

Media Patronage: Selection and Survival to Shape Foreign Media

Matt DeButts*

Jennifer Pan†

April 16, 2026

Abstract

Authoritarian regimes often seek to influence media beyond their borders, but applying transnational repression can be politically and economically costly. We argue that the financial strain facing media organizations worldwide has created a low-cost alternative: “media patronage,” the selective support of outlets whose coverage already slants toward the regime. Media patronage does not seek to change individual journalists or outlets but operates at the ecosystem level, increasing the share of pro-regime media over time. We document the Chinese Communist Party’s use of this strategy in global Chinese-language digital news, analyzing 14 million articles from 193 outlets over 25 years. We find that outlets with more positive coverage of China and less coverage of censored topics are selected for patronage. After selection, outlets maintain their pro-regime slant while producing local and global coverage resembling that of unaffiliated competitors. Patronage helps outlets survive, increasing web traffic by 27% and reducing the likelihood of closure by 57% relative to unaffiliated peers. Finally, we show overseas patronage has domestic payoffs: during China’s 2022 Zero-COVID protests, the absence of protest coverage from patronage outlets meant protest narratives did not dominate the flow of diaspora media back into China’s domestic information environment. These results contribute to research on transnational repression, political influence on media, and authoritarian politics.

Keywords: Media capture, information manipulation, transnational repression, authoritarian influence, China, diaspora

PRELIMINARY, PLEASE DO NOT CIRCULATE

*Ph.D. student, Department of Communication, Stanford University, mdebutts@stanford.edu

†Professor, Department of Communication, Stanford University, jp1@stanford.edu

Authoritarian regimes often seek to influence media production outside their borders. These efforts have varied aims—e.g., to discredit opponents, stop transnational movements, bolster domestic propaganda, or influence foreign elites and publics (Huang 2015; Huang and Yeh 2019; Kern and Hainmueller 2009; Kern 2011; Nugent and Siegel 2026; Peisakhin and Rozenas 2018; Repnikova 2022)—but are hindered by territorial boundaries. Inside their borders, autocrats have numerous institutional tools from regulatory agencies to the police to control media (Lucas 2022; McMillan and Zoido 2004; Rahmani 2025; Rosenfeld and Wallace 2024; Stockmann and Gallagher 2011). Outside state borders, foreign sovereignty and geographic distance can hinder attempts at control.

Existing research has focused on how authoritarian regimes can influence foreign media outlets with transnational repression. These strategies range from exploiting international legal institutions like Interpol to arrest journalists, to digital harassment, to applying relational pressure such as the detention of family members who remain in-country (Deng and O'Brien 2013; Dukalskis et al. 2022; Earl, Maher and Pan 2022; Hassan, Mattingly and Nugent 2022; Kao 2026; Moss 2016; Tsourapas 2021). Research has also shown how authoritarian regimes try to influence foreign information environments by directly founding and operating state-controlled media organizations abroad (Cull 2009; Fan, Pan and Sheng 2023). These strategies, while demonstrably effective in many contexts, carry substantial political and financial costs.

In this paper, we argue that *media patronage* is an alternative strategy for authoritarian regimes seeking to influence foreign media—one that is facilitated by the financial strain faced by media organizations worldwide. The advent of the internet and the shift to online advertising has decimated news media revenues, leading to reductions in staffing, declining coverage diversity, and the closure of thousands of news organizations around the world (Crain 2009; Nielsen 2015; Willnat and Martin 2020). In democracies and lower capacity states, where media organizations have historically relied on market revenues rather than domestic government funding (Hamilton 2004), there is increasing evidence that media or-

ganizations under financial strain are susceptible to influence by political interests (Besley and Prat 2006; Tella and Franceschelli 2011).

Media patronage is the selective provision of support to outlets whose coverage already slants toward the patron while shunning critical ones. Rather than changing reporting at the journalist or outlet level, media patronage operates on the ecosystem level, reshaping the competitive landscape of media markets (Weber and Monge 2017): it helps friendly media organizations crowd out critical ones in media markets, ensuring that pro-patron outlets increase their market share over time. This strategy, which can be pursued alongside the strengthening of state-run broadcasters that cover global or national affairs as well as targeted repression, helps to fill the void left by the decline in local journalism while advancing the agendas of authoritarian regimes.

We find empirical evidence of media patronage in the case of the Chinese Communist Party (CCP) and global Chinese-language digital news. By analyzing 14 million articles produced by 193 Chinese-language digital news outlets outside of mainland China as well as a census of all Chinese-language news outlets in the United States over a 25 year period, we find that the CCP is selectively supporting news outlets with a pre-existing slant toward China, but that these organizations do not change their coverage after securing CCP support. Rather, these outlets compete with unaffiliated outlets for market share by providing similar local and global news coverage while enjoying access to resources enabled by their linkage to the CCP. Patronage increases web traffic to patronage websites by 27% and reduces the likelihood of closure by 57% relative to unaffiliated peers. Finally, we find that the regime’s investment in overseas patronage has domestic payoffs during unforeseen crises: during China’s Zero-COVID (“White Paper”) protests in late 2022, patronage outlets largely declined to cover protest-related discontent even as unaffiliated outlets covered it widely, meaning that protest narratives did not dominate the flow of diaspora media back into China’s domestic information environment.

This research has implications for the study of transnational repression, media markets,

and authoritarian politics. On transnational repression, existing work has focused primarily on highly visible coercive tactics such as harassment, detention, and violence (Dukalskis et al. 2022; Moss 2016). This emphasis is driven in part by policy visibility and the attendant data availability: violent, repressive tactics draw attention, documentation, and coverage, increasing the urgency of academic research as well as facilitating its pursuit. Media patronage, by contrast, operates through market mechanisms that are often legal and routine. These transactions attract less scrutiny than assassinations or kidnappings, yet can nonetheless reshape information environments. Recognizing media patronage as a form of transnational influence—alongside overtly repressive tactics—offers a more complete picture of how authoritarian power operates beyond their borders (Earl, Maher and Pan 2022; Tsourapas 2015).

These results also contribute to a growing body of research on political influence on news media (Alonso and Padró i Miquel 2025; Besley and Prat 2006). Research on media in democracies has historically found market competition to facilitate media independence by providing media organizations with sources of revenue independent of domestic and foreign political actors (Gentzkow, Glaeser and Goldin 2006; Gentzkow and Shapiro 2008; Petrova 2011). Research on media in authoritarian and hybrid regimes shows how autocrats diminish media independence by undermining media competition, e.g., through state ownership, financial sanctions, or proxy influence (Egorov, Guriev and Sonin 2009; Guriev and Treisman 2019; McMillan and Zoido 2004; Pan 2017; Qin, Strömberg and Wu 2018; Rahmani 2025; Shirk 2010; Stockmann and Gallagher 2011). However, recent work challenges this distinction, showing how political influence can operate through the mechanisms of commercial news production, such as advertising and ownership, even in competitive markets (Alonso and Padró i Miquel 2025; Archer and Torres 2022; Grossman, Margalit and Mitts 2022). We contribute to this line of research by showing that well-resourced actors, including foreign governments, can exploit the financial vulnerability of media organizations to reshape media ecosystems in their favor.

Lastly, this paper has implications for research on authoritarian politics. First, these results are consistent with work showing how authoritarian regimes use propaganda to reinforce existing pro-regime sentiments (Shirikov 2024). As opponents disappear and supporters survive, authoritarian regimes can reshape the information environment without targeting individual changes in expression. Second, this extends the literature on how modern authoritarian regimes sustain power through media manipulation rather than overt repression (Guriev and Treisman 2019; Levitsky and Way 2010; Pan, Xu and Xu 2023) by showing that the informational strategies of modern authoritarianism are not confined to domestic information environments. Third, our finding that overseas media patronage can limit the reach of damaging foreign coverage of domestic crises among home audiences shows that transnational efforts of autocracies serve not only foreign policy goals but domestic political ones as well.

The paper proceeds as follows. In Section 1 we introduce the concept of media patronage. Section 2 describes our case of China and Chinese diasporic media organizations. Section 3 introduces the data and how we identify media patronage. The three subsequent sections present the results: Section 4 shows how outlets are selected for patronage, whether they change afterward, and how their coverage competes with unaffiliated outlet coverage; Section 5 focuses on patronage and survival; and Section 6 examines the implications of media patronage for information control within China’s domestic information environment. Section 7 concludes.

1. Media Patronage

The concept of media patronage adapts the logic of patron-client relationships, a form of mutually beneficial arrangement where a patron supports, but does not directly control, clients operating autonomously on his behalf (Scott 1972). Patron-client relationships are a pervasive feature of authoritarian politics and backsliding democracies, where those in power use selective resource allocation to build loyalty among local officials, business elites, and

other important actors (Manion 2023; Shih 2007). Despite entering into a relationship of obligation with the patron, clients retain local informational advantages, and maintain the relationship primarily due to common interests rather than the continuous threat of coercion (Kitschelt, Wilkinson et al. 2007; Munro 2010).

The presence of these common interests means that unlike strategies that seek to alter media behavior at the individual level via coercion or co-optation, media patronage operates at the ecosystem level through selection.¹ It sustains already supportive outlets to participate in an ongoing arena of political competition rather than trying to change critical ones (Nichter 2008). Media coverage is publicly observable, so patrons can infer an outlet’s orientation from its existing reporting, providing resources to supportive media outlets to help them survive financial pressures that might otherwise force their closure.²

Once media patronage is established, it operates primarily through preservation. Because the client’s interests are already aligned with the patron, it does not require altering editorial orientation but rather supports the client’s continued presence in the media ecosystem, advantaging it over competitors who do not receive such support. Over time, patronage thus secures changes in the overall market rather than at the level of each outlet.

The autonomy of the client is important as it also allows outlets to maintain sensitivity to local conditions and credibility with local audiences, which is precisely what makes them valuable to the patron. If a patronized outlet begins to only parrot regime talking points, it may lose local trust as it becomes indistinguishable from state-controlled media. If the outlet must solicit regime approval before each coverage decision, it cannot produce news with the alacrity sufficient to meet audience information needs and compete for attention.

¹Some studies use “co-optation” broadly to include what we call patronage (e.g., Kao 2026; Tsourapas 2021). We use the term patronage to draw an analytical distinction. While both patronage and co-optation entail the provision of benefits, the difference lies in *who* it is provided to rather than *what* is provided. Co-optation targets benefits to adversaries to change their behavior, while patronage selects allies whose behavior already aligns with patron interests. This distinction matters because they operate differently—co-optation leads to behavioral change at the outlet level while patronage shifts market composition at the ecosystem level.

²By selectively providing financial support to foreign media organizations whose coverage already aligns with regime interests, the interests of clients are already aligned with those of the patron—a result known as the “ally principle” (Bendor, Glazer and Hammond 2000). Because the patron and client share a common goal, the patron resolves the problem of monitoring by inferring the client’s latent preferences such that the client is presumed to act in their mutual interest in the absence of direct monitoring.

Media patronage is sustained by business transactions, whether direct transfers of resources or intermediate forms such as advertising purchases, content licensing, and content sharing. These resources aid client survival while allowing substantial autonomy in how to use them. This approach offers several advantages over alternative strategies. First, it carries lower political risk than repression. Transnational repression of media professionals can backfire when discovered. The murder of *Washington Post* columnist Jamal Khashoggi caused many foreign firms to reconsider investing in Saudi Arabia and generated substantial negative press for the Kingdom (McPherson-Smith 2021). The extraterritorial kidnapping of Hong Kong booksellers generated international condemnation and diplomatic tensions (Ramzy 2020). Although media patronage may over time acquire coercive dimensions as financial dependence deepens, this coercion is not as overt as in direct repression and is thus far less visible.

Second, media patronage may be more efficient than co-optation, which targets threats or potential threats and seeks to neutralize them. The relationship between patron and client is hierarchical and aligned, while the relationship between co-opter and co-opted is potentially adversarial. Inducing an opposed actor to cease opposition—and monitoring their ongoing compliance—requires more resources and carries greater risk of failure than sustaining a client who was never opposed to begin with (Eisenstadt and Roniger 1984; Scott 1972).³ Co-optation of hostile outlets may also generate backlash, as formerly independent organization lose credibility or as displaced media professionals create new dissident outlets (Ebsworth 2017; Kao 2026).⁴

Third, the commercial transactions underlying media patronage provide plausible deniability that direct state ownership does not. State broadcasters like RT and CGTN can be labeled as foreign agents, leading to government registration, platform labeling, and audiences to potentially discount their credibility (Mena 2020; Nassetta and Gross 2020). Covert state-controlled outlets, when exposed, suffer catastrophic loss in credibility (Cohen 2015;

³For more on the oppositional character of co-optation, see (Gandhi and Przeworski 2006).

⁴We find evidence of such backlash in Hong Kong, see Appendix A3.1.

Fittarelli 2024). With media patronage, client outlets retain editorial autonomy and continue to make independent decisions, particularly on matters outside of the patron’s core interests. This makes state sponsorship easy to deny. When media organizations decline patron support, they typically do so out of professional integrity and the desire to avoid real or perceived conflicts of interest rather than because the transactions themselves are illegal (of Professional Journalists 2014). The need to avoid even the perception of bias means that the patronage relationship is binary—it is the acceptance of resources from a political patron, rather than the magnitude of resources that constitutes the relationship. When patronage resources are accepted, client outlets are no longer fully independent, as they depend on maintaining a relationship with a political sponsor, but neither are they directly state-controlled (Bodrunova, Litvinenko and Nigmatullina 2021; Repnikova 2018; Stockmann and Gallagher 2011).⁵ This intermediacy is what makes media patronage effective. Clients retain sufficient independence to maintain local credibility while remaining sufficiently aligned to serve patron interests.

One aspect of media patronage we take as true by assumption is the financial vulnerability of media outlets: we assume all non-government supported outlets are financially vulnerable. Our theory of media patronage generates the following observable implications. First, clients should be **selected** for patronage on the basis of their pro-regime editorial inclination. After patronage, these organizations’ editorial inclinations should **not change** and should continue to compete with other media organizations by providing substantively similar news coverage. Second, patronage should help client outlets **survive** amid declining worldwide media revenues, especially compared to outlets which do not accept patronage support. Over time, this selective support should reshape the market for news as pro-regime outlets persist while unaffiliated outlets do not, advancing the regime’s strategic objectives.

Media patronage depends on several scope conditions. First, it requires the presence of

⁵In some ways resemble, patronage outlets resemble the intermediate status of semi-independent outlets within authoritarian borders. Hong Kong media outlets in the wake of the the territory’s 1997 handover to the PRC offer an illustrative comparison. The PRC during that time allowed the territory substantial autonomy, but the outlets nevertheless navigated implicit restrictions on reporting as the threat of PRC intervention loomed large, see (Lee and Lin 2006).

media organizations under financial strain. In contexts with robust media revenue or extensive domestic subsidies, the opportunities for patrons diminish. Second, media patronage requires heterogeneity in the media landscape such that potential allies exist for selection. In highly consolidated or homogeneous media environments, patrons may face difficulty finding suitable clients (in hostile environments) or have minimal incentives to conduct patronage as coverage is already positive (in supportive environments). Third, it requires a patronage actor willing to deploy resources to support media in competitive media markets. In theory, any well-resourced political actor can engage in media patronage, and some democracies have done so, such as when the U.S. subsidized anti-Communist journals during the Cold War (Saunders 2000).⁶ In practice, however, authoritarian regimes are more willing and better positioned to do so. Authoritarian regimes, especially those with strong domestic information controls, are less likely to face domestic constraints on selectively supporting aligned foreign media and have the institutional capacity for identifying and cultivating potential clients. They also often have strong strategic interests in shaping external information environments, where foreign media coverage can filter back to domestic audiences or influence transnational opposition movements. Lastly, the regulatory environment of the target market must permit foreign funding of media organizations without intense scrutiny such that patronage can operate as routine commercial activity. If foreign investment in domestic media is heavily monitored or restricted, patronage becomes riskier and less attractive compared with alternatives.

2. China and its Diaspora

We turn to China and Chinese-language diasporic media worldwide to evaluate our argument because this case meets the scope conditions for media patronage.

First, Chinese diasporic media are financially vulnerable. They serve relatively small

⁶More generally, democracies do fund foreign media through development assistance, but such programs claim to allocate resources based on journalistic quality or institutional capacity rather than content alignment, which would not constitute media patronage.

populations—Chinese diasporas are typically an ethnic and linguistic minority within the larger host society—which constrains both their advertising revenue and subscription base. Chinese diasporic media also face stiff competitive pressure from two directions: they compete for audience attention with both mainstream host country media and with Chinese mainland media increasingly accessible via digital platforms.

Second, Chinese diasporic media contains a wide distribution of media slants both for and against the PRC, allowing us to test a theory of media selection rather than media transformation. Some diasporic Chinese media outlets, such as those founded by practitioners of Falun Gong, a religious movement banned in China since 1999, are opponents of the Communist Party (Ownby 2008). Other Chinese media outlets are commercially-minded, seeking to turn a profit from informational needs of Chinese immigrant populations. Yet others with business ties to the Chinese mainland operate media organizations with a pro-regime slant. This diverse population presents a wide variety of media biases that can reveal a story of media selection.

Finally, the Chinese Communist Party has a strategic interest in influencing diasporic Chinese media. China’s overseas migrant population surpassed 8.6 million in 2024 (see Figure 1), with particularly large concentrations in established democracies (US 2.5 million, Japan 810,000) (United Nations, Department of Economic and Social Affairs, Population Division 2024).⁷ This population continues to consume media in Chinese: according to the National Asian American Survey, 66.3% of Chinese respondents read news in Chinese in 2008 (the last year in which the question was asked), with 41.0% preferring Chinese-language

⁷The U.N. defines “migrant” as individuals living for more than 12 months outside their country of birth or citizenship. When birth data is available, it is preferred to citizenship when calculating the statistics, see United Nations, Department of Economic and Social Affairs, Population Division (2024). The U.N.’s migrant population data is an imperfect proxy for Chinese news consumers for several reasons. First, some migrants may consume information in the host country language (e.g. a Chinese migrant in the U.S. consuming English-language news), leading this proxy to overestimate the Chinese news consuming population. On the other hand, some native-born speakers—who are not migrants and thus excluded from the U.N. count—may consume Chinese-language news alongside news in the host country language (e.g. a second-generation Chinese-Peruvian consuming Chinese-language news). This possibility would lead the U.N. estimate to undercount the Chinese news consuming population. In the Appendix, we also provide two other estimates for diasporic Chinese news consumers: unique website visits to Chinese websites and downloads of Chinese-language apps (see Appendix A1.4).

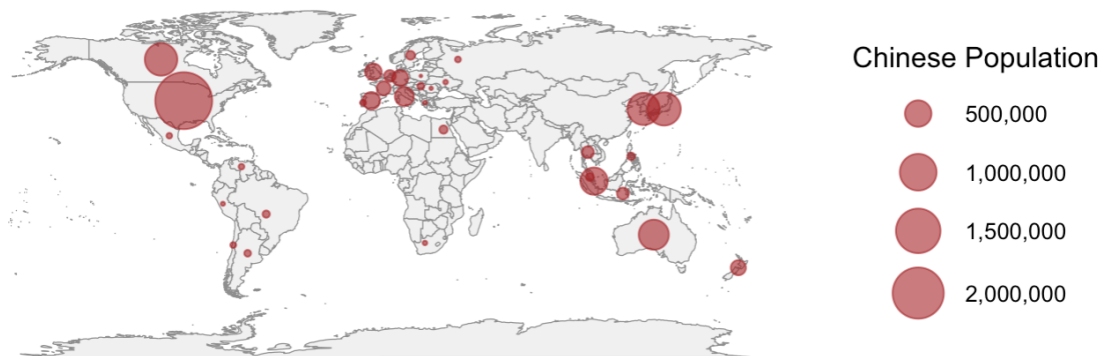


FIGURE 1. CHINESE MIGRANT POPULATION, 2024.

sources.⁸

Overseas Chinese have historically contributed to political movements on the mainland—from Sun Yat-Sen’s Nationalist fundraising among Hawaiian Chinese in the early 20th century to reform-era returnees contributing skills to China’s economic development in the post-Mao years. Chinese-American newspapers have been politically connected for generations: writing in the 1980s, [Lai \(1987\)](#) noted that Chinese-language media outlets often took advantage of America’s comparative media freedom to advocate their own vision for China’s development. More recently, overseas Chinese operating outside the Great Firewall have served as information conduits during political crises, such as the 2022 White Paper protests, sharing censored content to domestic audiences involved in the protests ([Chang et al. 2022](#); [Economist 2022](#)). Diasporic media organizations thus represent both an opportunity to cultivate favorable narratives among a strategically important population and a threat through which opposition movements can reach domestic audiences.

Beyond meeting scope conditions, focusing on the relationship between the CCP and Chinese-language diasporic media allows us to assess whether media patronage complements previously identified strategies. This is because the Chinese government can and often does target its own diasporic media producers with transnational repression and co-optation.

⁸The survey uses self-identification to determine Chinese identity.

In southeast Asia, the PRC has orchestrated the kidnapping of ethnic Chinese booksellers, including one with foreign nationality; in Australia, it has targeted ethnic Chinese journalists with coordinated campaigns of harassment (Hsiao-hwa 2024; Ramzy 2020). China has the willingness and capacity to practice transnational repression to influence foreign media when it chooses to. Does it also take advantage of the financial vulnerabilities of these news outlets through media patronage?

PRC policy documents confirm that resource provision to diasporic media outlets serves a strategic political interest. These documents mention two interrelated goals. One goal is to cultivate affective ties among overseas Chinese and foreign nationals of ethnic Chinese descent for use in the CCP’s political agenda, including support during political crises or geopolitical conflict (Chinese Radio Seattle 2023; World Chinese Media 2001). President Xi Jinping emphasized this goal by evoking the Maoist phrase “unite all that can be united” (团结一切可以团结的力量) at the 2015 United Front Work Conference (Xi 2015). Another is to use diasporic Chinese-language media as a channel to “transmit China’s voice” globally and strengthen China’s “international right to speak” (CNS 2012; OCAO 2007).⁹ While the first goal is focused on diaspora populations, the second also has implications for domestic audiences, given the importance of foreign sources during domestic crises (Chang et al. 2022; Chen and Yang 2019; Hanley, Lu and Pan 2025).

The Chinese regime pursues these goals through a centralized coordinating body, the Overseas Chinese Affairs Office (OCAO, 国务院侨务办公室), which was previously in the State Council (the PRC’s highest government organ) but has since been subsumed into the Party’s United Front Work Department (a Party organ dedicated to managing relations with non-Party actors). OCAO is explicitly tasked with influencing ethnic Chinese of foreign nationality. Beginning in 2001, OCAO and the China News Service (CNS) (*Zhongguo Xinwenshe*), a Chinese-language global news agency dedicated to producing “news and information for [...] overseas Chinese and Chinese with foreign nationalities,”¹⁰ jointly launched

⁹These objectives are consistent with a general shift toward aiming to use diasporic media, and diasporic populations more generally, as “public diplomats” on behalf of PRC strategic objectives (Thunø 2017).

¹⁰The CNS has been described by foreign observers as the CCP’s “main propaganda organ targeting overseas

the World Chinese Media Forum, a recurring summit that has convened hundreds of overseas Chinese media executives in China for sessions ranging from political orientation to content-sharing agreements to the construction of shared digital infrastructure (World Chinese Media 2001; Chinese Radio Seattle 2023), as well as other training programs. The integration of the OCAO into the United Front was part of a broader effort to institutionalize Party control of efforts related to overseas Chinese affairs. This structure, which includes dedicated bureaucracies and recurring coordinating forums and training programs, means that if media patronage occurs, it reflects a sustained top-down effort led by the CCP through the United Front, the OCAO, and the CNS.

Interviews with current and former representatives of Chinese-language diasporic media organizations confirm that PRC resource provision operates as our theory predicts (for more information on interviews, see Appendix A1.5). Organization F, located in a major Latin-American city, described being approached by a representative from the People’s Daily overseas edition: “In 2014, they [the representative] came by on a visit and saw our paper copy, which we had at the time, and saw that our position and point-of-view was proper so they invited us to partner with them” (2014年, 有《人民日报海外版》的人来访, 看了我们当时的纸质杂志, 也了解了我们的发行情况, 觉得我们的观点和立场都比较正面、比较对路, 所以后来就邀请我们合作).¹¹ Organization A, which is unaffiliated, recounted the sudden loss of advertising revenue because a PRC-owned company realized that A’s coverage choices were not appropriate: “Before Xi’s visit [to our country], Kweichow Maotai [a state-owned liquor company] had planned to co-sponsor an event with us. When they realized we were not *dangmei* [a pejorative term meaning ‘Party media’, 党媒] they withdrew in surprise [...] we’re quite fatalistic about the future of overseas media.”¹²

As indicated by the previous example concerning advertising, PRC patronage takes multiple forms. A representative of Organization I, located in a small European country, de-

Chinese, <https://jamestown.org/program/how-the-ccp-mobilized-a-cross-border-disinformation-campaign-against-the-czech-senate-speaker/>. See also <https://www.chinanews.com.cn/common/footer/aboutus.shtml>.

¹¹Interview with Organization F.

¹²Interview with Organization A, conducted in mixed English-Chinese.

scribed PRC article re-shares as key to generating traffic, noting that because the Chinese-speaking readership of their country was minimal—no more than 10,000—if an article has 40,000 viewers it must have views from mainland China: “Some [government] WeChat accounts will promote us and re-share our articles [...] good articles have to rely on Chinese audiences to generate real traffic” (有一些微信公众号会推广，会自动转发…好的文章要靠中国的读者上去，在这里是上不去的).¹³ Moreover, accepting or rejecting PRC resources is seen as a binary status: Organization H described getting social media contact requests from a representative from a PRC Propaganda Department (a potential precursor to patronage ties), but declined because “they’re very bureaucratic, and if you work with them you sometimes get pressure [...] we reduced costs to avoid that pressure.”¹⁴

3. Data

To evaluate our argument, we create two news datasets: a global dataset of 193 Chinese-language digital news organizations outside mainland China, Taiwan, and Hong Kong/Macau, and a historical census of 222 Chinese-language news organizations in the United States from 2000 to 2025. We describe each in turn before detailing how we operationalize media patronage across both.

3.1. Global Chinese-Language Digital News Dataset

To assess whether the PRC is conducting media patronage abroad, we construct a new dataset of Chinese-language digital news organizations outside of mainland China, Taiwan, and Hong Kong/Macau, which we refer to in this paper as the *Global Dataset*. We operationalize “digital news organization” as any *online entity* which (1) maintains an independent web domain hosting news content consistently accessible from U.S.-based web browsers as

¹³Interview with Organization I.

¹⁴Interview with Organization H, conducted in English.

of June 2024,¹⁵ and (2) features human creation or curation of news content.¹⁶ We further require that the organization (3) is not explicitly disclosed as owned or operated by a sovereign government;¹⁷ (4) covers general interest news rather than a single topic (e.g., the oil industry); and (5) is not a duplicate of content syndicated under another domain already in the dataset.¹⁸

We adopt these criteria for the following reasons. We require an independent web domain because domains proxy for a minimal level of organization capacity—the purchase, design, and upkeep of a domain name requires greater organizational capacity than opening a social media channel—and provide a more stable unit for cross-sectional analysis than social media accounts, which can vary across platforms, regions, and over time. Social media content also introduces platform-based variance both in format (e.g., character limits or unique symbols), and also, crucially, in degree of censorship. Chinese-language news production on social media frequently occurs on Chinese domestic apps such as WeChat or Weibo, which means that all news content, including content produced abroad, is already subject to platform-based censorship implemented on behalf of the CCP. Since we are interested in authoritarian influence on overseas news production, studying content on domestic Chinese platforms would mean that all units are in some sense “treated” due to compliance with domestic Chinese regulations. We exclude sovereign government-owned outlets because our interest is in whether authoritarian regimes can influence ostensibly independent news organizations.

We aimed to identify the global population of Chinese-language news organizations, many of whom serve small diasporic communities, which presented a substantial challenge. We employed five overlapping approaches to construct this study population. The first approach

¹⁵Four organizations were not consistently accessible but otherwise met the study criteria.

¹⁶This definition therefore excludes social media accounts or video channels which do not possess their own domain name, as is the case with many Chinese-language “self media” (*zimeiti*) accounts, often operated by an individual or small team of individuals. It includes outlets which primarily translate or curate news while producing minimal content themselves.

¹⁷This excludes outlets such as the British Broadcasting Corporation (BBC), Russia Today (RT), and the China Global Television Network (CGTN). If a website states that it is operated from both mainland China and abroad, it is excluded. If a website does not state from where it is operated, it is included.

¹⁸For example, *Sinchew Daily* (星洲网) is a Malaysia-based Chinese-language newspaper which operates in Indonesia under a different masthead. Its Indonesian version is considered a duplicate and is not included in the analysis.

used web traffic. We acquired a list of the top million global web domains ranked by website traffic from the network analysis firm, Cloudflare. We then subset to all domains labeled as “news” by Cloudflare, and further subset to all websites written in simplified or traditional Chinese.¹⁹ The second approach entailed a breadth-based search of one of the most popular diasporic websites, *Literature City*, identifying all news organizations it linked to between 2019-2021. We then located as many news domains as possible associated with these news organizations. In the third approach, we identified all diasporic news organizations indexed by WiseNews, a popular database of Chinese-language news content. Fourth, we contacted an *NPR* reporter, Emily Feng, the author of a 2018 *Financial Times* investigation into PRC influence abroad,²⁰ who graciously shared her database of news organizations with us. The fifth approach involved looking up each news website on SimilarWeb, a web data analytics company, to identify additional websites which attract similar users and web traffic patterns to the news organizations we had already identified, and recursively repeating this process until we did not identify any further organizations.²¹ After collecting these domains we manually checked each organization to ensure it met our criteria for news production. We identified 193 Chinese-language digital news organizations with this data collection process.²² Despite serving limited audiences, these sites remain quite popular: on average, they have a median of 6,309 monthly visitors (mean = 558,634).²³

3.2. US Chinese-Language News Census Dataset

Because we are interested in measuring survival at the media market level, we construct a second, new dataset that represents a historical census of both print and digital news outlets

¹⁹We used the Python library *langdetect* to identify the website’s language.

²⁰See <https://www.ft.com/content/f5d00a86-3296-11e8-b5bf-23cb17fd1498>.

²¹We refrain from providing the names of individual outlets to minimize risks of reputational, commercial, or regulatory harm.

²²In this process, we also identified 46 news organizations based in Hong Kong, Macau, and Taiwan; we use these outlets for an exploratory analysis of the effect of physical repression in the case of Hong Kong and dissident newspapers, see Appendix A3.1.

²³For robustness, we also gathered these numbers from alternate data provider, SimilarWeb.com, and find them broadly comparable (176/193 outlets, *median* = 13,200, *mean* = 516,464).

operating in the U.S. market since 2000, which we refer to as the *US Census Dataset*.²⁴

Unlike the Global Dataset where we identified outlets at the time of study and collected data going back in time, for this historical US news organization census, we began with historical sources and collected forward in time. Through extensive searches of archival and digital sources, we identified 222 outlets active in the United States anytime between 2000 and 2025, of which 93 were active through the start of our study period.

Our inclusion criteria for U.S. news organizations is broader than for global digital news organizations as disambiguating Chinese-language historical newspapers, many of which share the same English or Chinese names, was a formidable challenge, and many outlets have no website or produce only scanned versions of a print newspaper. Because we could not reliably determine whether a defunct outlet was based in the United States or merely stationed foreign correspondents there, we included any organization which archival sources mentioned as having a U.S. presence. We also assumed that an outlet covered general interest news unless their organization name explicitly named a specialty news interest (e.g. “New York Art Weekly”). For more information on how we assembled this historical dataset, including overlap to the Global Dataset, see Appendix [A1.2](#).

3.3. Operationalizing Patronage

We conceptualize patronage as a binary condition: having accepted CCP resources in any form. We rely on three sources to identify whether a news organization in either the Global Dataset or US Census Dataset has accepted PRC resources. If any of the three sources reveal that an outlet has accepted resources, we designate that outlet as linked via media patronage.

Our first source is investigative reporting. Specifically, we turn to the underlying data for *FT*’s 2018 investigative story into PRC content-sharing agreements, and record the year

²⁴We gather metadata on their existence over time, recording as their end date the date of their last published story or the reported closure of their newspaper, whichever is later. Where digital versions of newspapers are not available, we use the date of last archived print copy.

in which each content agreement was signed. After manually reviewing all content-sharing announcements, we identify 67 diasporic outlets in our dataset which have entered into partnerships with PRC organizations.

Our second source is the Chinese Communist Party, and in particular an organization called the Global Chinese Media Cooperative Union (GCMCU) (*Shijie Huawen Meiti Hezuo Lianmeng*), operated by the CNS within the Party’s United Front Work Department. The GCMCU publishes a list of all news organizations collaborating with the CCP, alongside the date that an agreement was announced. Using this method, we identify 112 news organizations partnering with the Chinese government, of which 57 also appear as government-linked via investigative reporting based on the first source.

Our final source is a “Internet Content Provider” (ICP) permit on the diaspora news organization domain pages. Here we exploit a feature of Chinese internet regulation: according to Chinese law, all China-based webpages must obtain an ICP permit from the Chinese Ministry of Industry and Information Technology. This ICP number indicates that the website is being operated in accordance with PRC regulations, including speech restrictions. This ICP number facilitates access to domestic Chinese audiences and the attendant advertising revenue, but comes at the cost of the website’s editorial independence by exposing it to legal liability. It indicates that the outlet accepts access to the Chinese market conditional upon circumscribed editorial decisions. On many webpages, the ICP number is displayed in small font at the website’s footer. By sifting through all websites homepages archived by the Wayback Machine, we can identify the first time that an ICP number appears on a news organization’s homepage (see sample footer Figure 2). We then mark that date as the date of first affiliation. Using this method, we identify 33 news organizations as government-linked, of which 27 are also identified via the other two approaches above.²⁵

The provision of patron resources is sometimes unobservable—direct cash transfers, for instance, leave no public trace—making a full accounting of the patronage relationship im-

²⁵Once affiliated, outlets do not appear to renege; of the 29 outlets revealed as linked in the 2018 FT reporting prior to 2013 (10 years before our study’s onset), 27 of them remained linked via GCMCU or ICP numbers at the time our data collection began.



FIGURE 2. WEBPAGE FOOTER WITH ICP NUMBER.

possible. Any operationalization based on visible markers is therefore likely to undercount patronage’s actual prevalence, as some outlets may have accepted resources through channels we cannot observe. Reassuringly, however, this undercounting biases our study against finding significant results, as patronage outlets incorrectly classified as unaffiliated would diminish the observed differences between the two groups in selection and survival.

4. Patronage as Selection

The first observable implication of media patronage is that outlets should be selected for patronage on the basis of pre-existing slant. After selection, their reporting should not change, and they should continue to compete with unaffiliated organizations for market share by producing similar news coverage.

To evaluate this, we gather all published articles from news organizations in the Global Dataset, including URL, date, title, and content, appearing on each news domain, reaching as many years back as each news organization’s website permitted us. Data collection began in March 2024 and ended in January 2025. After data cleaning and de-duplication (see Appendix A1.3), our data consisted of 13,990,206 articles spanning over 30 years, with a median of 16,145 articles (and mean of 72,488 articles) per organization. This source of data is rich but high variance, as articles from some news organizations were accessible for over 20 years while others were available only a few years into the past. On average, our dataset

consists of 6.8 years of data per outlet (median of 6 years).

We construct two variables to measure whether an outlet’s editorial slant favored China prior to the provision of resources. The first is stance toward China, as “telling China’s story well” is a primary strategic objective of the PRC. We measure an outlet’s stance toward China by using a three-step machine-assisted labeling process with human validation. First, we determined each outlet’s percent of China coverage by asking human coders to label a 1,600-article sample (stratified by outlet and year) as related to China (intercoder agreement = 95.2%), and use that sample to fine-tune a supervised learning model (DeBERTa) to replicate those labels on our full dataset (F1=0.91).²⁶ Second, we asked human coders to assess the stance toward China of each China-related article ($n = 654$, intercoder agreement = 81.8%), with 1 representing “positive” and 0 representing “neutral or negative.” Finally, we used GPT-4o to replicate those labels on China-labeled articles from Step 1.²⁷ For cost reasons, we sampled 10% of articles of outlets with >10,000 articles. For more details on coding rules, see Appendix A2.1, and for the generation of stance classifications, see Appendix A2.2. Our final labels contain three sources of error: DeBERTa’s misclassification of China relevance, GPT’s misclassification of stance, and sampling error for the 62 outlets from which we analyzed only 10% of articles. We treat DeBERTa’s China-relevance classifications as defining our analysis sample and do not propagate this uncertainty. We propagate the latter two sources using a non-parametric bootstrap (1,000 iterations) that randomly flips GPT labels at the estimated classifier error rate and re-samples from partially sampled outlets, constructing confidence intervals from the resulting distribution of values (see Error Propagation in Appendix A2.4 for details).

The second variable of editorial slant is censored topic coverage. We select four topics related to territorial separatism and accusations of Party wrongdoing that are considered potential “focal points” for protests (Stern and Hassid 2012; Truex 2019): Falun Gong, the Tiananmen Square massacre, the Dalai Lama, and Xinjiang concentration camps. In

²⁶<https://huggingface.co/IDEA-CCNL/Erlangshen-DeBERTa-v2-320M-Chinese>

²⁷GPT-4o-mini-2024-07-18

China, any mention of these topics by private or semi-private media outlets is forbidden; only “white-listed” (Party-owned) outlets are permitted to discuss them, and they seldom do so (Fu, Chan and Chau 2013; Yin, Fu and Zheng 2024). As a result, these topics are rarely discussed at all.²⁸ We check for the presence of any word related to these four events by compiling a list of related words for each event and averaging the total number of articles mentioning each of these keywords (see Appendix A2.5 for a full list of keywords). An outlet’s coverage rate of these censored topics is the percentage of its stories containing these keywords.

We also measure whether patronage outlets compete with unaffiliated outlets for market share post-patronage by offering similar local and global news coverage. To measure the proportion of each outlet’s coverage devoted to local news—defined as news pertaining to the organization’s country of residence rather than China—and use the same human-validated DeBERTa approach described above, achieving an F1 of 0.92 (see Appendix A2.3), and measure similarity using Welch’s two-sample T-test, where patronage and unaffiliated outlets each represent a sample and the outcome of interest is each outlet’s share of articles dedicated to local news. To measure similarity in global events, we compile a global news agenda from Wikipedia’s annual tally of globally important events and use a human-in-the-loop machine learning pipeline to match outlet articles to events (F1=0.97, see Appendix A2.6 for details). Our outcome measure is the percent of each outlet’s monthly articles dedicated to covering each global event.

One possibility is that patronage outlets may frame events differently from unaffiliated outlets, even if the proportion of coverage dedicated to global events is the same. To assess this, we embed the full text of the outlets’ articles about each event, residualize them against the average coverage embedding for that event, and aggregate these residuals by outlet to generate a framing embedding representing each outlet’s deviation from mean coverage. We benchmark against a commercially-oriented PRC state media outlet, *The Paper* (*Pengpai Xinwen*, 澎湃新闻), and assess whether patronage outlets are systematically closer in em-

²⁸The sensitivity of these terms were also confirmed during our qualitative interviews.

bedding space to this state media framing than unaffiliated outlets are (see Appendix [A2.7](#) for details).

4.1. Selection on Pre-existing Slant

Our results indeed show that Chinese language news outlets selected for patronage are significantly more positive toward China than unaffiliated outlets, and are more likely to avoid mention of censored topics prior to patronage. In Figure [3](#), each dot represents an outlet, while the larger dot is the annual average of each group: in orange are outlets that will later receive economic resources from the CCP (will be selected for patronage but have not yet been). After an outlet partners with the PRC, they are removed from the figure. In blue are outlets that remain unaffiliated with the PRC during the duration of the study period. Panel A shows that patronized outlets are more positive about China prior to patronization than unaffiliated outlets in all years during the study period. Similarly, Panel B shows that later-patronized outlets rarely—and in many cases never—discuss sensitive topics, in contrast to unaffiliated outlets’ occasional coverage of these issues. Together these results show that the CCP selectively supports outlets that are already sympathetic to PRC strategic objectives.

More formally, we also estimate a discrete-time hazard model to assess the likelihood that an outlet’s pre-affiliation news coverage stance toward China predicts its likelihood of affiliating in a given year (β_1), conditional on not having affiliated previously:

$$\Pr(Y_{ict} = 1 \mid Y_{ic,t-1} = 0) = \text{logit}^{-1}(\alpha_c + \beta_1 \text{ChinaStance}_i + \beta_2 \text{TimeAtRisk}_{it}) \quad (1)$$

We include a time at risk covariate to control for the amount of years an outlet has operated without patronage (β_2), and country fixed effects (α_c) to control for the risk that some country’s media environments are more likely to be associated with patronage. Standard errors are clustered at the outlet level. The results in Table [1](#) show that outlets with more positive China coverage face substantially higher hazards of affiliation ($\beta_1 = 4.201$, $p <$

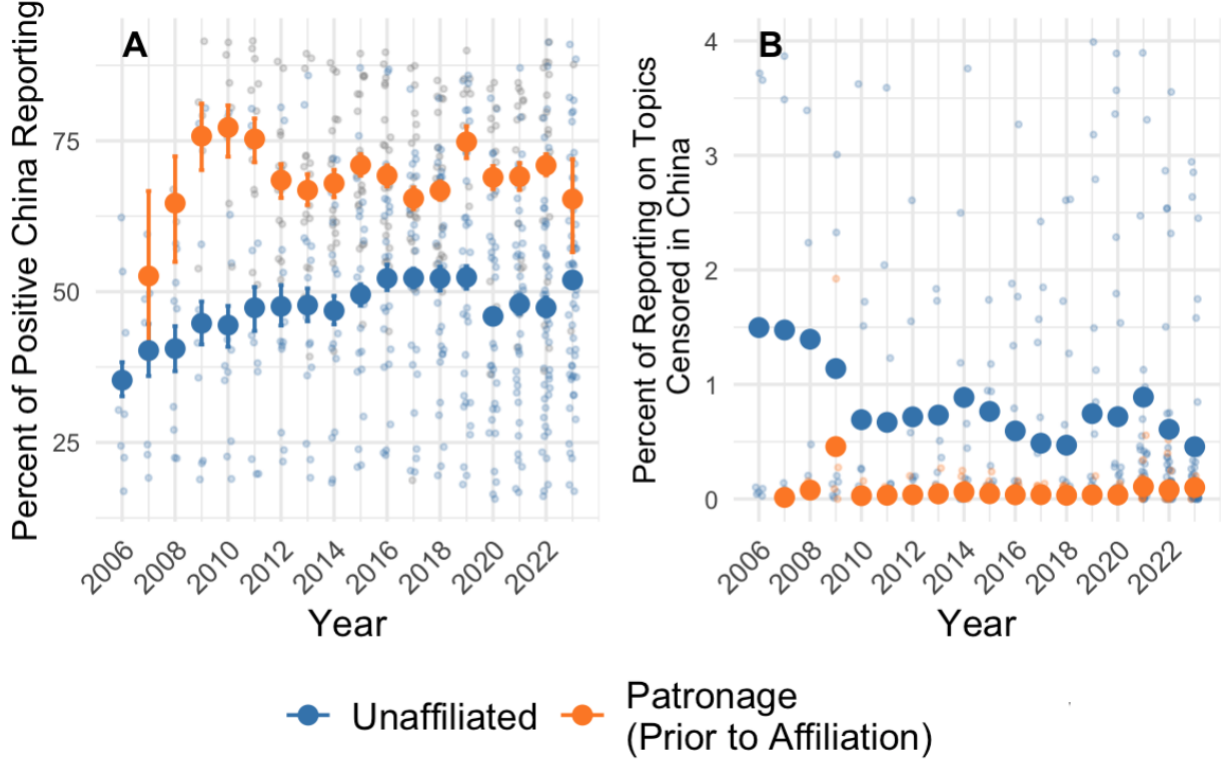


FIGURE 3. SLANT OF COVERAGE OF MEDIA OUTLETS PRIOR TO LINKAGE FORMATION. PANEL A: STANCE TOWARD CHINA. PANEL B: CENSORED TOPIC COVERAGE.

0.001), even after controlling for country fixed effects and time at risk. The positive coefficient on time at risk (β_2 , $p < 0.01$) suggests that outlets are more likely to affiliate the longer they operate without patronage.

TABLE 1. DISCRETE-TIME HAZARD MODEL WITH COUNTRY FES

	Affiliation Event
China Stance	4.201*** (0.561)
Time at Risk	0.051** (0.017)
Constant	−5.912*** (0.524)
Country FE	Yes
Number of Countries	39
Observations	1,629
Log Likelihood	−788.393
AIC	1,658.786

4.2. No Change in Coverage after Patronage

To test whether media patronage works by shifting market share (instead of changing content), we focus on the 28 outlets for which web-scraping provided at least one full year of full-text data available before and after patronage, which is a constraint driven by variation in number of historical articles stored by different outlet websites. For each outlet, we calculate the annual percentage of stories with positive stance toward China and containing censored topics, and compare these metrics to the average of all never-affiliated outlets in the same year.

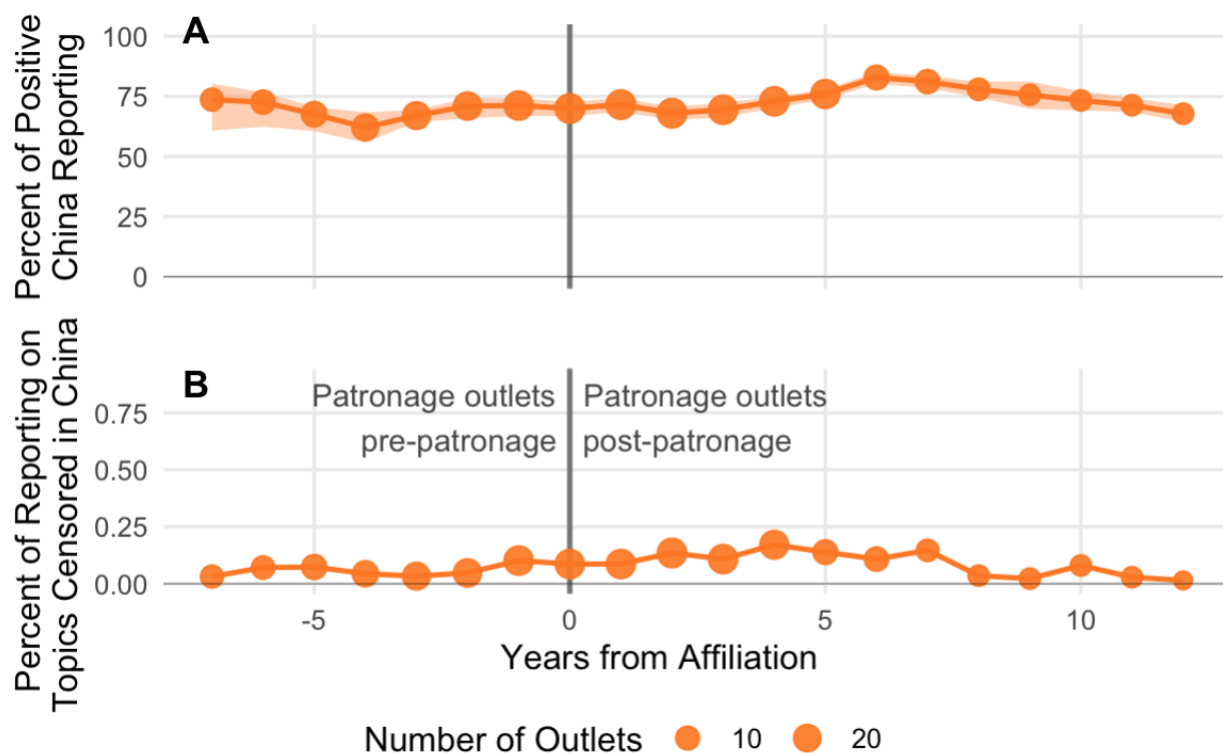


FIGURE 4. SLANT OF COVERAGE OF MEDIA OUTLETS BEFORE AND AFTER LINKAGE FORMATION. PANEL A: STANCE TOWARD CHINA. PANEL B: CENSORED TOPIC COVERAGE.

In alignment with our expectation, Figure 4 shows that post-patronage (to the right of the vertical black line marking the year of first affiliation, in red), the stance of patronized outlets remains consistent with coverage prior to patronage, with no increase in positive stance or decrease in coverage of censored topics after linkages are established. (We also

check whether outlets become more negative toward their host country and find that they do not, see Appendix A3.2.) These results show that media patronage seeks to support already friendly outlets rather than transforming critical ones.

4.3. Similarity in Local and Global News Coverage

Figure 5 shows that after patronage, patronage media outlets dedicate almost 50% of their news coverage to local news, nearly the same as unaffiliated outlets.²⁹ Across all years, patronage outlets dedicated 44.5% of their coverage to local news, compared to 44.9% in unaffiliated outlets (difference in means is not significant).

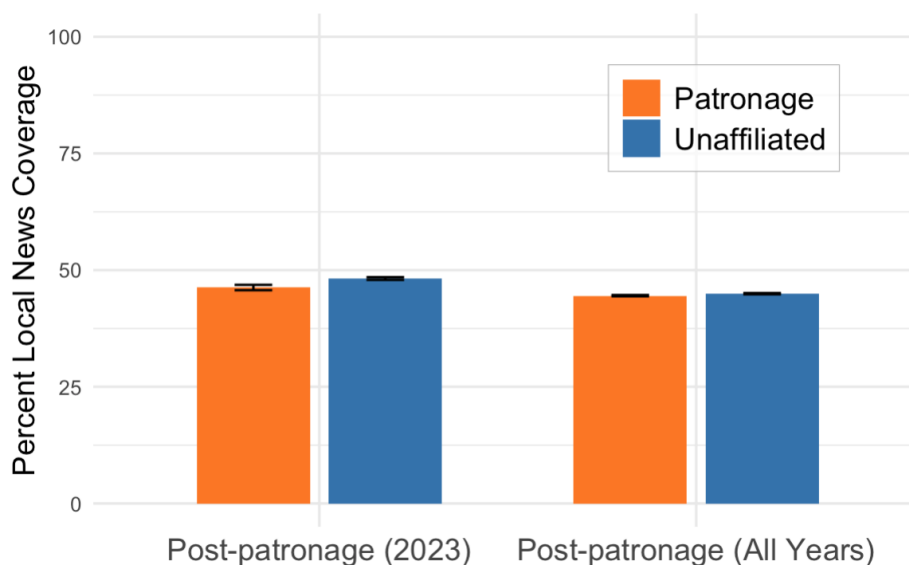


FIGURE 5. LOCAL NEWS COVERAGE.

On global events, Figure 6 Panel A shows that patronized outlets (orange) cover 2023's biggest global stories at rates more similar to unaffiliated outlets (blue) than to commercially-oriented CCP-owned outlet *The Paper* (mustard yellow). In 7 of 11 months, the coverage rate estimate between patronage and unaffiliated outlets overlap. Patronage media covered two events, the U.S.-China spy balloon incident in February and the Finland election in April, which did not appear in Party-owned media at all.³⁰

²⁹Confidence intervals derived from classifier error.

³⁰No events received coverage in August.

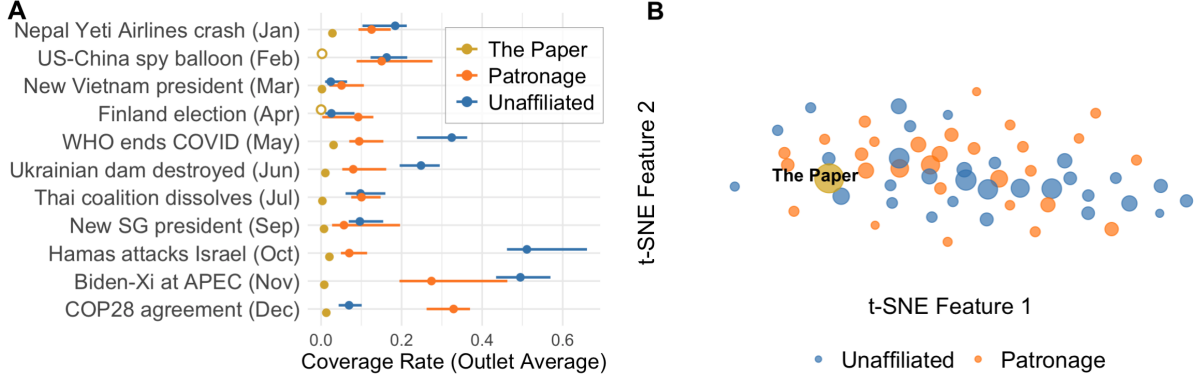


FIGURE 6. GLOBAL NEWS SELECTION AND FRAMING SIMILARITY. PANEL A: PORTION OF COVERAGE DEDICATED TO EACH MONTH’S TOP EVENT. PANEL B: AVERAGE FRAMING OF 2023 EVENTS.

When it comes to how these global events are framed, patronage outlets do not appear to cluster exclusively with *The Paper*. This is shown in Figure 6 Panel B, a two-dimensional t-distributed stochastic neighbor embedding (t-SNE) projection for the framing residuals for each outlet’s coverage of 2023 global events (≥ 5 articles covering events, $n = 60$). Patronage outlets’ framing is closer to *The Paper* than unaffiliated outlets, but the difference is not significantly larger than what might be observed due to random chance (difference in cosine distance between groups = 0.345, $p = 0.16$). This suggests that patronage outlets retain editorial independence in how they cover global events, consistent with the autonomy that makes them valuable to the patron.

While there remains substantial heterogeneity among outlets, overall these results suggest that patronage outlets continue to compete with unaffiliated outlets after the onset of patronage to provide local and international news coverage to audiences.

5. Patronage and Outlet Survival

The second observable implication of our theory is that patronage helps friendly outlets survive, gradually reshaping the composition of media producers within the market. We evaluate this using two measures: (1) the effect of patronage on web traffic for organizations in the Global Dataset, and (2) the association between patronage and outlet survival in the

US Census Dataset.

For digital news websites dependent on viewership and advertising sales, greater traffic numbers mean more advertising income and thus increased revenue necessary for survival. We gather data on estimated historical web traffic data dating to 2017 from Semrush, a web analytics company, retrospectively in February 2026 for outlets in the Global Dataset.³¹ Estimated web traffic was available for 187 of 193 news organizations, with 116 outlets available for the full nine year time span, and a mean of 7.6 years available for each website. If an outlet had no recorded traffic in all twelve months of a given year, we recorded that outlet as missing. Using this historic traffic data, we then measure the effect of patronage on treated outlets' historical web traffic using generalized synthetic controls (Xu 2017), which estimates latent factors to allow for unobserved time-varying heterogeneity. We aggregate traffic numbers at each quarter (e.g. Jan-Mar), and allow the estimator to search for up to two latent factors when estimating the following model:

$$Y_{it} = \delta D_{it} + \alpha_i + \xi_t + \lambda'_i f_t + \epsilon_{it} \quad (2)$$

where $Y_{it} = \log(1 + \text{visits}_{it})$ is log-transformed quarterly traffic for outlet i in quarter t in a five-year window around the treatment year, D_{it} is a binary treatment variable set to one beginning the first quarter after PRC affiliation year, α_i is an outlet fixed effect, ξ_t is a quarter fixed effect, $\lambda'_i f_t$ captures unobserved time-varying heterogeneity across outlets, and δ is the average treatment effect on the treated (ATT). Note that this specification drops observations from the treatment year because we do not know the exact quarter of affiliation and treatment may take time to take effect.

To assess whether patronage is associated with greater outlet survival, we estimate the protective effect of PRC affiliation with a Cox proportional hazards model on the US Census Dataset. The hazard function for outlet i at time t is specified as:

$$h_i(t) = h_0(t) \exp(\beta_1 \text{Affiliated}_{it} + \beta_2 \log(\text{Age}_{it})) \quad (3)$$

³¹Semrush uses anonymous click-stream data from third-party data providers, alongside proprietary modeling, to estimate global web traffic, see <https://www.semrush.com/kb/1211-how-semrush-turns-traffic-data-into-traffic-intelligence>.

where $h_0(t)$ is an unspecified baseline hazard, Affiliated_{it} is a time-varying binary indicator that switches from 0 to 1 in the year outlet i enters a PRC patronage relationship, and $\log(\text{Age}_{it})$ is the natural log of outlet age in years. Standard errors are clustered at the outlet level to account for within-outlet correlation across periods.

5.1. Patronage and Web Traffic

Our analysis using generalized synthetic controls confirm that PRC affiliation boosts web traffic to Chinese-language news organizations. Figure 7 shows the effect of PRC affiliation on affiliated outlets’ logged web traffic (only for outlets with one full year of pre-affiliation web traffic data, $n = 32$).³² The estimated ATT on logged web traffic across the eight quarters post-affiliation year is 0.239 ($p = 0.017$), equivalent to approximately 27% higher quarterly traffic relative to synthetic counterfactuals.³³ That increase in traffic may partly reflect increased story volume: almost a third of outlets for which pre/post patronage data is available doubled their story produced in the two years post patronage (8 of 27), relative to the two years prior to patronage, consistent with patronage resources enabling greater content production.³⁴

5.2. Patronage and Outlet Survival

Patronage also helps outlets survive longer than their unaffiliated peers (See Table 2). The estimated hazard ratio for affiliated outlets is 0.423, $p = 0.045$, indicating that affiliated outlets face approximately 57% lower hazard of closure relative to unaffiliated outlets, after controlling for outlet age (logged). The estimate is imprecise in part because so few affiliated outlets cease operations: of the 26 U.S. outlets active since 2000 that receive patronage, only

³²The model finds no latent time-varying factors, making it equivalent to a difference-in-differences estimator with staggered adoption.

³³Logging web traffic is common practice for long-tailed distributions such as web traffic. Nevertheless for robustness we also generated synthetic controls on unlogged quarterly web traffic, and found that results are directionally consistent (effect=956,785) but are not significant at the 95% level because a small number of high-traffic outlets inflate the standard error ($p = 0.18$).

³⁴The average volume increase post-patronage across all outlets was 2,393 stories/year.

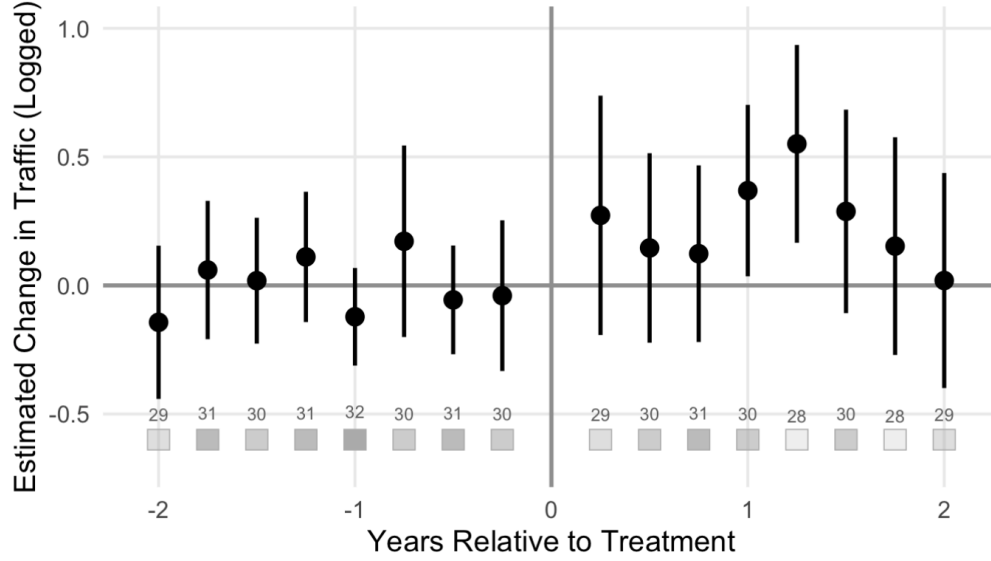


FIGURE 7. ESTIMATED EFFECT OF PRC AFFILIATION ON LOGGED WEB TRAFFIC.

VERTICAL LINE REPRESENTS YEAR OF AFFILIATION; EACH DOT IS A QUARTERLY OBSERVATION (PATRONAGE $n = 32$, UNAFFILIATED $n = 63$). SHADED BOX REPRESENTS TREATED OUTLET VOLUME.

five have ceased operations (19%), compared to 124 out of 196 unaffiliated outlets during the same time period (63%).³⁵

TABLE 2. COX PROPORTIONAL HAZARDS MODEL

	(1)
Affiliated	0.423* [0.183, 0.981]
Outlet Age	0.992+ [0.982, 1.002]
Num. Obs.	3375
Num. Events	131
LR test (χ^2)	7.14
LR p -value	0.028

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

³⁵This estimate of unaffiliated die-out roughly aligns with nationwide estimates of newspaper disappearances: Metzger and the Medill Local News Initiative (2025) estimates that 40% of American newspapers disappeared over the 10-year span between 2016 and 2025, about half the time frame of our analysis.

5.3. Patronage Reshapes Markets

Together, selection and differential survival reshape the composition of media markets over time. At the time this research began in March 2024, we identified 128 of 193 outlets in our Global Dataset as linked through media patronage. Looking retrospectively from 2024, the percentage of patronage outlets appears to have risen steadily since 2000 (see solid line in Figure 8). However because the Global Dataset only includes organizations still publishing as of 2024, it suffers from survivorship bias: defunct outlets are absent from the data set, inflating the apparent share of patronage outlets in earlier years (for more information on missingness see Appendix A1.1). To provide a lower bound on the share of historical patronage outlets, we construct an adjusted estimate that accounts for historical outlets that ceased operation prior to our data collection using the US Census Dataset, and conservatively estimate that these ceased outlets were unaffiliated (dashed line in Figure 8).³⁶ Using the US Census Dataset, we also construct a U.S.-only timeline of historical patronage rates in the U.S.

Figure 8 shows how, over time, media patronage has reshaped the global (solid line, with survivorship bias, and dashed, an alternative minimum) and U.S. (dotted line) shares of Chinese-language media production. These patronage outlets are also widely distributed across the globe. Figure 9 shows the number of outlets by country and proportion that are linked to the CCP. These data show that, by 2024, the Chinese government has established linkages with diasporic media organizations in nearly all countries with diasporic media outlets (38 out of 39 countries), with greater percentage of linkages in countries with fewer outlets overall. The breadth of these linkages suggests that media patronage is not confined to a few markets but operates as a systematic global strategy. We also find that outgoing web

³⁶We extract a year-by-year outlet die-off rate between 2000 and 2024 for all outlets in the US Census Dataset, with each year’s outlet count rendered as a percent of the 2024 total (e.g. 2021 contains 106% of the news organizations present 2024, with the additional 6% having later ceased publication). We conservatively assume that all of these defunct organizations are unaffiliated and then divide the known patronage count in each year by this larger denominator which includes soon-to-be-defunct organizations, to generate the adjusted minimum. Note that this estimate assumes the U.S. outlet die-out rate approximates the global die-out rate.

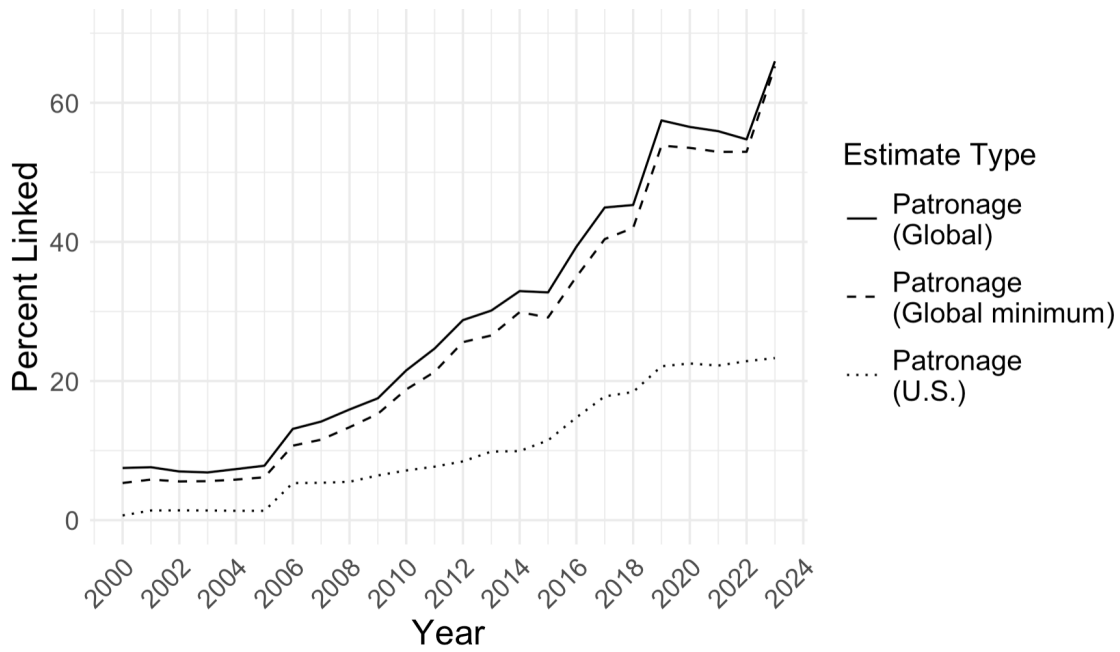


FIGURE 8. PERCENT OF DIASPORIC MEDIA ORGANIZATIONS WITH PATRONAGE LINKS TO THE PRC

links (i.e. HREFs) are increasingly directed toward China during this time (we demonstrate this with HREF data gathered from the Wayback Machine’s archived snapshots of digital outlets’ homepages, see Appendix [A3.3](#)).

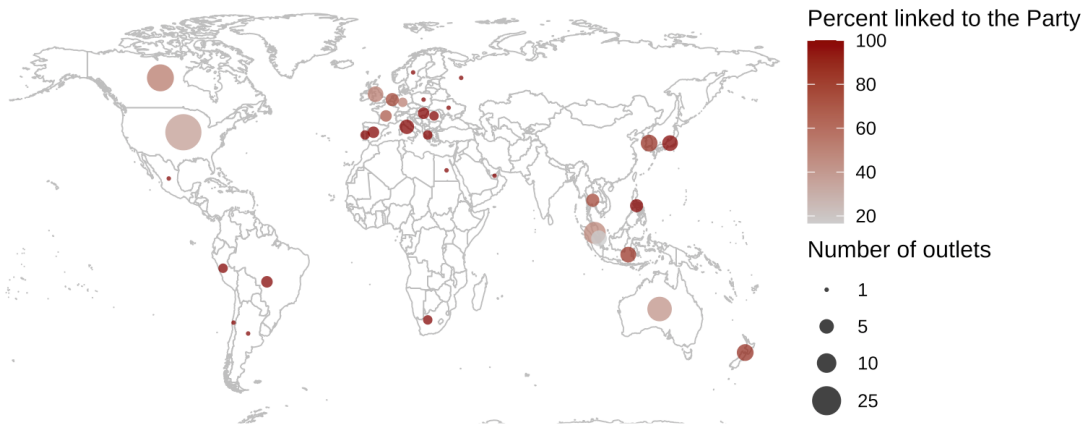


FIGURE 9. PRC LINKAGES BY COUNTRY, 2024

6. Implications of Patronage

What is the PRC’s strategic interest in influencing overseas media markets? In addition to external objectives such as increasing soft power or strengthening public diplomacy, one potential objective is mitigating global media outlets’ threat to domestic information control—and in particular preventing them from helping to coordinate collective action (Guriev and Treisman 2019; King, Pan and Roberts 2013). This threat is particularly acute during political crises, as the increased demand for credible information leads citizens to exert greater effort to seek foreign information sources (Chang et al. 2022). We examine the case of China’s White Paper Protests in late 2022, China’s most widespread public protests since the Tiananmen Incident in 1989. Citizens specifically sought out foreign information during this time, and the event was of likely interest to Chinese-speaking populations worldwide. The protests’ lack of recent precedent make it challenging for the PRC to anticipate how media outlets will respond. Do patronage media organizations cover the protests, and how widely does that coverage circulate domestically?

To answer this question, we examine both the production and reception of global Chinese-language news in the 40 days surrounding China’s White Paper protests (November 1 to December 10, 2022). On the production side, 153 of 193 outlets in the Global Dataset published in this period (total of 142,011 articles). On the reception side, we use a Chinese public opinion monitoring dataset that tracks hundreds of million of posts daily across all major Chinese social media platforms to identify of outlets’ domains and Chinese-language names, yielding 81,511 posts referencing 106 organizations.³⁷ We identify protest coverage using keyword searches for terms associated with the White Paper protests and the Zero-COVID policy.³⁸

³⁷In instances where names were a subset of each other (e.g. *Meiguo Huaren Zhoubao* 美国华人周报 and *Huaren Zhoubao* 华人周报) we used greedy matching to match the superset but not the subset. In instances where news organization names were identical (e.g. *Ouzhou Ribao* 欧洲日报) and thus posts could refer to multiple organizations, we used news organization country of origin to disambiguate between matching titles. If multiple countries were mentioned, we concluded that we were unable to disambiguate and removed them. $n = 7484$ posts could not be disambiguated and were removed from the analysis.

³⁸As of March 2026 we are in the process of hand-labeling a validation set for use in a machine-learning

6.1. Patronage Outlets Have Less Protest Coverage and Greater Share of Chinese Market

Of outlets publishing 10 or more articles over this period, patronage outlets ($n = 77$) covered the protests less widely than unaffiliated outlets ($n = 69$), with an average of just 4.2% of patronage outlet coverage (43 stories) dedicated to China’s Zero-COVID policies or subsequent protests, compared to 10.5% (86 stories) for unaffiliated outlets.³⁹ Those outlets were also over-represented in China relative to their global traffic volume. Among the 106 outlets that are mentioned in the Chinese public opinion monitoring dataset, patronage and unaffiliated coverage circulated at similar rates (patronage *mean* = 847, *median* = 101, $n = 52$; unaffiliated *mean* = 693, *median* = 85, $n = 54$) despite unaffiliated outlets receiving almost five times as many website visits globally as patronage outlets during a similar time (patronage=352,702, unaffiliated=1,532,940, $p = 0.037$).⁴⁰ In other words, patronage outlets are visible domestically in China far beyond what their global audience size would suggest.

Patronage outlets’ over-representation in China, relative to their global audience share, means that unaffiliated outlets’ protest coverage formed a smaller share of Chinese domestic circulation of foreign Chinese-language news during this White Paper protest period. Figure 10 shows protest-related news coverage from both types of outlets (Panel A), as well as how that coverage circulated domestically over the same time period (Panel B). Even before the Urumqi apartment fire on November 24 which is usually considered the catalyst of widespread protests (left dotted line), unaffiliated outlets account for a substantially larger share of Zero-COVID-related news posts than patronage outlets (69.5% vs 30.5%). After

pipeline to replace this keyword pipeline. Searched keywords were “protest” (抗议), “demonstration” (示威), “public procession” (游行), “white paper” (白纸), “reopening” (解封), “zero-COVID policy” (清零), “Urumqi” (where the fire that accelerated the protests took place, 乌鲁木齐), and “Xinjiang fire” (新疆火灾).

³⁹Of the 69 unaffiliated outlets at the time of the protests in 2022, 24 of them will later affiliate in either 2023 or 2024.

⁴⁰The difference in outlet means in the public opinion monitoring dataset is not statistically significant (Welch Two Sample t-test, $p = 0.759$). Results for the are consistent for mean of logged visits (patronage=7.24, unaffiliated=9.40, $p = 0.032$).

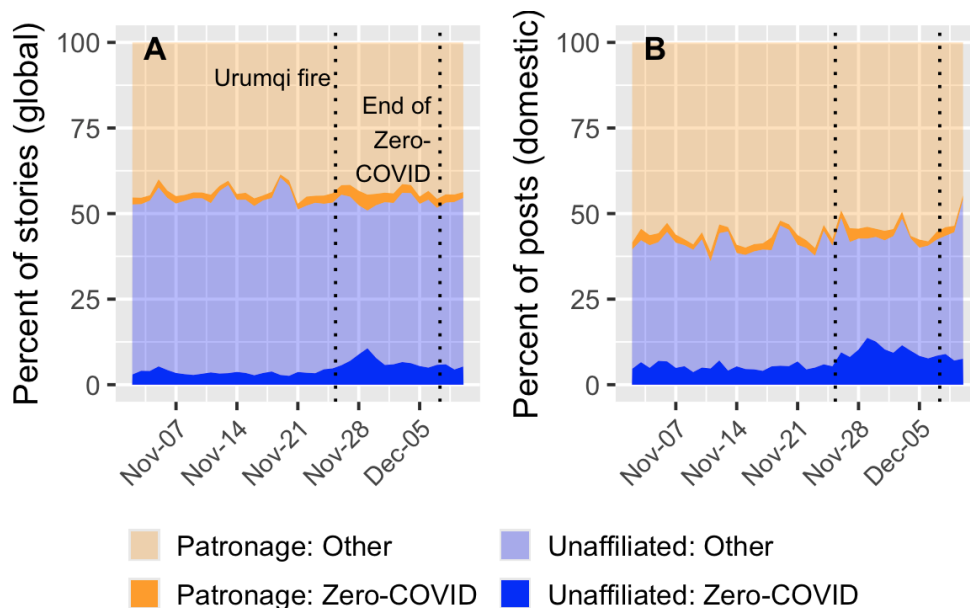


FIGURE 10. NEWS COVERAGE DURING THE ZERO-COVID PROTESTS NOV-DEC 2022. DOTTED LINES INDICATE THE BEGINNING AND END OF WIDESPREAD ZERO-COVID PROTESTS. PANEL A: PERCENT OF ALL STORIES PRODUCED. PANEL B: PERCENT OF POSTS IN DOMESTIC CIRCULATION.

the fire triggered widespread protests, this gap widened further—unaffiliated outlet constituted 80.2% of foreign-sourced Chinese-language protest coverage. When the regime needed domestic information control, patronage outlets did not report.

7. Discussion

This paper argues that authoritarian regimes can shape global news production by altering the economics of news supply to effect changes in overall market share. Across the new data we collected for this project, we find that the CCP uses media patronage to selectively sponsor news organizations with a pro-regime slant; patronage preserves existing pro-regime coverage; PRC-linked organizations compete with unaffiliated outlets by providing comparable levels of local and global news coverage; and PRC-linked organizations survive longer than outlets who do not receive state support. During crises, protest coverage from these PRC-linked organizations circulated far less widely than unaffiliated organizations.

These strategies help the CCP preserve pro-regime outlets while anti-regime outlets die

out due to existing global financial strains on local news. By using business transactions rather than overt repression or co-optation, authoritarian regimes diminish the risk of political, reputational, and financial harms while achieving country-wide changes in news production. Theoretically, these behaviors are consistent with the rise of “information autocrats” who seek to use lower-visibility tactics—in our case, market mechanisms—to undercut opposition without resorting to direct repression (Guriev and Treisman 2019).

While our study examines media patronage transnationally, the scope conditions of media patronage may exist in low-capacity authoritarian regimes and media institutions within their own borders. Where the state lacks capacity to enforce a command relationship with local media actors, or has limited ability to monitor local affairs, it is conceivable that they may turn to patronage to pursue state interests in the absence of the state’s unmediated presence. Indeed, such an approach bears similarities to recent work on the use of “proxy media” in an African context, where private intermediaries help to conceal the state’s role in influencing seemingly-independent press coverage (Rahmani 2025).

Our findings suggest that, in addition to nurturing media independence during boom times (Gentzkow, Glaeser and Goldin 2006; Petrova 2011), markets can also facilitate authoritarian influence when media markets are in crisis. In the absence of government support—a hallmark of media independence in democracies—media organizations under strain turn to other backers to ensure their survival. Autocrats’ willingness to fund foreign news, in turn, creates a structural advantage for autocracies in the market for news: pro-regime voices can find wealthy patrons while anti-regime voices cannot. This dynamic is associated with a rise in authoritarian influence on media production even outside territorial boundaries.

References

- Alonso, Ricardo and Gerard Padró i Miquel. 2025. “Competitive capture of public opinion.” *Econometrica* 93(4):1265–1297.
- Archer, Allison and Michelle Torres. 2022. Local News, Billionaire Owners, and Political Accountability. Technical report Working Paper.
- Bendor, Jonathan B, Amihai Glazer and Thomas H Hammond. 2000. *Theories of delegation in political science*. University of California, Irvine, Department of Economics.
- Besley, Timothy and Andrea Prat. 2006. “Handcuffs for the grabbing hand? The role of the media in political accountability.” *American Economic Review* 96(3):720–736.
- Bodrunova, Svetlana S, Anna Litvinenko and Kamilla Nigmatullina. 2021. “Who is the censor? Self-censorship of Russian journalists in professional routines and social networking.” *Journalism* 22(12):2919–2937.
- Chang, Keng-Chi, William R Hobbs, Margaret E Roberts and Zachary C Steinert-Threlkeld. 2022. “COVID-19 increased censorship circumvention and access to sensitive topics in China.” *Proceedings of the National Academy of Sciences* 119(4):e2102818119.
- Chen, Yuyu and David Y Yang. 2019. “The impact of media censorship: 1984 or brave new world?” *American Economic Review* 109(6):2294–2332.
- Chinese Radio Seattle. 2023. “The 11th Global Chinese Media Forum Was Grandly Held in Chengdu, China.” *Chinese Radio Seattle* Oct. 25.
URL: <https://chineseradioseattle.com/2023/10/25/world-chinese-media-forum-2023/>
- CNS. 2012. “The 9th Advanced Training Course for Overseas Chinese-Language Media Concludes in Beijing.” *China News Service* Sept. 16.
URL: <https://www.chinanews.com.cn/hr/2012/09-16/4186902.shtml>
- Cohen, Nicole S. 2015. “From pink slips to pink slime: Transforming media labor in a digital age.” *The Communication Review* 18(2):98–122.
- Crain, Matthew. 2009. “The rise of private equity media ownership in the United States: A public interest perspective.” *International Journal of Communication* 3:32.
- Cull, Nicholas John. 2009. *Public diplomacy: Lessons from the past*. Vol. 12 Figueroa Press Los Angeles.
- Deng, Yanhua and Kevin J O’Brien. 2013. “Relational repression in China: using social ties to demobilize protesters.” *The China Quarterly* 215:533–552.
- Dukalskis, Alexander, Saipira Furstenberg, Yana Gorokhovskaia, John Heathershaw, Edward Lemon and Nate Schenckan. 2022. “Transnational repression: Data advances, comparisons, and challenges.” *Political Research Exchange* 4(1):2104651.
- Earl, Jennifer, Thomas V Maher and Jennifer Pan. 2022. “The digital repression of social movements, protest, and activism: A synthetic review.” *Science Advances* 8(10):eabl8198.
- Ebsworth, Rowena. 2017. Not Wanting Want: The Anti-Media Monopoly Movement in Taiwan. In *Taiwan’s Social Movements under Ma Ying-jeou*. Routledge pp. 71–91.
- Economist, The. 2022. “How Chinese netizens breached the great firewall.” *The Economist* . Accessed: 2025-09-05.
URL: <https://www.economist.com/china/2022/12/08/how-chinese-netizens-breached-the-great-firewall>
- Egorov, Georgy, Sergei Guriev and Konstantin Sonin. 2009. “Why resource-poor dictators allow freer media: A theory and evidence from panel data.” *American political science*

Review 103(4):645–668.

- Eisenstadt, Shmuel Noah and Luis Roniger. 1984. *Patrons, clients and friends: Interpersonal relations and the structure of trust in society*. Cambridge University Press.
- Fan, Yingjie, Jennifer Pan and Jaymee Sheng. 2023. “Strategies of Chinese State Media on Twitter.” *Political Communication* pp. 1–22.
- Fittarelli, Alberto. 2024. “PAPERWALL: Chinese Websites Posing as Local News Outlets Target Global Audiences with Pro-Beijing Content.”. Accessed: 2025-12-12.
URL: <https://citizenlab.ca/2024/02/paperwall-chinese-websites-posing-as-local-news-outlets-with-pro-beijing-content/>
- Freelon, Deen. 2014. “On the interpretation of digital trace data in communication and social computing research.” *Journal of Broadcasting & Electronic Media* 58(1):59–75.
- Fu, King-wa, Chung-hong Chan and Michael Chau. 2013. “Assessing censorship on microblogs in China: Discriminatory keyword analysis and the real-name registration policy.” *IEEE internet computing* 17(3):42–50.
- Gandhi, Jennifer and Adam Przeworski. 2006. “Cooperation, cooptation, and rebellion under dictatorships.” *Economics & politics* 18(1):1–26.
- Gentzkow, Matthew, Edward L Glaeser and Claudia Goldin. 2006. The rise of the fourth estate. How newspapers became informative and why it mattered. In *Corruption and reform: Lessons from America’s economic history*. University of Chicago Press pp. 187–230.
- Gentzkow, Matthew and Jesse M Shapiro. 2008. “Competition and Truth in the Market for News.” *Journal of Economic perspectives* 22(2):133–154.
- Grossman, Guy, Yotam Margalit and Tamar Mitts. 2022. “How the ultrarich use media ownership as a political investment.” *The Journal of Politics* 84(4):1913–1931.
- Guriev, Sergei and Daniel Treisman. 2019. “Informational autocrats.” *Journal of economic perspectives* 33(4):100–127.
- Hamilton, James T. 2004. *All the news that’s fit to sell: How the market transforms information into news*. Princeton University Press.
- Hanley, Hans WA, Yingdan Lu and Jennifer Pan. 2025. “Across the firewall: Foreign media’s role in shaping Chinese social media narratives on the Russo-Ukrainian War.” *Proceedings of the National Academy of Sciences* 122(1):e2420607122.
- Hassan, Mai, Daniel Mattingly and Elizabeth R Nugent. 2022. “Political control.” *Annual Review of Political Science* 25(1):155–174.
- Hsiao-hwa, Hsia. 2024. “Australian Journalist Vicky Xu Starts a New Life in Taiwan.”.
URL: <https://www.rfa.org/english/news/china/vicky-xu-australia-taiwan-06062024164411.html>
- Huang, Haifeng. 2015. “International knowledge and domestic evaluations in a changing society: The case of China.” *American Political Science Review* 109(3):613–634.
- Huang, Haifeng and Yao-Yuan Yeh. 2019. “Information from abroad: Foreign media, selective exposure and political support in China.” *British Journal of Political Science* 49(2):611–636.
- Kao, Jay C. 2026. “How the Pro-Beijing Media Influences Voters: Evidence from a Randomized Field Experiment.” *American Political Science Review* p. 1–20.
- Kern, Holger Lutz. 2011. “Foreign media and protest diffusion in authoritarian regimes: The case of the 1989 East German revolution.” *Comparative Political Studies* 44(9):1179–1205.

- Kern, Holger Lutz and Jens Hainmueller. 2009. "Opium for the masses: How foreign media can stabilize authoritarian regimes." *Political Analysis* 17(4):377–399.
- King, Gary, Jennifer Pan and Margaret E Roberts. 2013. "How censorship in China allows government criticism but silences collective expression." *American political science Review* 107(2):326–343.
- Kitschelt, Herbert, Steven I Wilkinson et al. 2007. "Citizen-politician linkages: an introduction." *Patrons, clients, and policies: Patterns of democratic accountability and political competition* 2007:1–49.
- Lai, H. M. 1987. The Chinese-American Press. In *The Ethnic Press in the United States: A Historical Analysis and Handbook*, ed. Sally M. Miller. Greenwood Press.
- Lai, Him Mark. 2009. "Him Mark Lai Papers." Archival collection. AAS ARC 2000/80, The Ethnic Studies Library, University of California, Berkeley.
- Le Mens, Gaël and Aina Gallego. 2025. "Positioning political texts with large language models by asking and averaging." *Political Analysis* 33(3):274–282.
- Lee, Francis LF and Angel MY Lin. 2006. "Newspaper editorial discourse and the politics of self-censorship in Hong Kong." *Discourse & Society* 17(3):331–358.
- Levitsky, Steven and Lucan A Way. 2010. *Competitive authoritarianism: Hybrid regimes after the Cold War*. Cambridge University Press.
- Library of Congress. 2026. "Chronicling America: Historic American Newspapers." <https://www.loc.gov/collections/chronicling-america/>. Accessed 1 February 2026.
- Lucas, Edward. 2022. "How Autocrats Undermine Media Freedom." *Journal of Democracy* 33(1):131–146.
- Ma, Yan. 2003. "Chinese American newspapers and periodicals in the United States and their web presence." *Serials Review* 29(3):179–198.
- Manion, Melanie. 2023. *Political selection in China: Rethinking foundations and findings*. Cambridge University Press.
- McMillan, John and Pablo Zoido. 2004. "How to subvert democracy: Montesinos in Peru." *Journal of Economic perspectives* 18(4):69–92.
- McPherson-Smith, Oliver. 2021. "Diversification, Khashoggi, and Saudi Arabia's public investment fund." *Global Policy* 12(2):190–203.
- Mena, Paul. 2020. "Cleaning up social media: The effect of warning labels on likelihood of sharing false news on Facebook." *Policy & internet* 12(2):165–183.
- Metzger, Zach and the Medill Local News Initiative. 2025. "The State of Local News 2025." <https://localnewsinitiative.northwestern.edu/projects/state-of-local-news/2025/report/>. Accessed 1 February 2026.
- Moss, Dana M. 2016. "Transnational repression, diaspora mobilization, and the case of the Arab Spring." *Social Problems* 63(4):480–498.
- Muno, Wolfgang. 2010. Conceptualizing and measuring clientelism. In *Paper to be presented at the workshop on Neopatrimonialism in Various World Regions*, giga German Institute of Global and Area Studies, Hamburg.
- Nassetta, Jack and Kimberly Gross. 2020. "State media warning labels can counteract the effects of foreign misinformation." *Harvard Kennedy School Misinformation Review* .
- Nichter, Simeon. 2008. "Vote buying or turnout buying? Machine politics and the secret ballot." *American political science review* 102(1):19–31.
- Nielsen, Rasmus Kleis. 2015. *Local journalism: The decline of newspapers and the rise of*

- digital media*. Bloomsbury Publishing.
- Nugent, Elizabeth R and Alexandra A Siegel. 2026. "How Exiles Mobilize Domestic Dissent." *The Journal of Politics* 88(2):000–000.
- OCAO. 2007. "The 3rd Advanced Training Course for Overseas Chinese-Language Media' Opens in Beijing." *Overseas Chinese Affairs Office of the State Council* Nov. 26.
URL: <https://www.gqb.gov.cn/news/2007/1126/1/7210.shtml>
- of Professional Journalists, Society. 2014. "SPJ Code of Ethics." <https://www.spj.org/ethicscode.asp>. Revised September 6, 2014; Accessed: 2026-04-07.
- Ornstein, Joseph T, Elise N Blasingame and Jake S Truscott. 2025. "How to train your stochastic parrot: Large language models for political texts." *Political Science Research and Methods* 13(2):264–281.
- Ownby, David. 2008. *Falun Gong and the future of China*. Oxford University Press.
- Pan, Jennifer. 2017. "How market dynamics of domestic and foreign social media firms shape strategies of internet censorship." *Problems of Post-Communism* 64(3-4):167–188.
- Pan, Jennifer, Xu Xu and Yiqing Xu. 2023. "Disguised repression: Targeting opponents with non-political crimes to undermine dissent."
- Peisakhin, Leonid and Arturas Rozenas. 2018. "Electoral effects of biased media: Russian television in Ukraine." *American journal of political science* 62(3):535–550.
- Petrova, Maria. 2011. "Newspapers and parties: How advertising revenues created an independent press." *American Political Science Review* 105(4):790–808.
- Qin, Bei, David Strömberg and Yanhui Wu. 2018. "Media bias in China." *American Economic Review* 108(9):2442–2476.
- Rahmani, Bardia. 2025. "How autocrats use proxies to control the media." *Journal of Democracy* 36(2):92–105.
- Ramzy, Austin. 2020. "China Sentences Hong Kong Bookseller Gui Minhai to 10 Years in Prison." *The New York Times* . Accessed: 2025-09-05.
URL: <https://www.nytimes.com/2020/02/25/world/asia/gui-minhai-china-hong-kong-swedish-bookseller.html>
- Repnikova, Maria. 2018. "Contesting the state under authoritarianism: Critical journalists in China and Russia." *Comparative Politics* 51(1):41–60.
- Repnikova, Maria. 2022. *Chinese Soft Power*. New York.
- Reporters Without Borders. 2025. "World Press Freedom Index 2025." <https://rsf.org/en/world-press-freedom-index-2025>. Accessed September 3, 2025.
- Rosenfeld, Bryn and Jeremy Wallace. 2024. "Information politics and propaganda in authoritarian societies." *Annual Review of Political Science* 27.
- Saunders, Frances Stonor. 2000. *Who paid the piper?: the CIA and the cultural Cold War*. Granta Books.
- Scott, James C. 1972. "Patron-client politics and political change in Southeast Asia." *American political science review* 66(1):91–113.
- Shih, Victor C. 2007. "Factions and finance in China." (*No Title*) .
- Shirikov, Anton. 2024. "Rethinking propaganda: how state media build trust through belief affirmation." *The Journal of Politics* 86(4):1319–1332.
- Shirk, Susan L. 2010. *Changing media, changing China*. Oxford University Press.
- Stern, Rachel E and Jonathan Hassid. 2012. "Amplifying silence: uncertainty and control parables in contemporary China." *Comparative Political Studies* 45(10):1230–1254.

- Stockmann, Daniela and Mary E Gallagher. 2011. "Remote control: How the media sustain authoritarian rule in China." *Comparative political studies* 44(4):436–467.
- Tella, Rafael Di and Ignacio Franceschelli. 2011. "Government advertising and media coverage of corruption scandals." *American Economic Journal: Applied Economics* 3(4):119–151.
- Thunø, Mette. 2017. China's new global position: Changing policies toward the Chinese diaspora in the twenty-first century. In *China's rise and the Chinese overseas*. Routledge pp. 184–208.
- Truex, Rory. 2019. "Focal points, dissident calendars, and preemptive repression." *Journal of Conflict Resolution* 63(4):1032–1052.
- Tsourapas, Gerasimos. 2015. "Why do states develop multi-tier emigrant policies? Evidence from Egypt." *Journal of Ethnic and Migration Studies* 41(13):2192–2214.
- Tsourapas, Gerasimos. 2021. "Global autocracies: Strategies of transnational repression, legitimization, and co-optation in world politics." *International Studies Review* 23(3):616–644.
- United Nations, Department of Economic and Social Affairs, Population Division. 2024. "International Migrant Stock 2024: Key Facts and Figures." .
URL: <https://www.un.org/development/desa/pd/content/international-migrant-stock>
- Weber, Matthew S and Peter R Monge. 2017. "Industries in turmoil: Driving transformation during periods of disruption." *Communication Research* 44(2):147–176.
- Willnat, Lars and Jason Martin. 2020. "Foreign correspondents—An endangered species?" *The global journalist in the 21st century* pp. 495–510.
- World Chinese Media. 2001. "The World Association of Chinese-Language Mass Media Participates in the Inaugural World Chinese-Language Media Forum Held in Nanjing." *World Chinese Media* Sept. 17.
URL: <https://www.worldchinesemedia.com/2001/09/17/11926/>
- Xi, Jinping. 2015. "Consolidate and Develop the Broadest Patriotic United Front to Provide Broad Force Support for Realizing the Chinese Dream." Speech at the Central United Front Work Conference, <https://news.12371.cn/2015/05/20/ARTI1432111597695781.shtml>. Published on Communist Party Member Net ().
- Xu, Yiqing. 2017. "Generalized synthetic control method: Causal inference with interactive fixed effects models." *Political Analysis* 25(1):57–76.
- Xue, Zhao, Hanyu Zhao, Sha Yuan and Yequan Wang. 2022. "WuDaoCorpora Text.".
- Yin, Qi, Zhenhan Fu and Shiyu Zheng. 2024. "Meso news-space in China: peripheral news production of platform journalism." *Digital Journalism* 12(5):680–699.

Appendix

Table of Contents

A1 Data	A-2
A1.1 Global Dataset Data Availability	A-2
A1.2 US Census Dataset Construction	A-7
A1.3 Data Cleaning	A-9
A1.4 Alternate Estimates for Chinese News Consumption	A-11
A1.5 Qualitative Interviews	A-14
 A2 Methods	 A-15
A2.1 Coding rules	A-15
A2.2 Article classification	A-16
A2.3 Local News Classification	A-17
A2.4 Error Propagation	A-18
A2.5 Censored Topics	A-19
A2.6 Event Detection Pipeline	A-20
A2.7 Event Framing	A-21
 A3 Additional Analysis	 A-22
A3.1 Repression and Exit: The Case of Hong Kong	A-22
A3.2 Local News Stance	A-23
A3.3 Webpage linkages	A-24

A1. Data

A1.1. Global Dataset Data Availability

Outlet		Founded	Wayback From	Scrape From	Patronage
ID	Country				
1	USA	1851	2012	2012	0
2	United Kingdom	1888	2003	2005	0
3	USA	1889	2008	2014	0
4	Malaysia	1910	1998	2023	1
5	Philippines	1919	2012	2011	1
6	Korea	1920	2014	2012	0
7	Malaysia	1923	2015	2016	1
8	Malaysia	1929	2002	2007	1
9	Malaysia	1936	2024	2009	0
10	Malaysia	1946	2000	2017	1
11	Thailand	1950	2014	2017	1
12	Malaysia	1952	2009	2014	0
13	Brazil	1960	2004	2011	1
14	Korea	1965	2004	2008	1
15	Malaysia	1968	2011	2018	1
16	USA	1972	2021	2020	0
17	Philippines	1973	2016	2023	1
18	Malaysia	1976	2002	2016	1
19	USA	1976	1999	2019	0
20	USA	1976	2000	2020	0
21	Portugal	1977		2021	1
22	USA	1979	2000	2019	1
23	South Korea	1980	2018	2024	1
24	Indonesia	1981		2020	1
25	Philippines	1981	2014	2019	1
26	USA	1981	2021	2021	1
27	USA	1982	2004	2017	0
28	Singapore	1983	1998	2016	1
29	Australia	1987	2001	2019	0
30	Malaysia	1987	2003	2021	1
31	USA	1989	2008	2021	0
32	USA	1991	2020	2020	1
33	USA	1990	2019	2019	1
34	USA	1990	2010	2017	1
35	USA	1990	2004	2012	1
36	USA	1990	2007	2022	1
37	New Zealand	1991	2004	2022	1
38	Australia	1992	2008	2024	0

Outlet		Founded	Wayback From	Scrape From	Patronage
ID	Country				
39	Japan	1992	1997	2006	1
40	Netherlands	1992	2002	2023	1
41	Australia	1994	2014	2013	0
42	New Zealand	1994		2016	1
43	USA	1994	2003	2024	1
44	USA	1994	2015	2010	1
45	Canada	1995	1998	2023	0
46	USA	1995	2017	2017	1
47	Japan	1996	2021	2022	1
48	Germany	1997	2022	2017	1
49	USA	1997	2008	2017	0
50	USA	1997	1998	2014	1
51	USA	1997	1998	2003	0
52	Canada	1998	1999	2006	0
53	New Zealand	1998	2008	2021	1
54	USA	1998	2004	2021	0
55	Canada	1999	2005	2004	0
56	Hungary	1999	2013	2014	1
57	Japan	1999	2020	2003	1
58	Malaysia	1999		2005	0
59	Romania	1999	2016	2013	1
60	USA	1999	2021	2018	1
61	Cambodia	2000	2010	2018	1
62	England	2000	2016	2022	0
63	Hungary	2000	2001	2013	1
64	Indonesia	2000		2018	0
65	Indonesia	2000	2011	2024	1
66	Indonesia	2000	2010	2020	1
67	Peru	2000		2021	1
68	Romania	2000	2013	2013	1
69	Russia	2000	2014	2014	1
70	United Kingdom	2000	2017	2019	1
71	USA	2000	2000	2001	0
72	Global	2000	2004	2023	1
73	Canada	2001	2003	2023	1
74	Canada	2001	2011	2021	1
75	Canada	2001	2001	2024	0
76	Japan	2002	2017	2016	1
77	New Zealand	2001	2003	2018	1
78	USA	2001	2017	2019	0
79	USA	2001	2001	2001	0
80	USA	2001	2008	2002	0

Outlet		Founded	Wayback From	Scrape From	Patronage
ID	Country				
81	Canada	2002	2015	2015	1
82	Spain	2002	2003	2016	1
83	Canada	2000	2000	2003	0
84	Canada	2002	2002	2023	0
85	Malaysia	2002	2002	2012	1
86	Netherlands	2003	2008	2018	1
87	New Zealand	2003	2015	2017	1
88	Spain	2001		2010	1
89	USA	2001	2001	2019	0
90	Australia	2001	2001	2021	0
91	Brunei	2004	2009	2019	0
92	Italy	2004	2021	2020	1
93	Singapore	2004	2009	2008	1
94	USA	2004	2017	2016	1
95	USA	2004	2010	2010	0
96	USA	2004	2005	2017	0
97	France	1983	2009	2009	1
98	Greece	2005	2009	2009	1
99	Hungary	2005	2007	2011	1
100	South Africa	2005	2017	2017	1
101	South Korea	2005	2020	2020	1
102	USA	2005	2005	2011	0
103	Australia	2006	2018	2017	1
104	Australia	2006	2007	2010	0
105	Poland	2006	2006	2006	1
106	South Korea	2006	2009	2022	1
107	Canada	2007	2004	2017	0
108	Myanmar	2007	2012	2012	1
109	South Korea	2007	2019	2023	1
110	Italy	2008	2009	2014	1
111	Portugal	2006		2013	1
112	South Africa	2008		2008	1
113	Australia	2009		2021	1
114	Brazil	2009		2011	1
115	Canada	2009	2009	2020	1
116	United Arab Emirates	2009	2010	2010	1
117	Australia	2001	2001	2011	1
118	Canada	2010	2012	2020	0
119	Spain	2010	2008	2010	1
120	United Kingdom	2010	2012	2010	1
121	USA	2010	2021	2019	1
122	USA	2010	2010	2010	1

Outlet		Founded	Wayback From	Scrape From	Patronage
ID	Country				
123	Argentina	2010	2010	2017	1
124	Australia	2006	2006	2018	1
125	Canada	2008	2008	2017	1
126	France	2011	2011	2011	1
127	Italy	2011	2012	2017	1
128	Philippines	2011	2016	2020	1
129	Australia	2011	2019	2012	1
130	Madagascar	2012	2024	2018	1
131	Australia	2013	2015	2022	1
132	France	2014		2023	1
133	Peru	2013	2018	2016	0
134	Thailand	2013	2013	2018	1
135	Australia	2014	2005	2016	0
136	Indonesia	2014	2018	2018	1
137	Mexico	2014		2021	1
138	United Kingdom	2002	2013	2018	1
139	Canada	2015	2011	2016	0
140	Chile	2015	2015	2019	1
141	Singapore	2007	2007	2015	0
142	The Netherlands	2015	2015	2022	1
143	USA	2011	2011	2021	0
144	Australia	2016	2017	2016	0
145	Brazil	2016	2023	2021	1
146	Egypt	2012	2013	2020	1
147	New Zealand	2016	2018	2019	1
148	Pakistan	2016		2021	1
149	South Korea	2016	2016	2017	1
150	Sweden	2009	2015	2018	1
151	USA	2015	2021	2015	0
152	Cambodia	2018	2019	2023	1
153	Canada	2017	2011	2017	1
154	Singapore	2011	2011	2018	0
155	Germany	2020		2020	0
156	New Zealand	2020	2021	2020	1
157	USA	2020	2015	2020	0
158	Greece	2017	2023	2017	1
159	United Kingdom	2022		2022	0
160	Cambodia	2022	2022	2022	1
161	Hong Kong	2022		2022	0
162	Global	2022		2023	0
163	Australia	2014	2014	2017	1
164	Australia	2021	2021	2021	0

Outlet		Founded	Wayback From	Scrape From	Patronage
ID	Country				
165	Australia	2017	2017	2020	0
166	Cambodia	2017	2017	2017	1
167	Cambodia	1992	2016	2010	1
168	Canada		2021		0
169	Canada	2005	2010	2015	1
170	Canada	2008	2008	2008	1
171	Canada	2002	2004	2024	0
172	Canada	2023		2023	1
173	Indonesia	2007	2012	2020	1
174	Italy	2006	2018	2019	1
175	Italy	2004	2017	2010	1
176	Japan	2010	2010	2015	1
177	Japan	2018		2018	1
178	Malaysia	2012		2012	1
179	Malaysia	2017		2017	0
180	Netherlands	2014	2015	2016	1
181	Singapore	2014	2014	2014	0
182	Singapore	2018	2018	2018	0
183	Thailand	2024	2024	2024	0
184	Thailand	2008	2012	2009	1
185	Ukraine	2002	2006	2016	1
186	USA	2001	2015	2014	0
187	USA	1992	2021	2017	1
188	USA	2014	2015	2014	1
189	USA	1965	2001	2021	0
190	USA	2021	2021	2021	0
191	Venezuela	1999		2020	1
192	Global	1997	1998	2022	0
193	Global	1998	1999	2017	0

A1.2. US Census Dataset Construction

Data on historical Chinese-language newspapers are drawn from four sources. The first is the Him Mark Lai Collection at U.C. Berkeley, a record of Chinese-language newspaper clippings and newspapers records recorded by Him Mark Lai (1925 - 2009), a San Francisco-based historian of Chinese life in the United States (Lai 2009). The second is the Library of Congress’ “Chronicling America” collection of historic newspapers in the United States, which maintains a limited database of Chinese-American newspapers (Library of Congress 2026). The third is Ma (2003)’s excellent database of Chinese American Newspapers and Periodicals in the U.S. in 2003. The fourth is a 2014 P.R. Newswire index of Chinese-language media organizations in the United States.¹ Many outlets appeared in two or more of these sources.

To determine whether a newspaper was active in the United States, we adopted broader criteria than for the Global Dataset due to data scarcity concerns. In many cases it was challenging to determine conclusively where an organization was based and what it published about. In the end, we adopted the rule that if an outlet was: (A) publishing news, AND (B1) had staff based in the United States; OR (B2) published news about United States affairs drawn from a local correspondent, we considered the outlet active in the United States. If we could not view historical content but an outlet was described as a “newspaper” or “news organization”, we assumed that it produced general interest news. If an outlet was absorbed by another outlet, we considered that outlet closed. If an outlet remained active outside the U.S. (e.g. in Hong Kong) but the U.S. operations had ceased, we also considered an outlet closed. If an outlet once had a U.S. correspondent but we were unable to determine whether the correspondent had left, we considered it “active” if the overall organization was still producing news. If an outlet was a radio station dedicated to Chinese-language content we included it; if an outlet was a radio channel (within a non-Chinese language station) we excluded it. Finally, as with the Global Dataset, if the historical news organization openly acknowledged ownership by a state actor (e.g. *People’s Daily*), we excluded it.

To determine the date of closure, we adopted the following criteria:

1. If the date of outlet’s closure was published in archival materials or announced publicly by the outlet itself, we used that. If not:
2. A team of two research assistants (RAs) sought to determine via open source methods the date an outlet closed. If that date was available, we used it. If not:
3. We checked the Wayback Machine for the last date an outlet’s website was available, and used that as the date of closure. If the outlet did not appear in the Wayback Machine:
4. We used the last issue for each newspaper saved in archival resources, or the date 2014 if the PR newswire report was the last mention of the outlet on the internet. If one collection contained a later date than the other, we used the later date as an outlet’s official end date.

¹<https://www.webwire.com/prnewswire/PDF/384.pdf>, Accessed March 25 2026

Because the criteria for inclusion in the historical dataset was broader than the news-domain dataset, the two datasets have different counts of active outlets in the U.S. during 2024, the study’s onset. Of the 93 outlets active in the U.S. as of the end of 2024, 38 are in our news domain dataset while 55 are not. The remaining 55 were not included for several reasons. In nine cases, the outlet was based outside the U.S. (usually Hong Kong or Taiwan) but maintained a U.S.-based correspondent. An additional nine outlets produced, or could not be disproven as having produced, general interest news at some point prior to 2024 but no longer produce general interest news today. Eighteen outlets had no digital presence or produced only digital scanned copies of print newspapers, which were not scrapeable and thus excluded from the Global Dataset. The remaining 19 outlets appear to have been active in 2024 but were not included in the initial domain gathering process because they were not surfaced using the methods described in Section 3.

A1.3. Data Cleaning

For each full-text article we adopted the following cleaning and normalization steps:

- Removed stray encoding characters (e.g. sp?).
- Converted traditional Chinese characters to simplified characters
- Removed articles with impossible publication dates (e.g. prior to the invention of the internet or in the future).
- Filtered out articles containing less than 10% Chinese characters.

For each archived webpage snapshot, all hyperlinks (`<a>`) and image links (`<img>`) were extracted. Each link was recorded with associated metadata, including the snapshot date, the link itself, the element type (anchor or image), the order in which the tag appeared on the page, and any human-readable anchor text.

To ensure data quality, we applied a series of data-cleaning steps:

- Removed all self-references (e.g., homepage links to the same outlet).
- Removed vestigial Wayback Machine or Archive.org links.
- Removed JavaScript anchors and invalid domains (restricted to valid alphanumeric domain names with a top-level domain).
- Removed external links where a domain was referenced only once across the entire corpus (suggesting they functioned as self-links).
- Expanded shortened URLs using a comprehensive list of URL shorteners;² replaced the original shortened link with the expanded form, and re-ran domain lookups on the expanded addresses.
- Identified and removed advertising links, social media profile links, and social-sharing links using regular-expression matching.
- Consolidated domains where only the top-level domain differed (e.g., treating `example.com` and `example.net` as equivalent self-links).

To assign countries of origin to extracted URL domains, we employed a tiered procedure. Each domain was first queried against the most authoritative source available; if that approach failed, we proceeded sequentially to the next. Specifically:

1. Domains were first manually checked against the ICANN database.
2. If no ICANN record was available, we used the DomainTools API to query country-level registration data.
3. If DomainTools returned no result, we parsed WHOIS records directly.

²see <https://github.com/PeterDaveHello/url-shorteners/blob/master/list>

4. If WHOIS records were unavailable, we parsed Registration Data Access Protocol (RDAP) records.
5. If RDAP data could not be obtained, we consulted WHOIS history records.
6. As a last resort, we inferred the country based on the domain suffix (top-level domain).

Finally, we manually corrected a small number of high-profile cases where automated assignment was misleading. For example, `temu.com` was improperly assigned to the United States, and `weibo.com` was assigned as “Not Applicable.”

A1.4. Alternate Estimates for Chinese News Consumption

To provide alternate estimates of Chinese news consumption globally, we also look to two additional sources of Chinese-language media consumption: website visits and app downloads.

For website visits, we take the top 50 websites visited by Chinese users in November 2025³, and subset to all websites whose interfaces are Chinese by default when accessed from a U.S. IP address. This criterion yields 17 domains. We then use the Semrush Traffic & Analytics Toolkit to download country-by-country annual traffic data for unique visitors to each of the 17 domains from 2017-2024. We then subset to domains present in the data in both 2020 and 2024, and average the unique visitors to each site.⁴ The 17 domains were: feishu.cn, taobao.com, doubao.com, 163.com, xiaohongshu.com, weibo.com, 360.com, tmall.com, 1688.com, quark.cn, douyu.com, huya.com, douban.com, kdocs.cn, larkoffice.com, sohu.com, alipay.com.

For app downloads, we take the top 20 apps in China (by monthly active users), and subset to the 16 apps with current and historic country-by-country download data available from AppFigures.⁵ We then average across all 16 apps in each country to get the number of annual app downloads to users located in each country.

The results are largely consistent with the U.N. migration data.

³<https://www.semrush.com/trending-websites/cn/all>

⁴<https://www.semrush.com/>

⁵<https://appinchina.co/market/apps/>; <https://appfigures.com/>

TABLE A2. TOP COUNTRIES BY CHINESE POPULATION AND MEDIA REACH (2020)

Country	UN Rank	Traffic Rank	App Rank	Chinese Pop.	Unique Visitors	Downloads
United States	1	1	NA	2,255,170	6,486,587	NA
Japan	2	3	NA	917,070	2,309,252	NA
South Korea	3	6	NA	751,745	998,007	NA
Canada	4	5	NA	721,158	1,274,896	NA
Australia	5	8	NA	632,690	884,418	NA
Singapore	6	4	NA	509,007	2,090,483	NA
Italy	7	18	NA	241,671	201,844	NA
United Kingdom	8	10	NA	213,560	675,855	NA
Spain	9	17	NA	190,373	230,712	NA
Germany	10	7	NA	152,761	896,216	NA
New Zealand	11	20	NA	144,287	177,153	NA
France	12	12	NA	114,237	452,734	NA
Thailand	13	11	NA	100,327	580,086	NA
Netherlands	14	16	NA	73,837	260,688	NA
Indonesia	15	15	NA	47,633	265,901	NA
Sweden	16	23	NA	35,677	113,554	NA
Brazil	18	14	NA	32,715	273,688	NA
Malaysia	19	2	NA	30,606	3,057,346	NA
Philippines	20	13	NA	27,676	274,047	NA
Russia	21	9	NA	24,009	871,319	NA
Hungary	22	35	NA	17,648	28,386	NA
Argentina	23	28	NA	17,325	43,216	NA
Chile	24	26	NA	14,364	46,112	NA
Venezuela	25	37	NA	13,071	5,684	NA
Portugal	26	33	NA	12,000	34,368	NA
Mexico	27	21	NA	10,667	145,018	NA
Egypt	28	29	NA	7,407	41,925	NA
Ukraine	29	19	NA	6,668	192,207	NA
South Africa	30	34	NA	6,371	30,193	NA
Peru	31	31	NA	4,355	38,939	NA
Greece	32	32	NA	3,462	37,840	NA
Romania	33	25	NA	2,323	51,403	NA
Brunei	34	36	NA	2,238	10,246	NA
Madagascar	35	38	NA	2,193	3,086	NA
Cambodia	36	27	NA	1,627	45,468	NA
Poland	37	22	NA	1,069	134,173	NA
Pakistan	38	24	NA	311	54,013	NA

TABLE A3. TOP COUNTRIES BY CHINESE POPULATION AND MEDIA REACH (2024)

Country	UN Rank	Traffic Rank	App Rank	Chinese Pop.	Unique Visitors	Downloads
United States	1	1	1	2,488,775	11,661,178	502077.444
Japan	2	3	5	809,893	3,943,075	87372.636
Canada	3	10	12	771,973	1,474,382	38734.091
South Korea	4	7	4	756,684	1,934,805	90370.500
Australia	5	6	9	656,264	1,935,462	50587.455
Singapore	6	4	7	533,745	3,558,922	63095.167
Italy	7	20	17	240,895	452,955	31517.857
Spain	8	16	19	198,909	715,942	23680.286
United Kingdom	9	11	8	179,337	1,254,851	62499.000
Germany	10	12	11	168,314	1,073,070	45214.429
New Zealand	11	21	26	145,906	416,954	8647.000
France	12	17	13	120,595	652,260	37986.750
Thailand	13	5	3	85,764	2,014,758	102455.364
Indonesia	14	9	10	82,568	1,604,619	48229.800
Netherlands	15	15	23	82,477	856,870	12798.750
Sweden	16	23	27	40,898	191,771	7012.000
Egypt	18	31	24	34,659	95,705	11724.000
Malaysia	19	2	6	30,524	4,688,864	87104.667
Philippines	20	13	18	27,676	991,236	31083.000
Brazil	21	14	16	24,462	976,185	33643.000
Hungary	22	35	35	20,818	65,961	2232.800
Argentina	23	30	30	19,955	98,769	4409.875
Russia	24	8	2	13,724	1,900,487	160491.750
Portugal	25	32	29	13,500	91,559	5224.286
Chile	26	26	32	13,126	124,790	3862.444
Venezuela	27	36	NA	12,470	37,424	NA
Mexico	28	18	20	12,070	613,278	18903.429
Ukraine	29	22	15	6,757	272,266	36704.000
South Africa	30	33	28	6,189	73,524	6330.100
Peru	31	29	33	4,560	103,535	2745.556
Greece	32	34	34	3,678	69,570	2327.625
Pakistan	33	24	22	3,568	174,199	13833.889
Brunei	34	38	36	2,397	4,769	691.900
Madagascar	35	37	NA	2,382	11,319	NA
Romania	36	25	31	2,359	152,490	4305.429
Cambodia	37	27	14	1,721	114,568	37275.818
Poland	38	19	25	703	554,059	9020.750

A1.5. Qualitative Interviews

Interviews were conducted in Mandarin, English or both. Between October 2024 and February 2026 we spoke with 10 current or former representatives from diasporic media organizations, which yielded 7 oral interviews, 3 written interviews, 2 organizations declining an interview and the remainder failing to respond (see Table A4). We contacted media or-

TABLE A4. OVERVIEW OF INTERVIEWS CONDUCTED

Interview	Date Conducted	Mode
Interview A	October 2024	In-person (oral)
Interview B	November 2024	In-person (oral)
Interview C	March 2025	Written
Interview D	June 2025	Phone
Interview E	July 2025	Written
Interview F	July 2025	Phone
Interview G	August 2025	In-person
Interview H	November 2025	Phone
Interview I	November 2025	Phone
Interview J	February 2026	Written

ganizations via email accounts or submission portals located on news organization websites. We contacted a total of 151 news organizations for interviews (all outlets for which contact information was available); each oral interview lasted between 30 and 60 minutes, and were conducted on Zoom, telephone, and in-person, and were conducted in either English, Mandarin Chinese, or a mixture of both. Our questions concerned organizational policies and practices, including news selection, target audiences, editorial decisions, and partnerships with domestic Chinese organizations. We applied for IRB approval from Stanford University (Protocol number: 76683) and were informed that because our study investigates organizations, it does not meet the definition of Human Subjects Research as codified in federal regulations at 45 CFR 46 and thus IRB approval was not required.

A2. Methods

A2.1. Coding rules

In this project, we are examining how diasporic Chinese media outlets (Chinese-language media organizations based outside of mainland China) report on China. We will ask you to read the titles of several news articles, as well as the first paragraph if necessary, and we will ask you to evaluate each article on a number of different dimensions (listed and described below). If the lede contains additional details that are unrelated to the title of the article, do not consider those additional details.

You will be filling in four columns:

- Related to China (1, 0).
- Related to Host Country (1, 0).
- If Related to China $\neq$ 0: Stance_C (-1, 0, 1).

1. Related to China: • (1) if the article is related to China (the nation, its government, institutions or policies, territory, or any event that occurs in mainland China, or a multinational organization in which the country participates).

- “Institution” refers to any state-owned or state-operated institution, and any individuals or policies affiliated with that institution.
- State-owned Chinese companies, such as steel manufacturers or banks, are considered a “Chinese” institution even if operating abroad.
- If an individual appears in a formal capacity representing China (the country), then code (1).

- (0) if the article is unrelated to China.

2. Related to Host Country: • (1) if the article is related to the host country (the nation, its government, institutions or politics, territory, or any event that occurs in the host country, or a multinational organization in which the country participates).

- “Institution” same as above; “individuals” representing the host country same as above.

- (0) if the article is unrelated to the host country.

3. Stance_C (only if Related to China $\neq$ 0): • (1) if the article is supportive of the Chinese Communist Party, the Chinese government, its leaders, its institutions, or its policies; OR likely to leave a reader with a positive impression of the People’s Republic of China. (Note: Chinese culture, language, or history is not necessarily related to the Chinese government.)

- (-1) if the article is critical of the Chinese Communist Party, the Chinese government, its leaders, its institutions, or its policies; OR likely to leave the reader with a negative impression of the People’s Republic of China.
- (0) if the article is neutral (neither supportive nor critical), or unlikely to positively or negatively influence the reader’s impression of the People’s Republic of China.

A2.2. Article classification

Our hand-labeled sample was generated via stratified sampling by outlet and year. To ensure that our sample was not dominated by outlets with large numbers of articles, we took 1% or 50 articles from each outlet (whichever is greater), with a maximum of 500 articles per outlet. Two trained, independent human coders read the title and lede (first 500 characters) of all articles, determining both whether an article was related to the People’s Republic of China (1 or 0) and, if related, the stance of that article’s stance toward China (-1, 0, or 1). After refining the codebook (see Appendix A2.1), we achieved an intercoder agreement of 95.2% on China relevance and 81.8% on stance toward China ($n = 100$), indicating sufficient construct validity. These coder then coded a total of 1,600 articles for China relevance and stance toward China. Because only about a third of articles relate to China, our gold-standard set of China-stance labels was relatively limited ($n = 654$).

In the second step, we fine-tuned a supervised learning model (DeBERTa) based on the human coded data to classify the remainder of the dataset. After testing multiple models, we found the best performing model to be Erlangshen-DeBERTa-v2-320M-Chinese,⁶ a DeBERTa model trained on the WuDao Corpora with whole world masking.⁷ Using a 80/20 train-test split, we achieved a binary F1 of 0.91 on our holdout set ($n = 250$) on the China relevance outcome, indicating strong model performance.

In our third step, we used large language models to assess the stance of China-related articles. Due to the small number of hand-labeled China-related articles, we felt using DeBERTa model to apply surrogate stance labels ran the risk of overfitting. We thus turned to large language models (LLMs). While the use of large language models in data-labeling is still relatively new, a few best practices have emerged: use a set checkpoint, set model randomness (“temperature”) to zero, validate against human labels, and run the model repeatedly on the same data to ensure reproducibility (Le Mens and Gallego 2025; Ornstein, Blasingame and Truscott 2025). After following these best practices, we found GPT to be the best performing model family. Yet even after adopting these best practices and experimenting with multiple checkpoints, we found that the best performing model checkpoint at a reasonable cost (gpt-4o-mini-2024-07-18) still performed inadequately when detecting stances, displaying an “agreement bias” where it affirmed whatever category label was posed to it in the prompt. To compensate, we collapsed our three stance categories into two, with “1” representing “positive” coverage while “0” represented “negative or neutral”. After this amendment, GPT’s labels achieved satisfactory agreement with human labels (F1=0.89). Because querying GPT for all China-related articles in our dataset would be prohibitively expensive, we once again sampled from larger outlets (>10,000 observations) while preserving all articles from few observations; we performed sampling on 62 of 193 outlets. We then twice queried gpt-4o-mini to assess stance toward China for a sample of China-related articles in our dataset; approximately 5% of rows received different labels from each GPT run. For disagreements, we performed a third GPT run to break the tie. After obtaining stance labels for all outlets, we calculate the stance measure by aggregating annually for each outlet the average percent of articles each year with positive stance toward China.

⁶<https://huggingface.co/IDEA-CCNL/Erlangshen-DeBERTa-v2-320M-Chinese>

⁷The WudaoCorpora is a large (5TB) Chinese-language pre-training corpus compiled by the Beijing Academy of Artificial Intelligence and used widely in Chinese-language artificial intelligence research (Xue et al. 2022).

A2.3. Local News Classification

To classify local news coverage, two trained independent coders assessed whether each article covers news relating to the publishing outlet’s residence country, where “local” means news pertaining to the organization’s country of residence rather than China. We were able to identify residence countries for 188 of 193 outlets; outlets with unknown residence countries were excluded from this analysis. After achieving human inter-coder agreement of 89.4%, we fine-tuned a DeBERTa model using the hand-labeled data to label remaining articles, achieving an F1 of 0.92 on our holdout set.

A2.4. Error Propagation

Our final labels contain three sources of error. Two are measurement errors: the first results from DeBERTa’s misclassification of news article relevance, while the second results from GPT’s misclassification of China-related articles’ stances toward China. The third is sampling error for the 62 outlets from which we analyzed only 10% of articles. We treat DeBERTa’s China-relevance classifications (the first error) as defining our analysis sample and do not propagate this source of uncertainty in our confidence intervals. In effect, our sample thus includes approximately 5.2% false positives (non-China articles misclassified as China-related) and excludes approximately 16.1% of true China articles (false negatives). We lack sentiment labels for articles DeBERTa classified as non-China-related, making it infeasible to fully propagate DeBERTa classification uncertainty without additional labeling that is prohibitively costly. We propagate the latter two sources of error using a non-parametric bootstrap procedure (1000 iterations) that accounts for GPT sentiment classification error by randomly flipping the label classifications at the same rate as the classifier error rate (FPR: 13.7%, FNR: 8.6%, estimated from five independent validation runs) and accounts for sampling error by re-sampling from outlets where only 10% of articles were analyzed. After each iteration we recalculate the outcome variables and construct confidence intervals from this bootstrapped range of values.

A2.5. Censored Topics

To identify topics censored in China, we adopted dictionaries of Chinese-language keywords and phrases. Each topic label was assigned based on the presence of one or more of the following terms in article titles or full text using regex.

Falun Gong: 法轮功, 李洪志, 法轮大法

Tiananmen Massacre: 天安门事件, 天安门屠杀, 天安门大屠杀, 六四事件, 六四屠杀, 六四

Dalai Lama 达赖喇嘛, 丹增嘉措

Xinjiang concentration camps: Articles that contain the term 新疆 (Xinjiang) together with at least one of the following: 集中营 (concentration camp), 再教育营 (re-education camp), 强制劳动 / 强迫劳动 (forced labor), 压迫 (oppression), 文化灭绝 (cultural extermination), 拘禁 / 拘押 (detention), 人权 (human rights), 宗教压制 (religious suppression), 种族清洗 (ethnic cleansing), 失踪 (disappearance).

A2.6. Event Detection Pipeline

We begin by gathering the top events each year from Wikipedia from 2000 to 2023, yielding 2,059 events.⁸ Next, we generate text embeddings for all 2,059 events of Wikipedia event descriptions using m3e-base, a bilingual embedding model built on Chinese RoBERTa and trained with contrastive learning that is compatible with the Sentence-Transformers framework⁹. Using the same approach, we then generate embeddings for all newspaper articles in our dataset ($n = 19,832,534$), using only the title and first 500 characters of each article.¹⁰ Next, we compare each article’s embedding to that of all events occurring within 1-day of the article’s publication date (e.g. November 1-3 for an article published on November 2), and subset to all events with a cosine similarity greater than 0.7).

Using human coders, we then manually evaluate the top 200 most-similar articles for a sample of 100 events, and manually determine whether each article is indeed reporting on the Wikipedia event. If it is, we mark “1”, if not, we mark “0.” We then use these labels to fine-tune a DeBERTa classifier to predict whether each story covers each Wikipedia event ($F1 = 0.97$).

The output of this pipeline is the set of articles, across all outlets (including Party media), covering each event in the Wikipedia list. Our outcome measure for global event coverage is the percent of each outlet’s monthly news articles dedicated to covering each global event, with bootstrapped confidence intervals that inherit DeBERTa classifier error (using the same procedure as described in Appendix A2.2).

⁸See for example: <https://en.wikipedia.org/wiki/2023>

⁹We use moka-ai/m3e-base.

¹⁰We impose this limit as our model has a context window of 512 tokens.

A2.7. Event Framing

To measure each outlet’s framing of global events, we embed the full text of each outlet’s articles covering each event in the Wikipedia dataset using m3e-base. For each event, we calculate an average coverage embedding across all outlets (aggregated at the outlet level first). We then compute the residual of each outlet’s coverage embedding against this event-average, capturing how each outlet’s framing deviates from the mean. After aggregating these residuals by event and then by outlet, we generate a single framing embedding for each outlet representing its average deviation in coverage from mean coverage across all events in the Wikipedia dataset.

To provide a benchmark, we generate a framing embedding using the same procedure from a commercially-oriented PRC state media outlet, *The Paper* (*Pengpai Xinwen*, 澎湃新闻). We then assess whether patronage outlets are systematically closer in embedding space to *The Paper*’s framing embedding than unaffiliated outlets are, conditional on covering the same event. We used a permutation test to calculate p-values: we randomly shuffled outlet group labels (patronage and unaffiliated), measured the group mean cosine distance to *The Paper*, and subtracted means to determine how much more similar patronage outlets are than unaffiliated outlets. We repeated this 1,000 times and used the distribution to construct p-values.

A3. Additional Analysis

A3.1. Repression and Exit: The Case of Hong Kong

Hong Kong media outlets offer suggestive evidence about the consequences of direct repression on media outlets rather than patronage. In 2020, the PRC government passed the Hong Kong national security law, establishing punishments for secession, subversion, terrorism and collusion with foreign organizations, the stipulations of which were soon applied to local media organizations. Hong Kong’s media freedom declined precipitously in the years after the law’s passage, dropping 80 places in Reporters Without Borders’ Freedom Index and prompting widespread self-censorship among the media ([Reporters Without Borders 2025](#)). Our data (see Figure A1) confirms that the crackdown coincided with an increase in positive reporting toward China in Hong Kong media. All Hong Kong media outlets increased the percentage of positive China stories in the wake of the crackdown. However, Figure A1 also depicts the birth of several new “dissident” media outlets based outside of Hong Kong that trace their origin to the Hong Kong crackdown, and are staffed by journalists who fled Hong Kong in the wake of the crackdown.¹¹ These outlets are characterized by a much lower fraction of positive stories about China, with no dissident outlet surpassing more than 25% positive China stories. While this evidence is only suggestive, it indicates that physical repression — coupled with territorial control — may succeed in changing news coverage, but may also lead to backlash in the form of new critical outlets.

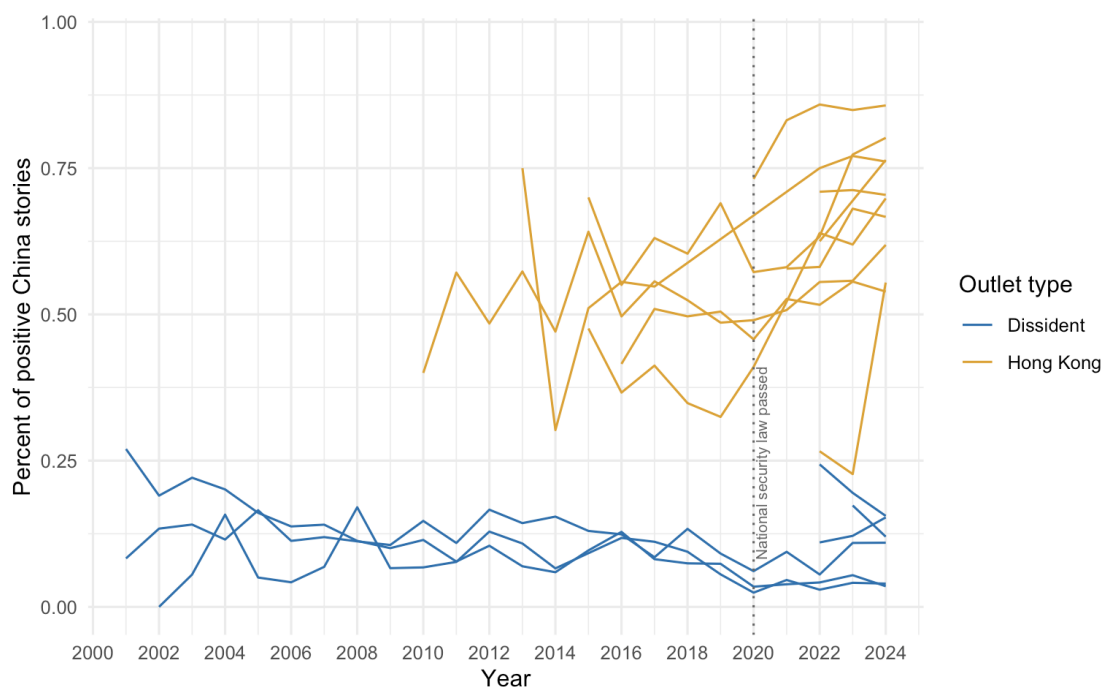


FIGURE A1. HONG KONG AND DISSIDENT MEDIA OUTLETS IN THE WAKE OF THE 2020 NATIONAL SECURITY LAW; OUTLETS WITH FEWER THAN 250 STORIES ANNUALLY OMITTED

¹¹We operationalize “dissident” media as outlets whose “About us” page or founding mission states an explicit opposition to the CCP.

A3.2. Local News Stance

Patronage outlets are not negative toward the host countries where they are reporting. Using the same DeBERTa pipeline described in Appendix A2.2 and A2.1, we also assess the percentage of outlets' article stances that are negative to the country in which the newspapers are located. (1 is negative; 0 is neutral/positive.) We find that there is no change to newspaper coverage after receiving PRC patronage. Comparing the two years pre and post affiliation (dropping the year of affiliation), a paired t-test reveals no significant difference in negativity post-affiliation.

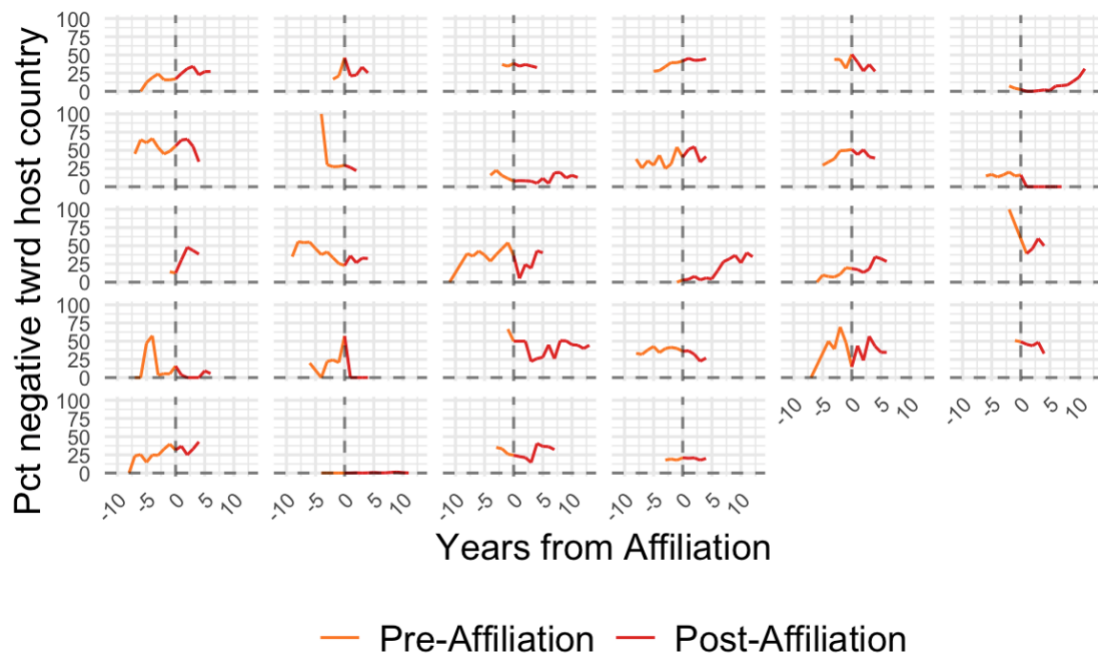


FIGURE A2. PERCENT OF NEGATIVE STORIES ABOUT A HOST COUNTRY, PRE/POST NEWS OUTLET AFFILIATION.

A3.3. Webpage linkages

To validate our measure of media patronage, we assess the frequency that diaspora news organization websites redirect through hyperlinks to websites on the Chinese mainland, following approaches in communication and social computing for assessing interconnectedness (Freelon 2014). We extracted all URLs from all news organization website homepages and, after cleaning them and expanding shortened URLs, identified their country domain by cross-referencing the ICANN database, the website’s domain suffix (e.g. “.cn”), and using a database of website registrations from DomainTools, to look up the website’s country of registration (see Appendix A1.3). After averaging the percentage of links to China in each outlet in each year, we then average across all outlets with homepages available in that year to assess the average percentage of news organization homepages links that direct visitors back to the Chinese mainland.

This analysis confirms our measure of PRC linkage. Figure A3 shows that over the past twenty-three years homepage linkages to China have nearly tripled, from 6.8% in 2000 to 22.3% in 2023, indicating a growing share of global Chinese-language news linking to content hosted in mainland China.

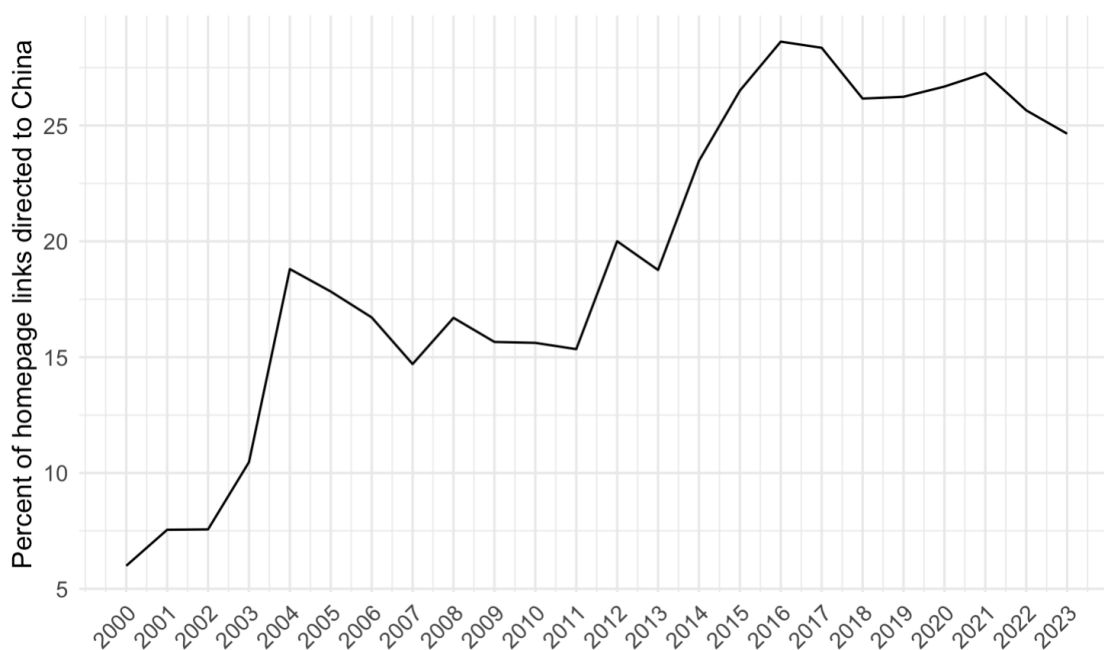


FIGURE A3. AVERAGE PERCENTAGE OF HOMEPAGES LINKING TO WEBSITES REGISTERED IN CHINA.