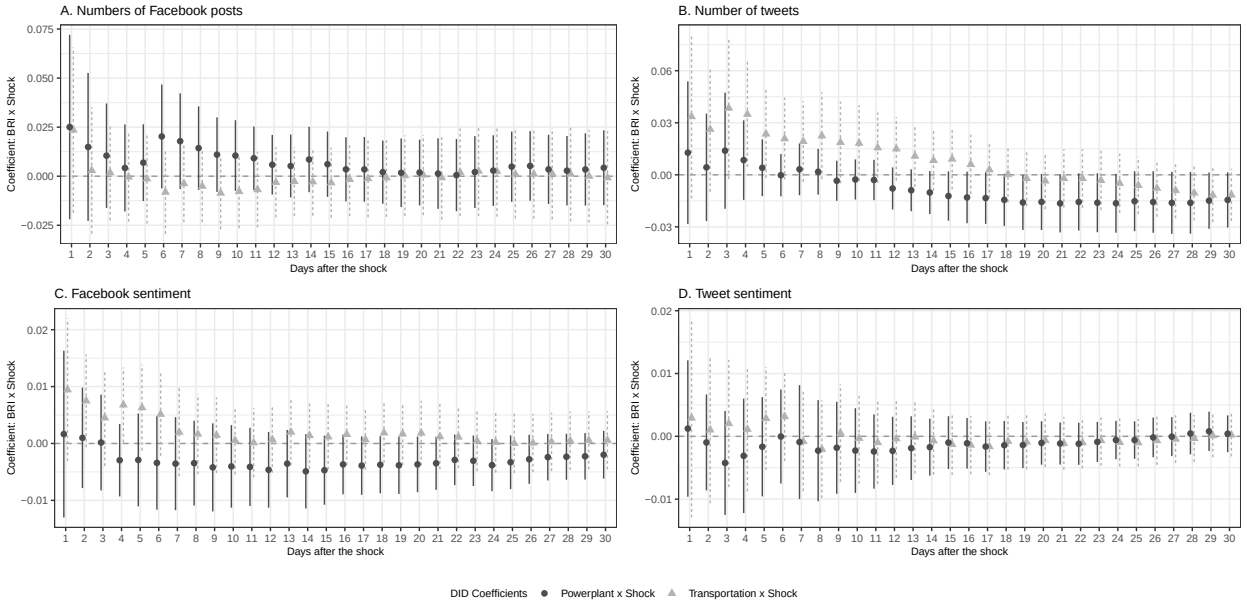
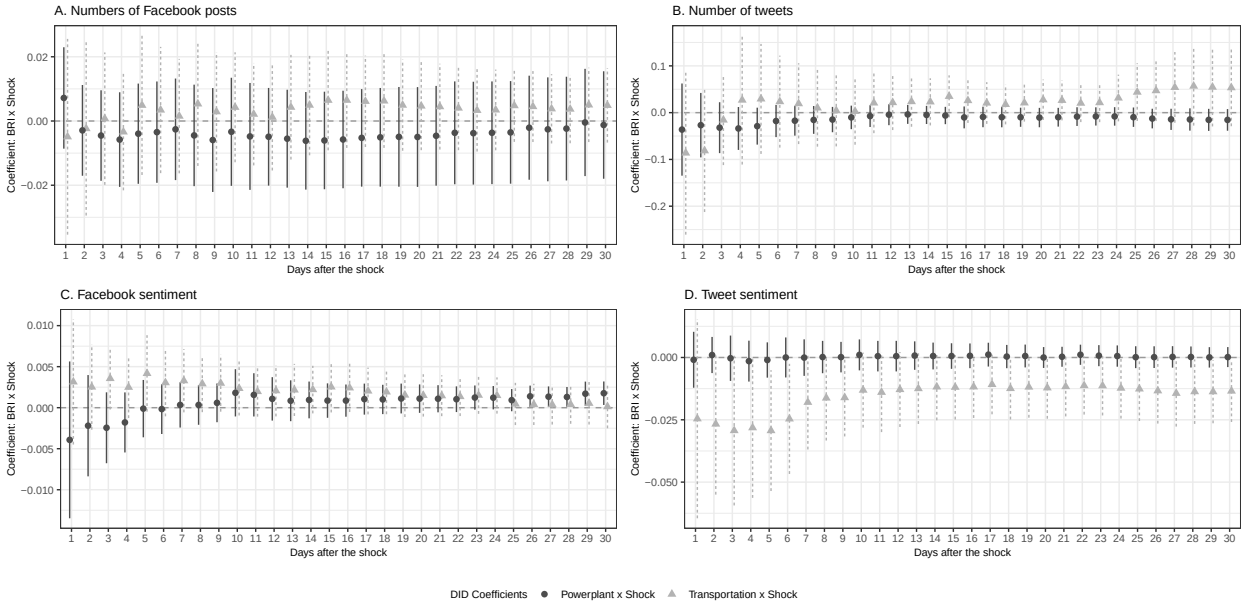


Figure A-50: Public Transportation vs. Powerplants: Constituency-date level (continued)

(c) Indonesia



Note: Graphs illustrate the heterogeneity effect of BRI project types. The control group (dropped out of the model) is constituencies without BRI projects. Dots represent the coefficients from stacked difference-in-difference regressions of outcome variables on A. number of Facebook posts mentioning China, B. number of tweets mentioning China, C. sentiment of Facebook posts about China, and D. sentiment of tweets about China. Vertical lines refer to the 95% confidence intervals for each coefficient using robust standard errors clustered at the constituency level. Pre-window is set to 30 days before the shock and does not vary. The horizontal axis labels denote the number of days after the shock, which varies from 1 day to 30 days post-event. For example, 5 refers to the *averaged* outcome variables *within the 5 days after the event date* compared to the *averaged* outcome variables within the 30 days before the event date. All regressions control for the constituency-fixed effect and the time-fixed effect.

6 Power Analysis

Figure A-51: Power Plots: Minimum Detectable Effect

(a) The Philippines

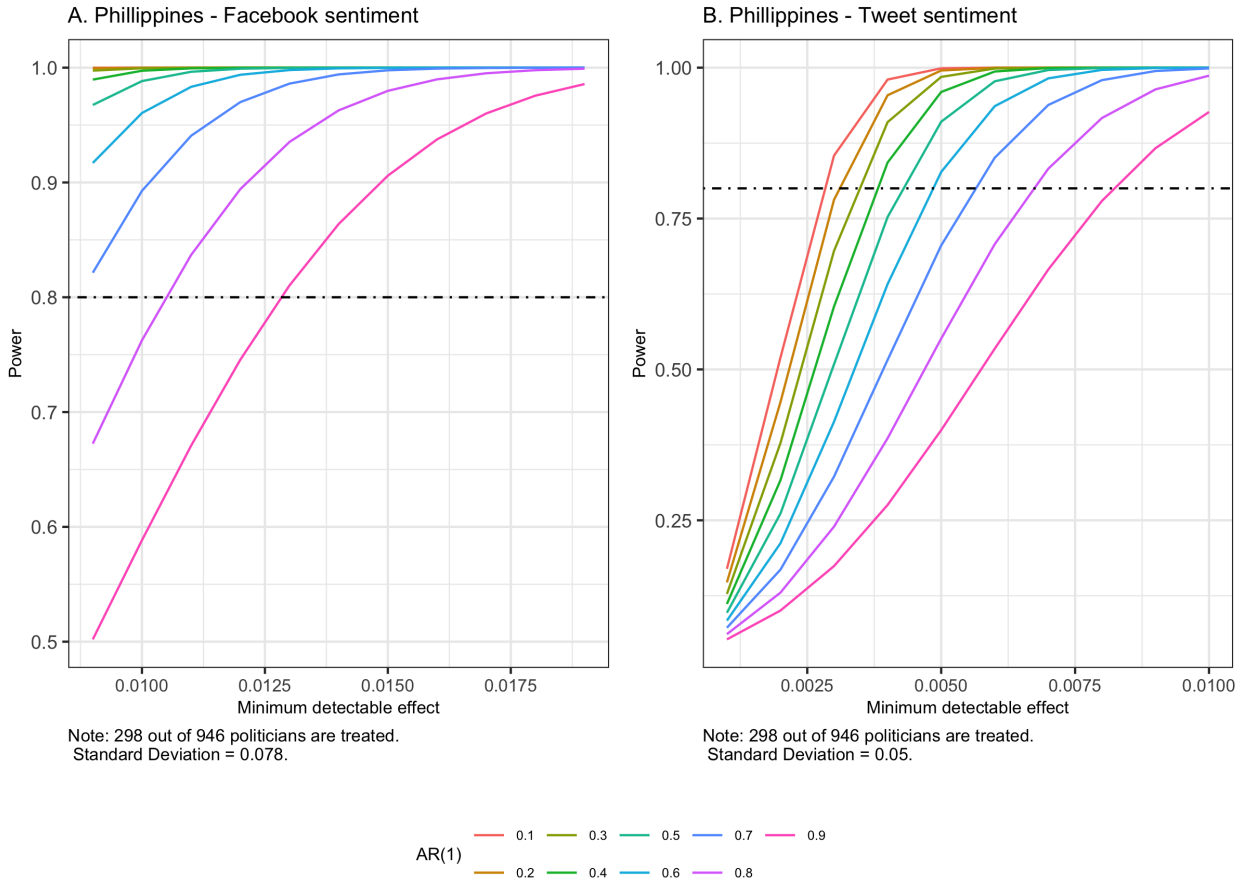


Figure A-51: Power Plots: Minimum Detectable Effect (continued)

(b) Malaysia

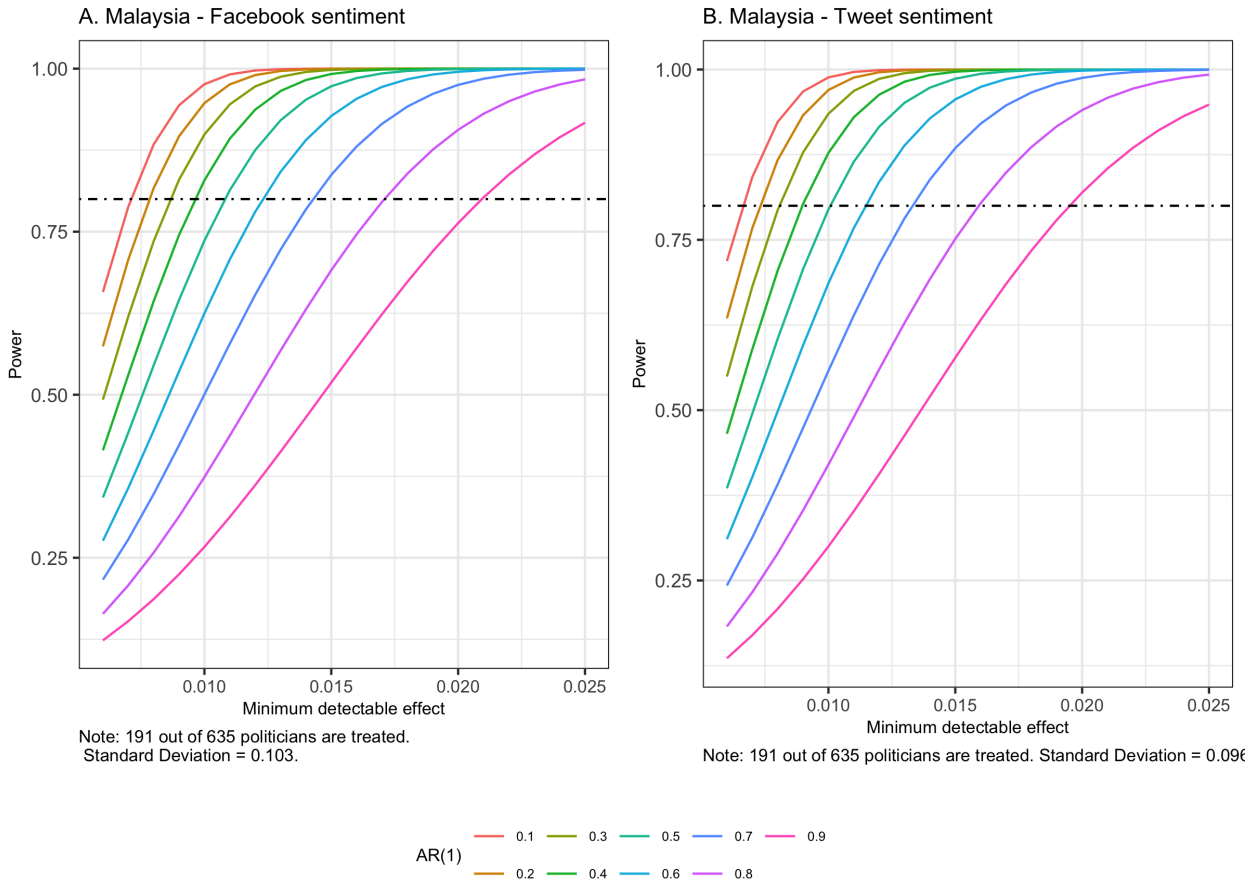
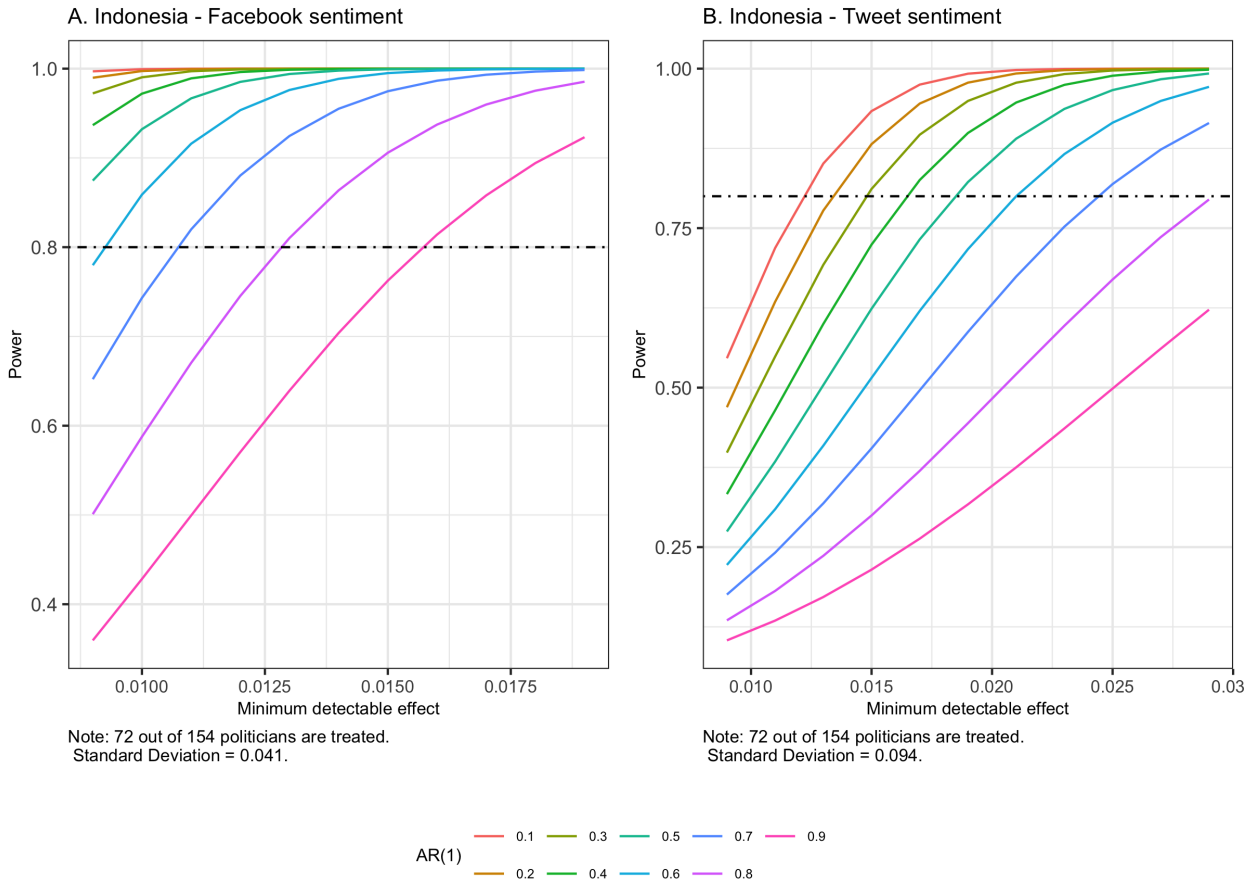


Figure A-51: Power Plots: Minimum Detectable Effect (continued)

(c) Indonesia



Note: Powers are obtained with *pcpanel* package in Stata, the analytical power calculations for panel difference-in-difference experiments with serially correlated error structures. 1) Power calculations assume Type-I error rate of alpha=.05. Errors are assumed to follow an AR(1) process (may poorly approximate more complex covariance structures). 2) Both pre-treatment and post-treatment periods are 30 days.

References

Milam Aiken et al. An updated evaluation of google translate accuracy. *Studies in linguistics and literature*, 3(3):253–260, 2019.

Center for Strategic and International Studies. Reconnecting Asia Project Database. `https://reconasia.csis.org/reconnecting-asia-map/`, 2020. [Last accessed: September 29, 2022].

Axel Dreher, Andreas Fuchs, Bradley Parks, Austin Strange, and Michael J Tierney. *Banking on Beijing: The Aims and Impacts of China's Overseas Development Program*. Cambridge University Press, 2022.

Ken Kollman, Allen Hicken, Daniele Caramani, David Backer, and David Lublin. *Constituency-Level Elections Archive*. Ann Arbor, MI: Center for Political Studies, University of Michigan, 2019. `http://www.electiondataarchive.org`.

Rebecca Ray, Kevin P Gallagher, William Kring, Joshua Pitts, and B Alexander Simmons. Geolocated dataset of chinese overseas development finance. *Scientific Data*, 8(1):1–12, 2021.

Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, et al. Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*, 2016.