

Appendix A: Facebook impressions dataset

Impressions

The API gives the lower and upper bound of impressions; we used the mean of the two bounds. When there was no upper bound (over 1 million impressions) we used 1.25 million impressions instead. This is the mean between the actual lower bound of 1 million and a derived upper bound of 1.5 million, extrapolated from the increments between preceding upper bounds. If an ad was shown over multiple days, we assumed the ad received a uniform number of Impressions on each day over its lifetime.

Pages

The table in this section lists the Chinese state media Facebook pages from which we collected advertisements. We identified the Facebook pages labeled as China state-controlled media and then identified several related pages without such labels using the Ads Library API. Most of the related pages represent localizations of another page’s content in a different language. We considered all of the pages in the table to be Chinese state media pages for the purposes of our experiments, even if they were not officially labeled as such by the Facebook Ads Library.

In Table 1, the first column lists the name of the Facebook page. The second column shows whether or not the Facebook Ads Library lists the page as being state-controlled, as of December 2020. Because Facebook updates these labels on a rolling basis, some pages are currently labeled as state-controlled in the Facebook Ads Library but were not labeled during the study period. Advertisements from these pages would have appeared without labels during the study period, even though *new* advertisements purchased today from these pages would appear with labels. The table below lists each page as it appeared during the study period to make clear whether the page belongs to the treated (label added on June 4) or untreated (no label from 2018 to 2020) group in our study of labels.

We only list pages that have posted advertisements that are available in the Facebook Ads Library during our study period from 2018 to 2020. Pages owned by China’s local governments (e.g., Shanghai United Media Group) were excluded.

Table 1. List of Facebook pages used in this study.

Facebook Page Name	Labeled as State-Controlled in Facebook Ads Library?	Page ID Number
China Xinhua News	Yes	338109312883186
China Xinhua Sci-Tech	Yes	1175630379148974
China Xinhua News (French)	Yes	1453950118203438
China Xinhua News (Spanish)	Yes	484016595013939
China Xinhua News (Myanmar)	Yes	422671591235160
China Xinhua News (Portuguese)	Yes	1629228347341832
China Xinhua News (Swahili)	Yes	844193555664475

China Daily	Yes	191347651290
People's Daily	Yes	188625661189259
China Daily Hong Kong	Yes	161095810581910
People's Daily App	No	2002267650037647
China.org.cn	Yes	371171589575669
China.org.cn (French)	No	1448903671990351
China.org.cn (Russian)	No	152362801582819
China.org.cn (German)	No	295446797213542
China.org.cn (Japanese)	No	526327500799853
CGTN	Yes	565225540184937
CGTN America	Yes	300653326696613
CGTN Français	Yes	169799706487474
CGTN en Español	Yes	227622537655549
CGTN Europe	No	1798882800328461
CGTN App	No	234535423558415
CCTV	Yes	188872144758
CCTV+	No	756877521031964
CCTV (Arabic)	Yes	578470445572869
CCTV (Spanish)	Yes	611178888961399
CCTV 中文	Yes	761870613900500
CCTV (French)	Yes	313374855482172
CCTV (Russian)	Yes	278778585525815
CCTV (Korean)	Yes	718077228263799
Global Times	Yes	115591005188475
China Plus News	Yes	223495844457800
Discover China	Yes	678954262167350
Indonesian Service, China Radio International	No	1414515485454171
CRI – Radio Internacional da China	No	1518216038506999
Radio Cina Internazionale	No	299227550122220
Radio Internacional de China en Espanol	No	153955617984812
CRIarabic*	Yes (currently)	232678173543084
CRI日本語*	Yes (currently)	488965394532508
China Plus America*	Yes (currently)	539149559464605
Xinhua Culture & Travel*	Yes (currently)	535503280269676
CGTN Global Business*	Yes (currently)	837157493122072
Fly Over China	No	1006829009506673
Belt and Road News Network	No	405199520209766
XinhuaHK	Yes	1798435720478073
CRI German	No	340577086006224
Radio Chine Internationale	No	438443612916112
China Mosaic by China.org.cn	No	1877361012493253

Note: Some pages have identical names but post content in different languages. In these cases, we list the language in parentheses to distinguish the two pages.

Pages marked with an asterisk were excluded from the analysis of Facebook's labels because we were unable to determine whether they initially appeared with or without labels. The second column indicates whether they are currently labeled as state-controlled.

Cross-platform content

For each advertisement, human coders manually searched for the advertisement text on other platforms, either using Google or following hyperlinks contained in the advertisement text (which typically led to state media articles). Text was considered identical if an exact or near-exact match was found for two or more sentences (or the full text of the advertisement or cross-platform content, if one was shorter). If the identical text appeared on a state-media platform other than Facebook, the advertisement was marked as having nearly identical content on another platform, and the publication date of that content was recorded. For each ad, we only aimed to identify whether nearly identical content existed, not to provide an exhaustive list of all such content on all platforms. Therefore, only one piece of cross-platform content was included for each advertisement; if there were multiple matches, only the state media article was listed.

Some cross-platform content that was initially available when the Facebook advertisement was originally posted may have been unavailable during data collection. For instance, state media accounts can delete tweets or news articles after posting them. There were four instances in the dataset where the advertisement contained a broken hyperlink. As such, the figures in Finding 1 on the proportion of advertisements linked to off-platform content should be taken as a lower bound rather than an exact estimate. The advertisements with annotations can be found in the Harvard Dataverse.