Resurgent Authoritarian Influence: Evidence from Machine-Generated, Cross-National Data*

Jeremy Springman[1,*] Andreas Beger[3] Erik Wibbels[1,2]

August 21, 2023

In recent years, scholars and policymakers have expressed growing concern about powerful non-democracies becoming more assertive in their foreign policy. According to popular accounts, this has included increased attempts by Russia and China to influence political outcomes in less powerful nations. To date, the absence of quantitative data has limited systematic investigation of this resurgence of authoritarian influence (RAI). This report has six objectives. First, we describe limitations of extant data on Russian and Chinese influence. Second, we introduce a unique, high-frequency dataset that relies on machine learning to track 22 different modes of influence by Russia and China in 56 low- and middle-income countries between 2012 and 2023. Third, we describe the construction of indices that summarize foreign influence in distinct RAI themes. Fourth, we describe trends in RAI since 2012, including the specific tools used to influence target countries. Fifth, we use our high-frequency data to forecast RAI activity; we show that in many cases, the forecasts are quite accurate. Sixth, we introduce a new set of online dashboards that analysts and researchers can use to visually explore RAI activities and forecasts in unprecedented detail.

[1] PDRI-DevLab, University of Pennsylvania
[2] Department of Political Science, University of Pennsylvania
[3] Basil Analytics

[*] Correspondence: Jeremy Springman <jspr@sas.upenn.edu>

*This report and the digital tools described herein were funded by the United States Agency for International Development (USAID) under the Analytic Task on Authoritarian Resurgence and Influence (ATARI) projects.

Introduction

The third wave of democratization was accompanied by a spectacular decline in the international influence of authoritarian governments. Culminating with the collapse of the Soviet Union, the ascendance of advanced democracies resulted in pressure on smaller and poorer countries to liberalize both their economies and their political systems. However, the last 15 years has seen the most powerful non-democracies become both more autocratic and more assertive in their foreign policies (Diamond 2020). According to popular narratives, this has included increased attempts by Russia and China to influence political outcomes in less powerful nations, ranging from the dissemination of propaganda targeting citizens' attitudes to technological support for governments seeking to suppress political competition. Over this period, concerns voiced by scholars and policymakers about a resurgence of authoritarian influence have driven high-level decision-making in foreign policy, including major investments by Western country governments in strengthening ties with strategically important countries.

Despite the importance of these issues for foreign policy, the absence of quantitative data on foreign authoritarian influence has limited systematic investigation. Assembling data on foreign influence activity is difficult due to the sporadic nature of common RAI tools and the absence of publicly available administrative data. Furthermore, while major events, such as military confrontations or major treaties, are covered widely in the international press, less dramatic but nonetheless important events are often not reported beyond the domestic papers in the countries targeted for influence (which we refer to as 'target countries').

This report has six objectives. First, we briefly describe existing research on Russian and Chinese influence and discuss the limitations of extant data. Second, we introduce the Resurgent Authoritarian Influence (RAI) dataset. This dataset tracks the incidence of reporting on events indicative of 22 different types of Russian and Chinese influence across international, regional, and domestic online news sources for 55 countries. To date, we have scraped more than 100 million articles published by 16 international, 12 regional, and 252 domestic online newspapers. We then fine-tuned an open source, transformer-based large language model to identify reporting on a wide range of foreign influence events and developed an extensive list of keywords to identify influence events involving Russian and Chinese agencies, companies, or individuals. Finally, we apply open-source geo-parsing tools to locate events and identify those happening within a target country. Importantly, these data are updated every 90 days, ensuring their utility for studying current events in additional to medium-term patterns and trends.

Third, we describe the construction of indices that summarize foreign influence according to six distinct RAI themes. We discuss the trade-offs involved in different approaches to summarizing our measures and the justifications for the specific method that we use. Fourth, we describe RAI activity since 2012, identifying the target countries where influence is most intense, the relative amount of influence activity between influencing countries (Russia and China) across the world, and the specific tools being used to influence target countries. Finally, we assess empirically whether popular narratives about a resurgence in foreign authoritarian influence are accurate. In doing so, we provide evidence that authoritarian influence has increased dramatically in recent years, providing one of the first tests of an empirical claim driving high-level decision-making in foreign policy and international advocacy. Fifth, assess our ability to forecast future levels of RAI activity using the RAI dataset. We show preliminary evidence that there are many countries where we can accurately forecast RAI activity, especially for Russian influence. Sixth, we introduce three publicly accessible data dashboards that analysts and researchers can use to visually explore RAI activity in unprecedented detail.

The RAI dataset presents a valuable and timely resource for both academic researchers and policymakers. For researchers, this dataset offers a unique opportunity to study the dynamics of global power politics, enabling a deeper understanding of how Russian and China employ strategies to shape the geopolitical landscape and how those strategies have changed over time. By uncovering patterns, trends, and correlations in their influence efforts, scholars will be able to make inferences about the underlying motives, mechanisms, and potential consequences of foreign authoritarian influence. Furthermore, insights gleaned from this dataset can aid policymakers in crafting diplomatic and strategic approaches that align with ever-changing geopolitical realities, enhancing the ability to anticipate or mitigate the influence exerted by these countries. As global dynamics continue to evolve, this dataset stands as a powerful tool, fostering both scholarly discourse and effective policy formulation in an increasingly complex international arena.

Russian and Chinese Foreign Influence

In foreign policy and international advocacy circles, there has been rampant speculation that resurgent authoritarian influence has been increasing and that this influence has been reshaping global power dynamics. Beyond the flurry of technical reports and opinion pieces, concerns about RAI are driving major policy decisions. For example, in 2019, the United States Agency for International Development launched the *Countering Malign Kremlin Influence Development Framework* and the $300 million *Countering Chinese Influence Fund* and listed both among it's key accomplishments for 2020.[1] However, quantitative data on RAI are rare.

To date, the most comprehensive data on RAI comes from AidData. AidData provides detailed, project-level data on China's foreign aid and state financing to all low- and middle-income countries from 2000-2017 and more recently, annual event data on 10 types of diplomatic events in 25 in East Asian and Pacific countries (2000-2016) and 13 South and Central Asian countries (2000-2018). Both datasets received major updates in 2021 (Custer et al. 2021).

AidData's data on foreign aid and state financing comes from an exhaustive human review of official government sources from China and aid/finance receiving countries. For each project identified during this review of official sources, coders seek to collect information on 70 unique characteristics, including project start and end dates, fees, terms, and implementing organizations. When possible, this includes the specific dates on which projects begin and end. This process is complemented by a systematic search of media content sourced from Factiva, a media monitoring company, to collect information on additional project details. Public diplomacy event data is compiled using a similar methodology. These data are available at the country-year level, meaning that AidData provides the number of times each event occurred in each country in each year. These data have been a boon to researchers and policymakers alike.

Yet because AidData's is the only systematic data available, the majority of research on foreign authoritarian influence has focused on the impact of Chinese development assistance and foreign direct investment. This nascent literature has often pointed to China's use of aid and investment as a means of obtaining political support from, and increasing the political stability of, strategically important countries (Dreher et al. 2019; Ma and Teng 2018). Accordingly, China has pursued these

[1] 'USAID Key Accomplishments,' USAID

objectives by strengthening incumbents with fungible economic and political resources that can be used to spur development, reward supporters, or suppress oppositions (Dreher et al. 2019).

Recent work suggests that these forms of Chinese influence have been quite successful in their objectives, finding that Chinese aid and investment reduce conflict in recipient countries (Gehring, Kaplan, and Wong 2019; Strange et al. 2017), increase economic growth (Dreher et al. 2021; Knutsen and Kotsadam 2020), increase government repression (Gehring, Kaplan, and Wong 2019; Kishi and Raleigh 2017), decrease membership in trade unions (Isaksson and Kotsadam 2018), and undermine citizen support for democracy (Gehring, Kaplan, and Wong 2019).

While the existing data on Chinese influence has yielded valuable research and policy insights, there are several important blindspots. First, these data focus exclusively on Chinese influence. While China is undoubtedly the greater global power, Russian influence has often proven more aggressive and destabilizing in some countries, especially in recent years. Second, these data are updated only sporadically, limiting their utility for understanding current events, emerging crises, or changes in the tactics of influencing countries. Third, these data focus on a narrow range of influence tools that represent a small set of the tactics that Russia and China deploy to exert influence abroad.

Indeed, foreign authoritarian influence encompasses a broader set of policy tools that are likely deployed for different objectives at different times and in different places. Indeed, this influence includes a variety of military/security, cultural/social, and foreign policy tools designed to provide support for, or apply pressure to, governments across the world. For example, Russia and China build the security capacity of their partners through the export of military hardware,[2] collecting intelligence on the political opposition,[3] engaging in cyber attacks against independent news sources,[4] and transferring surveillance technology.[5] Alternatively, Russia and China also deploy similar tactics to pressure target governments, including everything from surveillance and cyber-attacks[6] to the movement of troops and military assets.[7]

Similarly, Russia and China promote their interests abroad using diplomatic maneuvers that boost their political allies, including formal statements and official visits to express support,[8] trade agreements and coordinated sanctions,[9] and diplomatic mediation and recognition.[10] They also frequently deploy social and political tools to strengthen political allies and boost perceptions of authoritarian governance through, for example, establishing cultural centers,[11] media campaigns,[12] and interventions in domestic politics and policies.[13] These same tools are also used to weaken

[2] Andy Wrong. 'Understanding Russia's Foreign Policy Through International Arms Sales.,' Wavell Room, April 30, 2020.

[3] 'Uganda and Zambia rejects Huawei spying allegations.,' BBC, August 16, 2019.

[4] 'Serbia's Independent N1 Portal Buffeted by Cyber-Attacks.,' BalkanInsight, January 31, 2020.

[5] Stephen Kafeero. 'Uganda is using Huawei's facial recognition tech to crack down on dissent after anti-government protests.' Quartz Africa, November 27, 2020.

[6] Steve Holland, Doina Chiacu. 'U.S. and allies accuse China of global hacking spree.'Reuters, July 20, 2021.

[7] Rajesh Roy. 'China, India Move Tens of Thousands of Troops to the Border in Largest Buildup.,' The Wall Street Journal, July 2, 2021.

[8] Maja Zuvela, Aleksandar Vasovic. 'Beset by protests, Serb leader stages lavish reception for Putin.,' Reuters, January 17, 2019.

[9] Yuras Karmanau. 'Belarus, Russia announce retaliatory sanctions against EU.,' AP, October 2, 2020.

[10] Helena Legarda. 'China as a conflict mediator.,' MERICS, August 22, 2018.

[11] 'Cultural Cooperation.,' Embassy of the Russian Federation in the Kingdom of Cambodia.

[12] Lorne Cook. 'EU report takes aim at Russia over vaccine fake news.,' AP, April 28, 2021.

[13] Krishna Pokharel, Eric Bellman. 'Nepal Elections Could Set Back China's Goals.,' The Wall Street Journal, December 23, 2020.

perceived adversaries or undermine public support for resistance in target countries.[14]

Relatedly, attempts to exert influence can also have unintended consequences. For example, Chinese influence has been linked to increased corruption (Isaksson and Kotsadam 2018) and the spread of Chinese organized crime.[15] Consistent with those findings, perceptions of excessive Russian and Chinese influence have caused anti-incumbent political mobilization in many countries.[16]

As these examples illustrate, Russia and China use an impressive array of tools to exert influence abroad. For many of these tools, their use is deliberately covert, rarely disclosed in official government records, and only exposed through investigative reporting. Furthermore, RAI tools are often deployed sporadically in response to short-term contingencies such as domestic political crises in strategically important countries, rather than on a predictable, ongoing basis.[17] These characteristics have made tracking RAI extremely challenging and limited the ability of researchers to identify broader trends in, or investigate the impact of, these tactics.

The Resurgent Authoritarian Influence Dataset

To address the limitations of existing data on foreign authoritarian influence, we introduce the Resurgent Authoritarian Influence (RAI) dataset. RAI tracks the incidence of reporting on 22 events indicative of Russian and Chinese influence. To develop the list of RAI events, we reviewed existing research and consulted with partners in civil society and the United States Agency for International Development (USAID) to identify the events and behaviors seen as most indicative of attempts by powerful governments to exert influence over foreign countries.[18] We then group these 22 events into six themes, which we use to create index variables to summarize levels of activity across substantive domains. Table 1 provides a complete list of the 22 event types and how they are grouped into the six themes. In the following section, we discuss the construction of these index measures in greater detail.

To track reporting on these events, we scrape news stories published by 16 international,[19] 12 regional, and 252 domestic online newspapers. We begin with the most popular international sources, which tend to have extensive international coverage, deep historical archives, and reliable publishing habits, thereby eliminating the need for human oversight of scraping. We complement these major international sources with regional sources, which are important for countries with particularly shallow or repressive domestic media markets. While international and regional sources publish a large volume of news and are relatively easy to scrape, their coverage of events in most countries focuses on only the most significant developments. For this reason, the vast majority of coverage for the countries in our data comes from domestic sources.

For domestic sources, we aim to maintain a minimum of 3 sources per country; we are actively scraping from an average of 4.6 sources per country. To select domestic sources, we identify the top digital news sources by web traffic by consulting lists of newspapers maintained by university library

[14] 'Moldova Bars Russian Delegation for Attempted Election Interference.,' Balkan Insight, April 18, 2023.

[15] 'Triad Societies and Chinese Organised Crime in South Africa.,' U.S. Department of Justice, September 2003.

[16] 'Georgia protests: Thousands storm parliament over Russian MP's speech.,' BBC, June 21, 2019.; Abu-Bakarr Jalloh, Fang Wan. 'Resistance growing to Chinese presence in Zambia.,' DW, April 9, 2019.

[17] Siegle, Joseph. 'Russia in Africa: Undermining Democracy through Elite Capture.,' Democracy in Africa, September 23, 2021.

[18] Please see the project website for a list of definitions for each event.

[19] We include two Chinese (scmp.com and xinhuanet.com) and one Russian (themoscowtimes.com) sources aimed at international audiences.

Table 1: RAI events theme

Theme	Definition	Events
Soft Power	Attempts to change attitudes or beliefs of publics or influence policy through the mobilization of citizens.	Diaspora Activation Media Campaign Cultural Activity
Hard Power	Attempts to strengthen or weaken the military capacity of or military ties with incumbent regimes.	Security Transfer Security Exercise Security Engagement Security Presence
Economic Power	Attempts to strengthen or weaken the economic capacity of or economic ties with incumbent regimes.	Aid Operation Investment Action Trade Action Trade Agreement
Diplomacy	Attempts to strengthen or weaken the diplomatic standing of or ties with incumbent regimes.	Diplomatic Engagement Diplomatic Relations Diplomatic Action Diplomatic Statement Diplomatic Visit
Domestic Interference	Attempts to directly influence the policies or capacity of incumbent regimes through non-military actions.	Intelligence Operation Policy Intervention Cyber Operation Tech Transfer
Backlash	Unintended outcomes of attempted influence.	Corruption Organized Crime

guides and publicly available media reports. We also include sources recommended by our partners working in international NGOs, USAID country offices, and local civil society organizations. We then check each source to ensure that it primarily publishes original content, is machine scrapable, and has a sufficiently large archive to justify scraping (with frequent publications over at least several years). Finally, we consult our partners and conduct desk research to identify the independence of each source.

For each source, we then develop and deploy a custom scraper to accommodate the website architecture and a custom parser to extract the publication date, title, and story text from each article. Depending on website architecture, we obtain news articles by scraping sitemaps, newspaper archives, or by simulating infinite clicking/scrolling using Selenium. In order to avoid storing the same article multiple times, we de-duplicate based on URL similarity and title similarity for articles published on the same day. To date, we have scraped and stored more than 100 million news articles covering 2012 to 2023. 81% of these articles come from domestic sources, 4% from regional sources, and 15% from international sources.

After articles are scraped, we identify the language and use neural machine translations (NMT) through Hugging Face or OpenNMT to translate more than 30 languages into English. We test the efficacy of all translation models by extracting sample text from articles published in each language and running the text through all available translation models on the Hugging Face open database. We then assess whether the translations are sufficiently comprehensible to identify the main event being reported on. If they are not, we compare the performances with those of our other APIs and choose the one that yields the optimal sentence-to-sentence translations with sufficient human readability. We have not found translation to be an obstacle to successful event classification in any language to date.

Once they are translated, we classify each article according to the main event being reported on in the text. To do so, we fine-tuned an open source, transformer-based large language model to identify reporting on 22 foreign influence events. Specifically, we fine-tuned a RoBERTa model for this task using a double human-coded dataset of 3,400 news articles (including a large sample of articles translated from various languages into English). Overall out-of-sample accuracy is above 0.8 on human-coded event data, with most misses coming from the presence of multiple events in a single entry or from partially overlapping event categories. To identify foreign influence events involving Russian or Chinese influence, we developed an extensive library of Russian and Chinese agencies, companies, or individuals. Articles that contain at least one of these keywords are counted as capturing Russian or Chinese influence, respectively.

To locate events and identify those happening within a target country, we implement Named Entity Recognition (NER) for locations using the CLIFF-CLAVIN geoparsing API service with the GeoNames ontological gazetteer. For international and regional sources, articles must mention a location within a target country to be classified as an event for that country. For domestic sources, articles must either mention a location within the target country or mention no locations in order to be classified as an event for that country.

Finally, we aggregate these data to the country-month level, normalizing the count of articles reporting on each event by the total number of articles published in that country-month.[20] The final RAI event measures correspond to the monthly share of all news articles reporting on a country that are reporting on each RAI event type. This ratio tells us how frequently each RAI event type is reported-on relative to the total volume of news in a given month. While this method does not directly allow us to code *individual* RAI events, it does provide information on the *importance* of each type of RAI activity in a given month. To construct summary indices, we sum the individual events according to their theme (referenced in Table 1). We also combine all RAI event measures into indices capturing the total amount of RAI activity from Russia, China, and both influencing countries combined.

Creating RAI Index Measures

In this section, we explain our approach to summarizing the 22 RAI event types into six summary indices. The goal of these indices is to allow RAI activities across a broad range of events to be summarized in the simple way. We began by comparing inductive and deductive approaches to constructing index measures for the RAI data. An inductive approach uses unsupervised methods to detect patterns in the data and constructs an index around these empirically identified patterns. A deductive approach uses the substantive understanding of researchers as the basis for summarizing data around a set of theoretical identified themes. Ultimately, our analysis suggested that the unsupervised methods that characterize an inductive approach detected relatively weak patterns in the data, suggesting a deductive approach is more appropriate for summarizing RAI activities.

[20]For countries where non-independent sources consistently produce a greater volume of news than independent sources, we impose equal weighting on independent and non-independent sources by multiplying the count of articles for each RAI event by a weight calculated by subtracting the number of articles from non-independent sources from the result of dividing the number of non-independent articles by 0.5 and dividing this value by the number of articles from independent sources for each month.

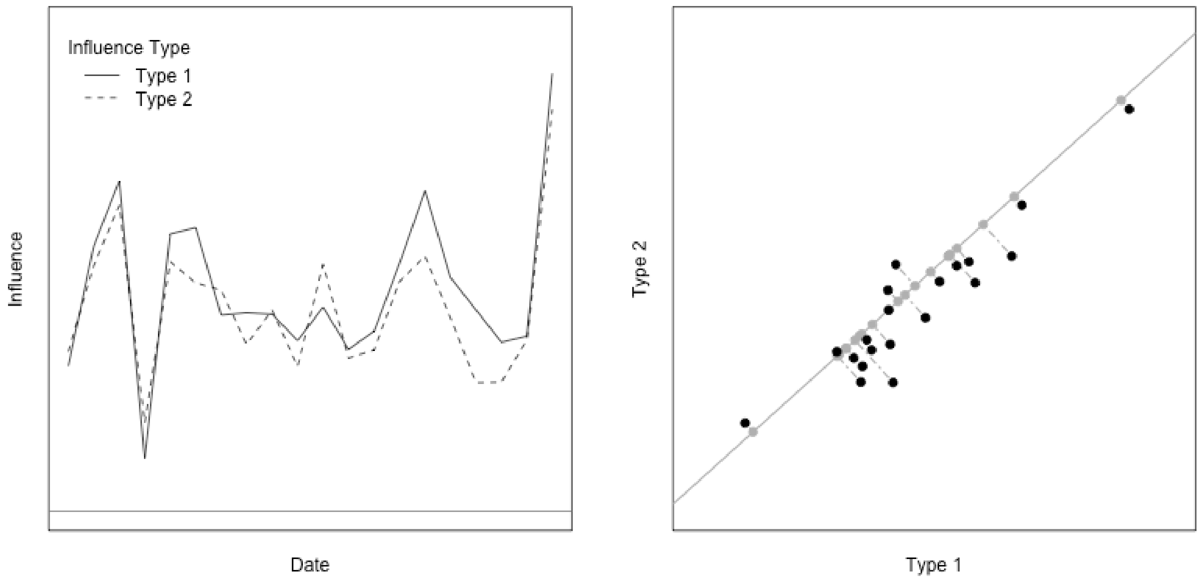

Figure 1: Strongly Correlated Influence Measures

Broadly speaking, inductive approaches seek to find a way to summarize variation across a number of variables using a reduced number of latent variables that capture a large proportion of the variation in the original variables. Figure 1 illustrates how this approach works using a contrived example of two simulated measures of influence (Type 1 and Type 2). Changes in the value of these simulated measures of influence over time are plotted in the subfigure on the left. These measures are highly correlated (0.9); the values tend to increase and decrease at the same time. In the subfigure on the right, we plot these two measures against each other. Each black point is the value of Type 1 and Type 2 on a given date. Because the two types are so highly correlated, knowing the value of one on a given data tells you a lot about the value of the other on that date. On dates when Type 1 has a low value, Type 2 is also likely to have a low value. An inductive approach is well suited to summarize Types 1 and Type 2 using the gray line in the figure that describes the relationship between them. Projecting the values of Type 1 and Type 2 onto the gray line gives us a new set of values marked by the gray points. These gray points have reduced a two-dimensional data (Type 1 and Type 2) to a one-dimensional space. This approach loses some information (the distance between the black points and the gray points, captured by the dashed lines), we are able to capture most of the information about the value of the two-dimensional data using only one dimension.

However, inductive approaches do not work in cases where the measures being summarized is not highly correlated with each other. Figure 2 show Type 1 when compared with a new simulated measures of influence: Type 3. These measures are weakly correlated with each other (0.15). Technically, we can still project the two-dimensional data down to a single dimension, but because the value of Type 1 does not tell us much about the value of Type 3, we lose a great deal of information (note the greater length of the dashed lines connecting grey dots to black dots). In other words, the projection does a poor job of summarizing the original data.

Turning to our analysis of the RAI data, we find that our 22 measures of RAI are not highly correlated with each other. Figure 3 shows the pairwise correlations of all 22 RAI variables against

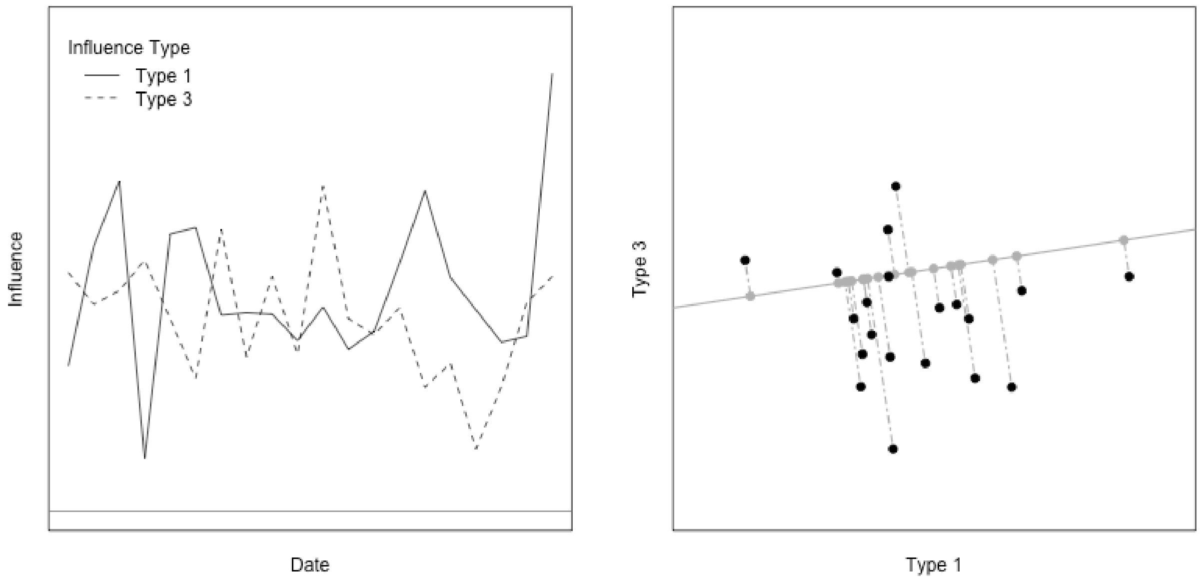

Figure 2: Weakly Correlated Influence Measures

each other. The cells on the diagonal capture each measure's correlation with itself, which are all necessarily equal to 1 (dark blue). Looking at the non-diagonal cells, we see that very few of these measures are highly correlated with another measure. Specifically, there are very few cells indicating pairwise correlations above 0.5, and only one pair of cells indicating a pairwise correlation above 0.75. In fact, the average pairwise correlation across all RAI variables is only 0.33, suggesting that inductive measures are not appropriate for summarizing our RAI measures.

Looking separately at Russia and China, we see a similar picture. Figure 4 shows the pairwise correlations of all 22 RAI variables calculated separately for Russian and Chinese influence measures. Chinese influence measures are remarkably uncorrelated, with an average pairwise correlation of only 0.2. Although Russian influence measures are much more correlated with one another, the average correlation is still only 0.38.

So far, we have only considered correlations at the global level (pooling data across all 56 countries). However, it may be the case that RAI measures have much higher correlations when we compute these correlations separately for each country. Specifically, it could be the case that the specific measures which are highly correlated with each other varies across countries, leading to low correlations when calculated on the pooled sample but higher correlations when calculated separately for each country. This is particularly likely for Russian influence, where we see some evidence for higher correlations relative to Chinese influence. To investigate this, calculated the pairwise comparisons separately for each of our 56 countries. Looking at the average pairwise correlation when each correlation is calculated separately for each country, we see that these correlations decrease rather than increase, registering at 0.18 for the combined sample, 0.16 for Russia, and 0.1 for China.

Importantly, even if we did see higher correlations when calculated separately for each country, an inductive index calculated separately for each country would pose severe challenges to interpretability. Inductive approaches identify clusters of highly correlated measures and construct a single summary index. When calculated separately for each country, these methods would identify

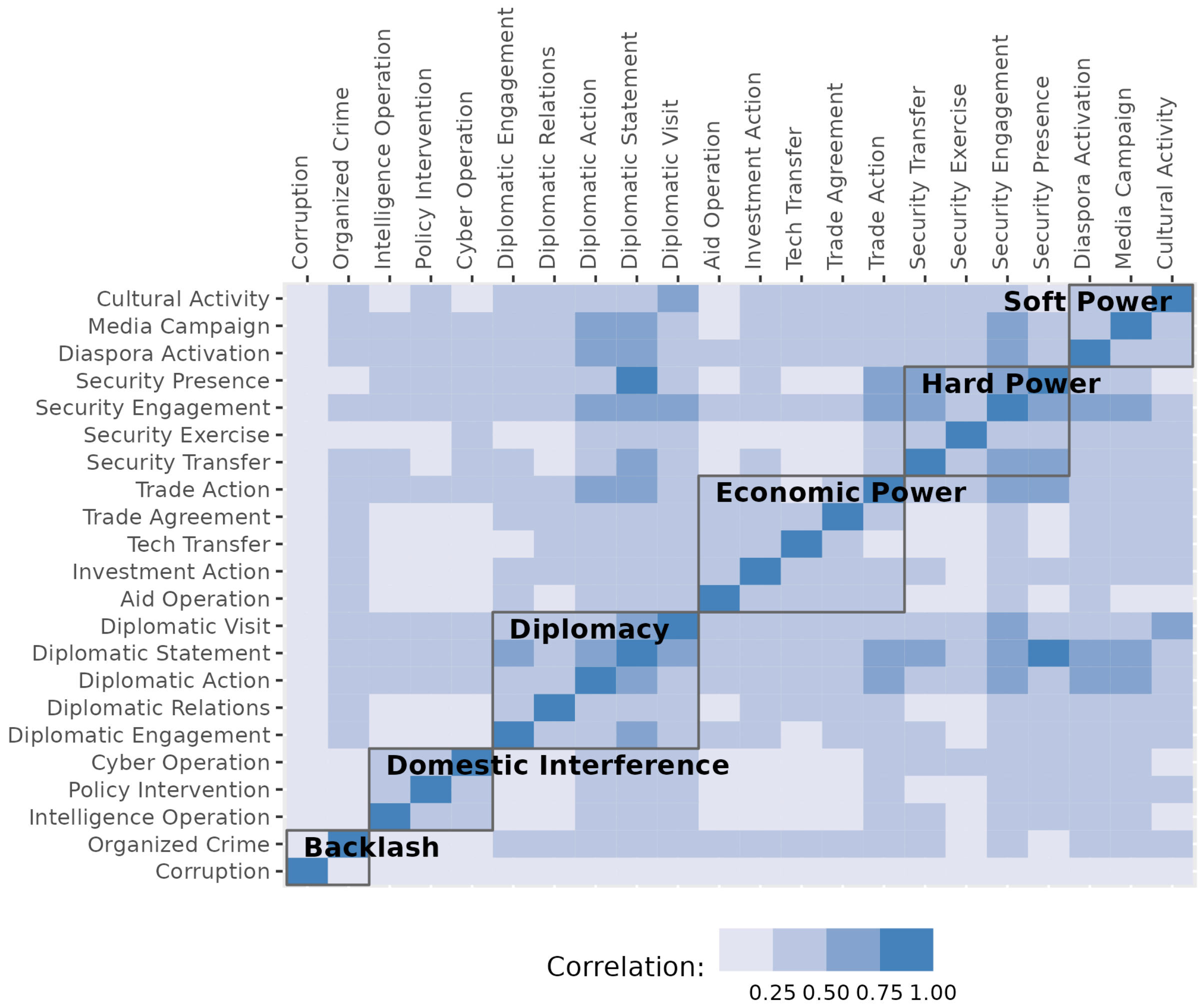

Figure 3: Heatmap of pairwise correlations across 22 RAI variables (combining influencers). Boxes capture our six index themes.

different clusters of highly correlated measures from which to build indices. The result would be country-specific index variables that would contain different underlying influence measures, preventing any comparison across countries.

To illustrate this, Figure 5 shows how to popular inductive approaches, inverse covariance weighting (ICW) and principal component analysis (PCA), would treat the Security Engagement measure in different countries and for different influencers. Each grey point is the weight the indicator would receive when constructing either a single, overarching ICW, or the weight for the first principal component in a PCA. In other words, each grey point captures how these methods would represent the importance of Security Engagement in the month-to-month values of a single index measuring overall levels of influence activity across countries and influencers. The red line highlights the weights for Algeria, specifically. These weights differ wildly across countries and influencers. We can see that for each country, changes in the Security Engagement measure would have completely different implications for changes in the value of the index. Even within a single country, changes in

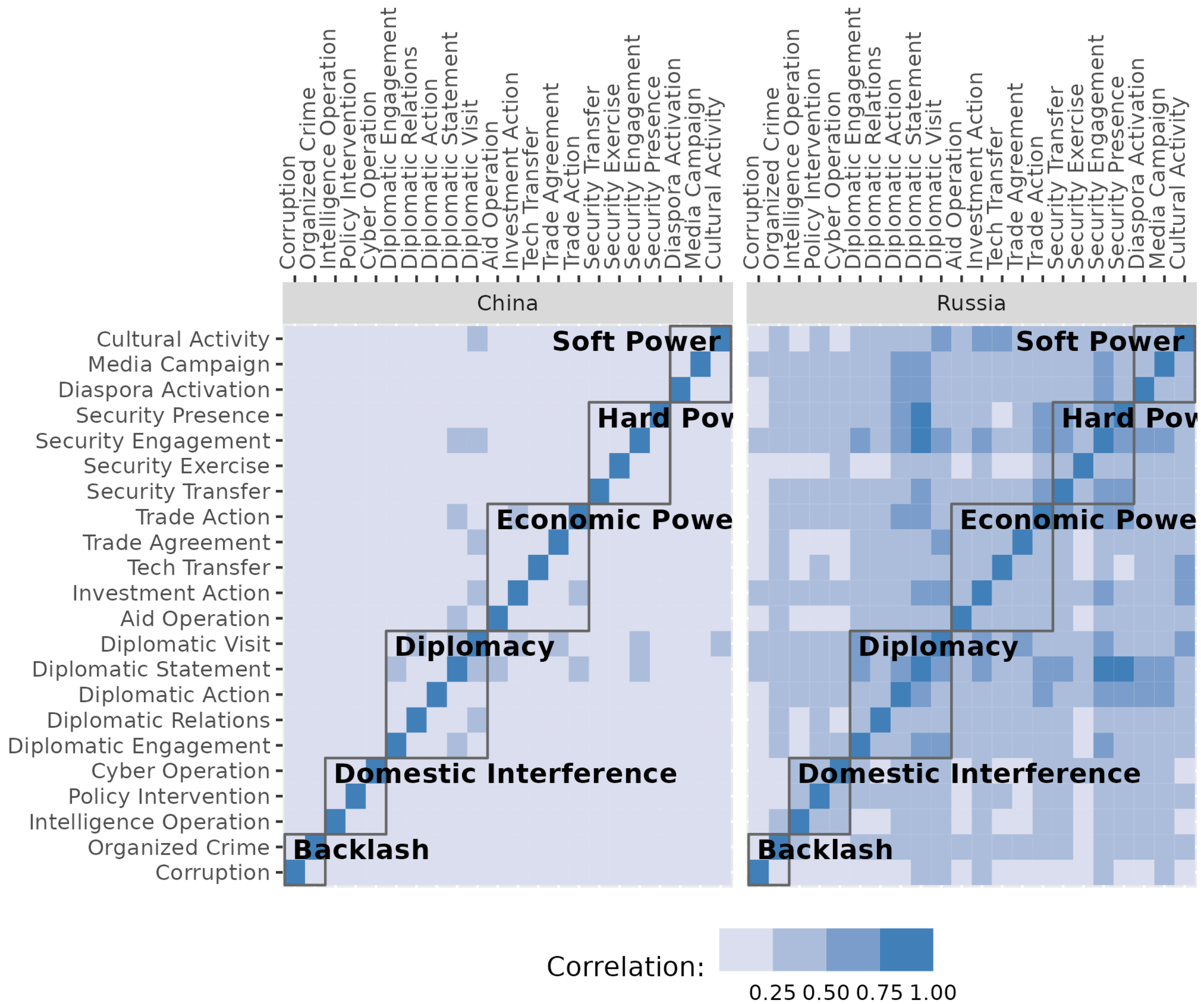

Figure 4: Heatmap of pairwise correlations across 22 RAI variables (disaggregating influencers). Boxes capture our six index themes.

Security Engagement would have very different implications for an aggregegate measure of Chinese influence than it would for an aggregate measure of Russian influence.

Because pairwise correlations between influence measures in both the pooled data and the country-level data are weak, we adopt a deductive approach to summarizing RAI activities. To do so, we construct a total of 21 index variables across two levels of aggregation. First, we construct one index variable measuring the total amount of influence activity in each country in each month. This measure of overall activity combines all of our 22 influence measures into a single, additive score. Second, we construct six measures that group our 22 influence measures into influence themes (see Table 1). These themes combine similar types of influence based on the substantive knowledge of the research team. Third, we calculate these seven measures separately for Russia and China, as well as for their combination.

By adopting these well-defined categories, we offer a framework to analyze the diverse manifestations of influence utilized by Russia and China to impact incumbent governments during elections.

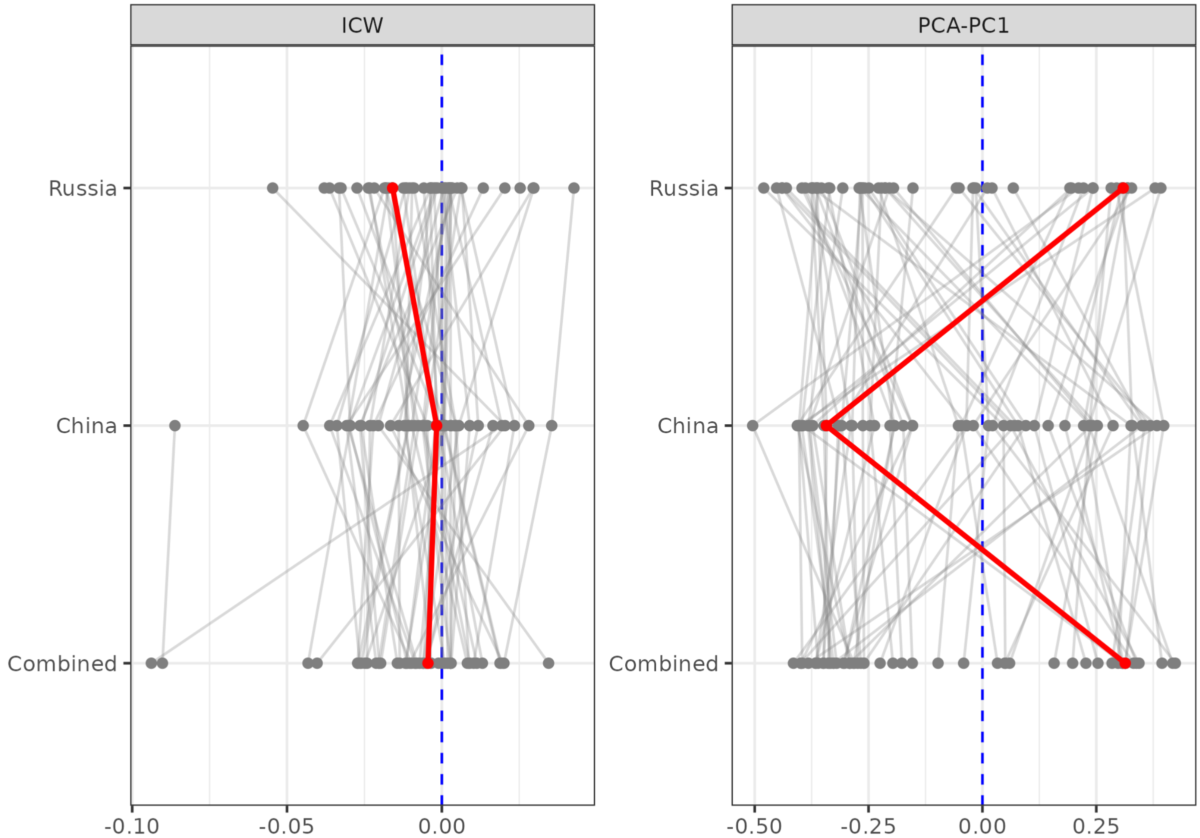

Figure 5: Weights for Security Engagement from ICW and PCA across countries and by influencer. Red captures weights for Algeria.

As discussed in the second section, we aggregate the individual influence measures into an index by summing them and dividing by the total number of all articles per country-month and multiply this value by 10,000. We normalize our measures by the total number of articles in order to account for monthly fluctuations in the volume of news published in each country. We multiply the final number fo 10,000 in order to increase the readability of our figures and summary statistics.

Describing Authoritarian Influence Using RAI Data

In this section, we describe monthly RAI activity since 2012, identifying the target countries where influence is most intense, the relative amount of influence activity between influencing countries (Russia and China) across target countries, and the specific tools they use to influence target countries. We report RAI activity as the number of articles reporting on RAI events per 10,000 articles published on a country. To facilitate analysis of how RAI activity is changing over time, we break our data into three historical periods: 2012-2016, 2017-2021, and 2022-2023. This provides two five-year windows at the beginning of our series followed by a smaller snapshot of the very recent past. 2012-2016 includes Russia's annexation of Crimea; 2016-2021 includes the onset of the COVID-19 pandemic, which originated in China prompted Russia and China's 'vaccine diplomacy'; and 2022-2023 includes Russia's invasion of Ukraine.

Evolving Spheres of Influence

We begin by looking at the share of all RAI articles that are reporting specifically on influence events involving Russia. It is important to note that this approach counts all influence events equally; for example, a report on a diplomatic statement by Russia would receive the same amount of weight as a report on a major transfer of technology by China. However, we expect that events will receive coverage in accordance with their significance, facilitating a rough measure of the share of influence coming from each influencing country.

Figure 6 plots Russia's share of the RAI influence for each country for each of our three historical periods. We see evidence for two important descriptive facts. First, Russian influence events are more concentrated in a sphere of influence based primarily on geographic proximity. China's sphere of influence is less geographically proscribed. While Russian influence dominates Eastern and Central Europe, Chinese influence dominates its nearest neighbors in Southeast Asia but also most of Sub-Saharan Africa. Latin America and North Africa are more evenly split. Second, we see evidence that these spheres of influence are shifting over time. Comparing the first period (2012-2016) to the intermediate period (2017-2021), we see stability, with Chinese influence growing in several geographic pockets, including much of East and Southern Africa, Southeast Asia, and Latin America. However, in recent years, Russia has expanded its own sphere of influence into the Sahel and North Africa, while also challenging China's dominance in Sub-Saharan Africa, and to a lesser extent, Southeast Asia. Interestingly, Russia's expansion does not appear to have spread in Latin America.

These descriptive results correspond with popular narratives that focused on growing Chinese influence in Africa and Latin America during the 2017-2021 period, as well as more recent accounts focused on a surging Russian influence campaign in the wake of the Ukraine invasion (McGlynn 2023). Importantly, our event classification models are trained to detect distinct events, rather than broad discussions of Russian or Chinese influence, minimizing concerns that the salience of these issues in international discourse is driving increased coverage. Furthermore, the vast majority of articles in our database come from domestic rather than international news sources, which we expect to be less susceptible to the changing whims of international discourse.

Stable Geographic Hotspots

We look next at the salience of RAI events across countries. In Figure 7, we combine Russian and Chinese influence events into a single measure tracking the total level of influence across countries. Each facet shows the relative prominence of reporting on RAI activities compared to all other reporting for every country in our data across the three historical periods. We see that the locus of RAI activity has remained fixed in Russia's immediate sphere of influence in East and Central Europe, with the specific country experiencing the highest level of RAI events shifting from Uzbekistan, to Azerbaijan, to Ukraine across the three periods. Interestingly, this suggests that although Chinese influence has been more widespread, especially in the first two periods, Russian influence is more intensely focused on specific nearby countries. Note that the scale (the count of RAI articles per 10,000 published articles) is set independently for each of the three periods, with the maximum value changing dramatically in the third period. Unsurprisingly, a large share of this increase appears to be coverage of events in Ukraine.

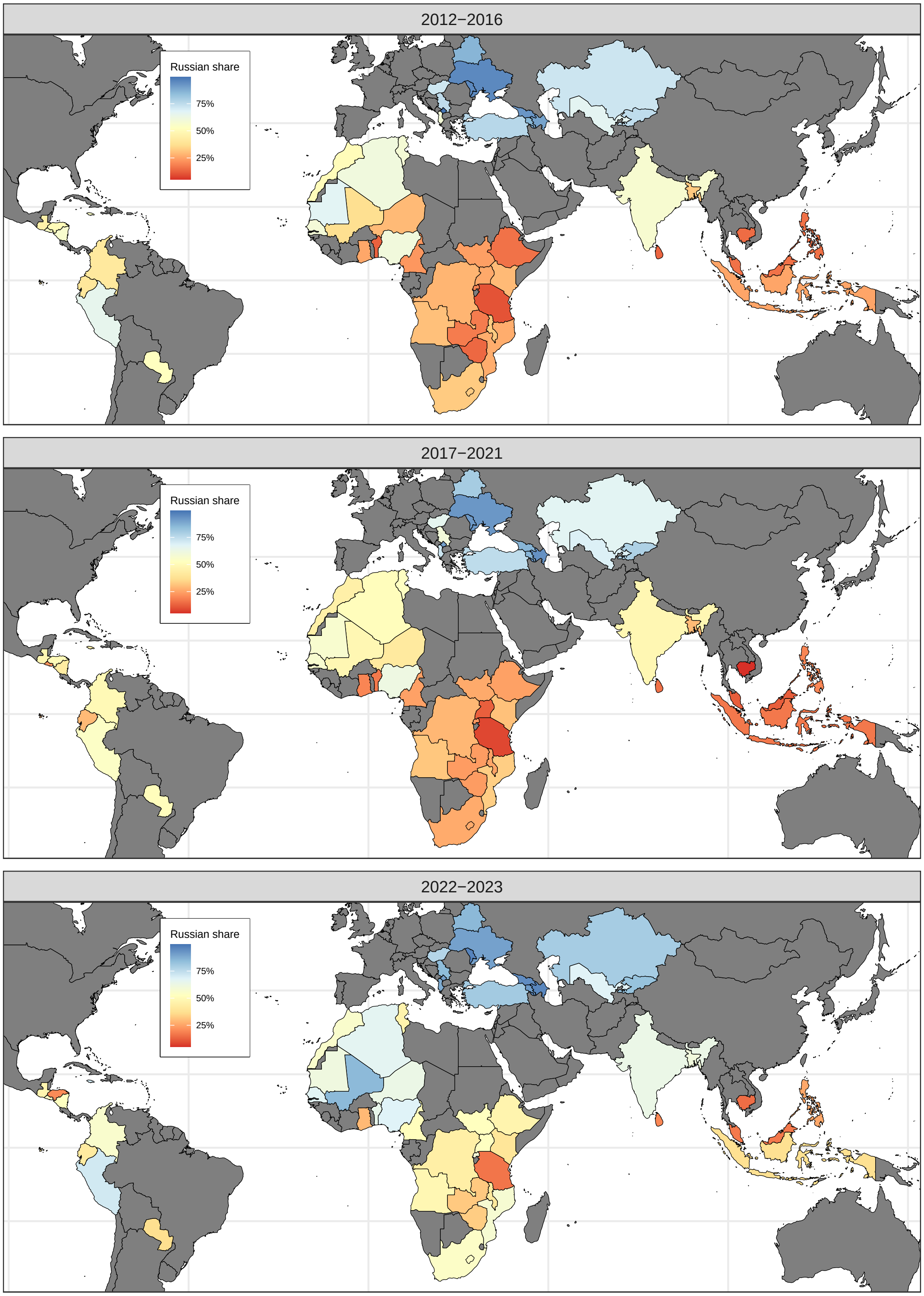

Figure 6: Share of RAI activity between influencing countries, measured as the normalized share of all articles reporting on RAI activities that focus on Russian influence events.

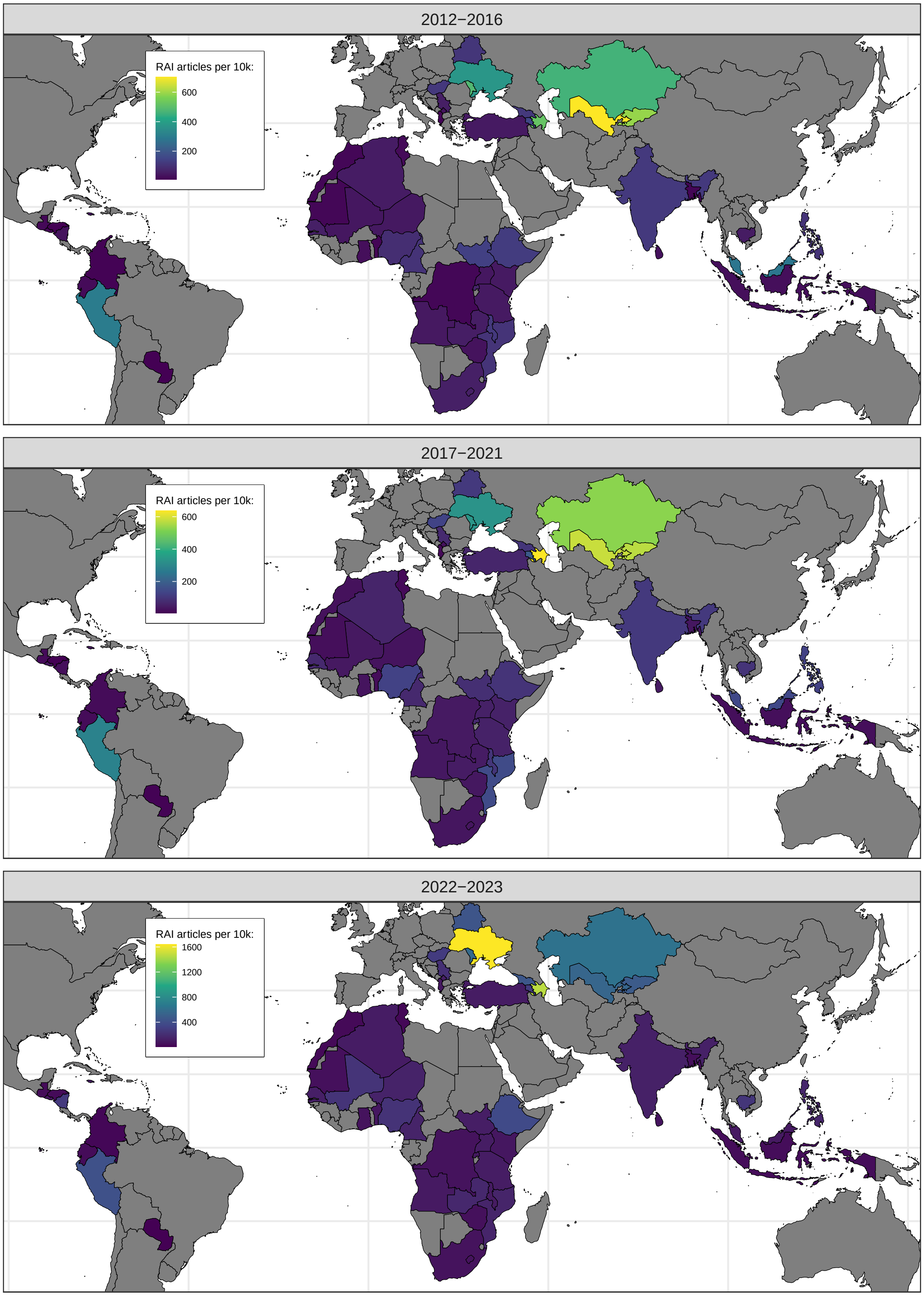

Figure 7: Total RAI activity across countries, measured as the number of articles reporting on RAI events per 10k articles published.

Shifting RAI Tools

Finally, we use the six themes described in Table 1 to identify the specific tools that Russia and China deploy. Again, it is important to note that this approach counts reporting on all influence events equally, but we expect that events will receive coverage roughly in accord with their significance, with more significant events being covered by multiple articles published in multiple sources over multiple days or weeks. Figure 8 shows the distribution of reporting on our RAI themes across countries and over the three historical periods. The relative prevalence of themes appears mostly stable over the first two historical periods. Economic Power is the most prevalent theme in the majority of countries in each region and both spheres of influence. However, we see distinct evidence for a growth in the use of diplomacy after 2022. This growth is especially apparent in the Russian sphere of influence and in places where Russia's share of influence saw the largest increases. This provides further evidence for an influence campaign by Russia beginning at roughly the same time as the invasion of Ukraine. Given the focus on diplomacy, this campaign appears aimed at bolstering international support from non-Western governments.

Looking at Figure 9, we see additional evidence for a Russian influence campaign beginning after the country's invasion of Ukraine. Figure 9 shows how the share of all RAI reporting dedicated to events under each theme changes over time. We see that Chinese influence has been dedicated to the use of Economic Power over the full time-series. The prevalence of each of the themes appears relatively stable over time, with a brief surge in Diplomacy and Hard power from mid-2022 which subsided in the first few months of 2023. For Russia, we see a much smaller share of influence events classified as Economic Power, with the share shrinking since the invasion of Ukraine. At the same time, Diplomacy appears to constitute a larger and growing share of Russia's RAI activity, especially since 2022. Across our three historic periods, we see Chinese Diplomacy increase from 20% in 2012-2016 to 21% in 2017-2021 to 26% in 2022-2023, while Russian Diplomacy increase from 25% to 29% to 34% over these three periods. This finding suggests that while a great deal of attention has gone to covert Russian influence operations, such as misinformation, traditional diplomacy is perceived as the most significant form of influence in the day-to-day politics of most countries (Eady et al. 2023; Karlsen 2019). This evidence that increased reporting on Russian influence after the invasion of Ukraine was driven by a focus on bilateral diplomacy should further diminish concerns that the salience of these issues in international discourse is driving increased coverage.

In summary, we see three main descriptive findings from the data. First, Russian influence events have been more geographically concentrated in their geographic neighborhood, while China's have been less geographically proscribed. However, Russia's share of RAI influence has expanded geographically and significantly eroded Chinese dominance in recent years. Second, the highest levels of RAI activity occur in countries within Russia's immediate sphere of influence. Third, there has been a growing emphasis on bilateral Diplomacy over other modes of influence; this appears to be especially true for Russia.

Is Authoritarian Influence Resurgent?

In this final section, we turn to the question of whether RAI activity has increased in recent years. We see some evidence for a large increase since 2022 in Figure 7, but it is difficult to tell whether this increase is driven by Russian influence, Chinese influence, or both, and whether it is concentrated in Ukraine or extents to other countries or regions. Turning to Figure 10, we see evidence for increased

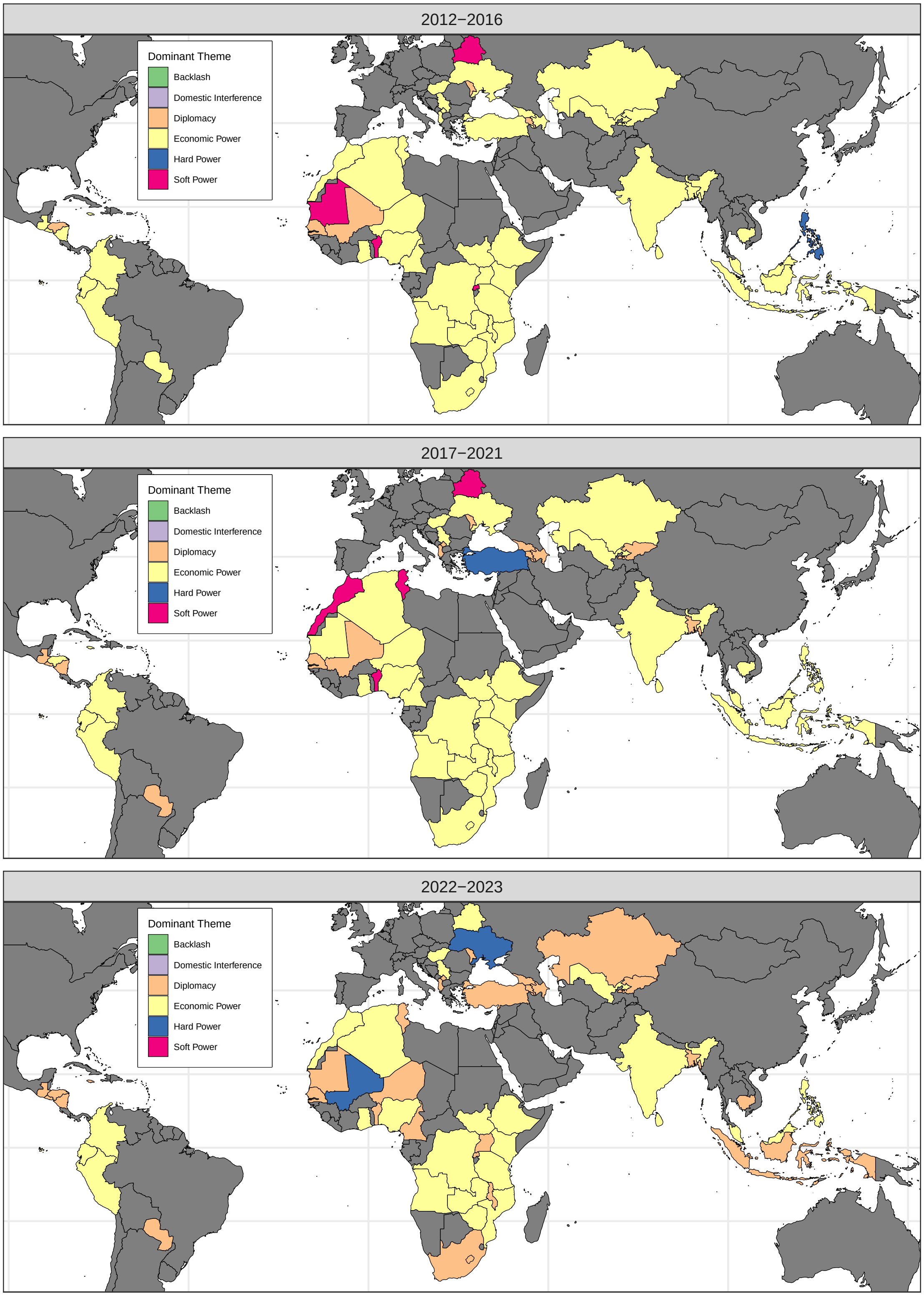

Figure 8: Share of RAI activity across RAI themes by country, measured as the normalized share of all articles reporting on RAI events that correspond with each theme.

Figure 9: Share of RAI activity across RAI themes over time, measured as the normalized share of all articles reporting on RAI events that correspond with each theme.

RAI activity by both Russia and China. The surge in Russian influence coincides with the invasion of Ukraine in February 2022 and continues at an elevated level through mid-2023.[21] By contrast, Chinese influence surges several months later, in August 2022. Surprisingly, this surge in Chinese influence follows a sustained period of decreased activity which corresponds with the declaration of COVID-19 as a pandemic in March 2020 and lasts through June 2022. Indeed, there is little evidence of China systematically increasing foreign activities around COVID diplomacy, despite international concerns to the contrary. If anything, China seems to have withdrawn internationally in the face of the pandemic.

Given that we also an increase in both Russian and Chinese influence, it seems unlikely that the increased RAI activity after 2022 is a product only of events in Ukraine. Combined with evidence

[21]Not all countries have been updated to include data for the last 2 months of the time-series, so the dramatic decrease in Russian activity for June 2023 is likely the result of composition changes.

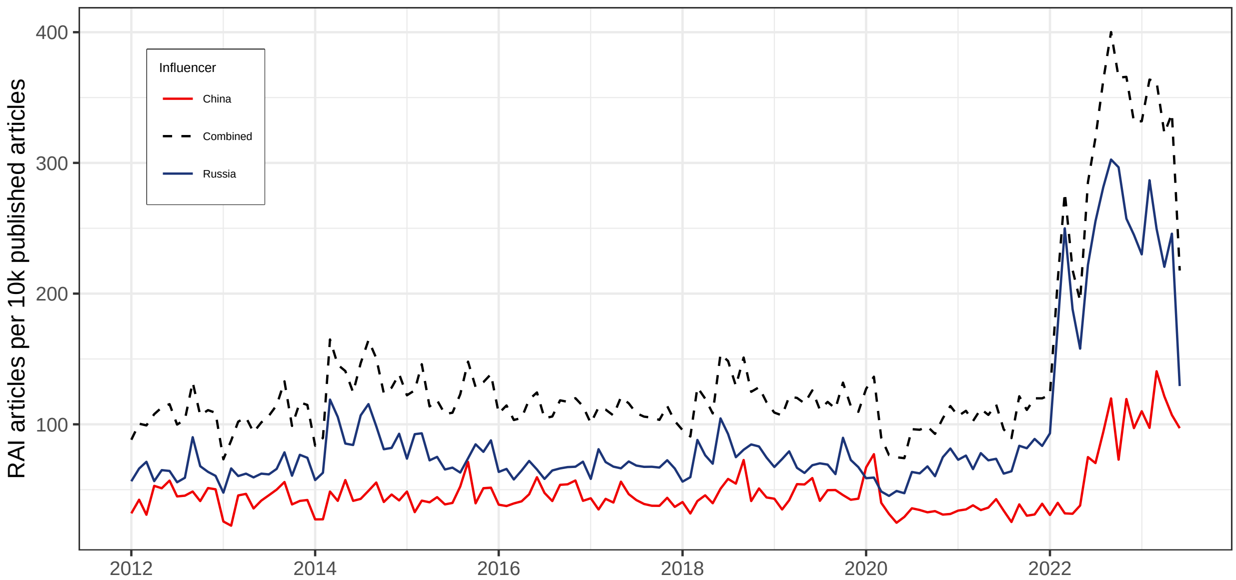

Figure 10: Levels of RAI activity over time, measured as the normalized share of all articles reporting on RAI events. Not all countries have been updated to include data for the last 2 months of the time-series, so dramatic changes may be the result of composition changes.

that this increased activity was the result of increased bilateral diplomacy, rather than minimizing concerns that the salience of these issues in international discourse is driving increased coverage.

To further isolate where these increases are coming from, Figure 11 plots the percentage change in average monthly RAI articles per 10,000 published articles across our three periods. Yellow points capture the percent change between the first period and the intermediate period, while green does capture the percent change between the intermediate and the final period. Looking at the yellow points, we see that for both Russian and Chinese influence, changes in the share of articles reporting on RAI activity between the first and the second period are relatively small and are roughly evenly split between positive and negative changes. Rather than seeing a broad surge in Chinese influence, 24 of 55 countries in our sample saw a decrease in activity between these periods. Meanwhile, 29 of 55 countries saw a decrease in Russian activity.

Turning to the change between the second and third period, captured by the green points, Russia's surge in influence events is clearly larger and more widespread than China's. 12 countries saw a decrease in Chinese influence, but only 3 in Russian influence. Although Ukraine had by far the largest absolute increase in RAI articles reporting on Russian influence events, it did not have the largest percentage increase. Instead, the largest percentage increases came in countries with relatively little Russian influence in previous periods. In fact, the percentage increase in articles reporting on Russian influence was negatively correlated with the level of Russian influence in the first period at -0.27, suggesting that Russian influence events increased most dramatically in places where Russia was not previously exerting influence. Importantly, this surge was only weakly correlated with China's share of total influence in the intermediate period at 0.25, suggesting Russia was not aggressively targeting its new influence at places where China held disproportionate sway.

Figure 11: Percent change in average monthly RAI articles per 10k published articles. Countries ordered by size of percent change in Russian influence between last two periods.

However, the increase in the percentage change of Russian and Chinese influence between the second and third period was correlated at 0.38, suggesting that both powers were increasing their influence in many of the same places.

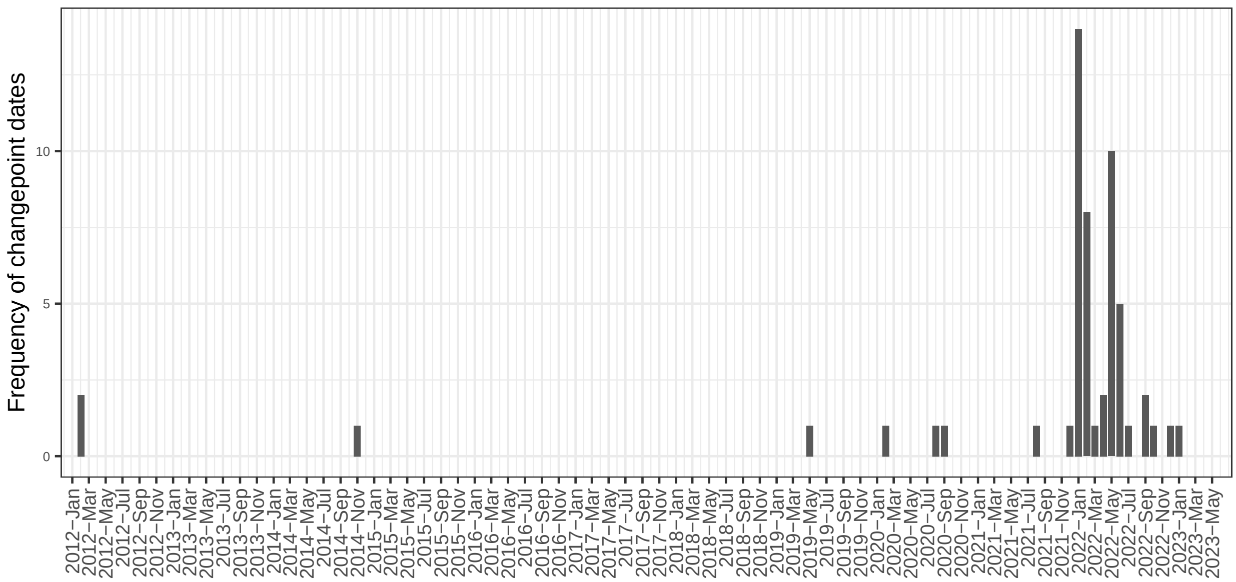

Figure 12: Frequency of changepoint dates in Russian influence time-series for all target countries. We identify one changepoint for each country. Changepoints capture increases in the mean of the time-series.

To provide additional evidence on the timing of this increase across countries, we conduct a changepoint analysis using binned segmentation (Fryzlewicz 2014). Changepoint methods divide a time series into two segments where each segment has its own statistical characteristics (e.g., mean, variance, etc.). Thus, the changepoint is located where the underlying characteristics change abruptly. In this case, we focus on abrupt changes in the mean level of influence for each influencing country. This exercise serves two functions. First, it gives a more precise estimation of when increased RAI activity started. Second, it provides additional evidence that the salience of Russian influence in international discourse is not driving increased coverage of influence events by domestic sources. In Figure 12, we present the result of models that allow each each country to have only one changepoint. The most common changepoint across countries is January 2022, the month before the February 24 invasion.[22] The two most common months are January and May, where we see 14 and 10 (including Ukraine) changepoints, respectively. We see only one changepoint (Nicaragua) in March 2023, the first full month of the invasion. Furthermore, while we see a similar concentration of changepoints in January for Diplomacy, Economic Power, and Hard Power, changepoints for Soft Power and Backlash, and for Domestic Interference are concentrated several months later in Mayy and July, respectively. Interestingly, the January surge might indicate that, despite the widespread surprise at the invasion among experts, Russian behavior was signalling an effort to prepare for the blowback its aggression would cause.

[22] This was the most common changepoint regardless of the number of changepoints that we allowed, running from 1-10.

Forecasting Russian and Chinese Influence

In this section, we assess whether it is possible to accurately forecast future levels of RAI activity. This question is critically important for for policymakers. Accurate forecasts of RAI activity could provide a valuable tool for strategic planning by actors working to counter authoritarian influence in strategically important countries. For this preliminary assessment, we limit our forecasts to the total levels of activity by each influencer country.

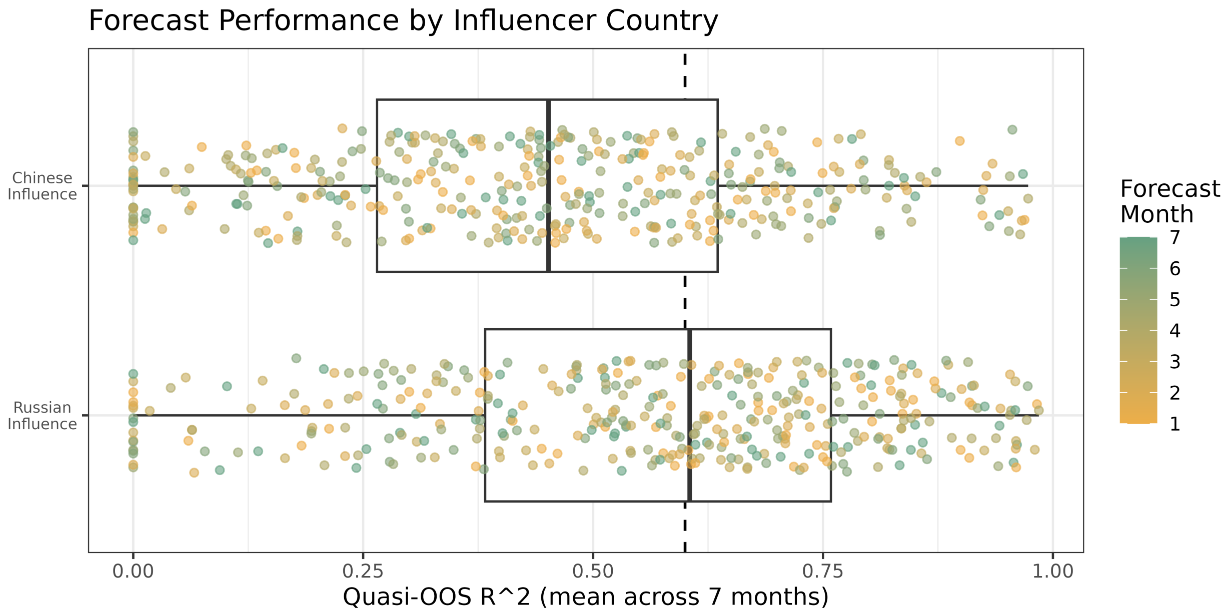

Figure 13: Forecast performance by influencer country and forecast length. The dashed line indicats an R^2 of 0.6.

To forecast RAI activity, we combine our RAI data with data on domestic political events from the Machine Learning for Peace project and high-frequency data on local economic conditions from TradingEconomics.[23] We then use machine learning to train statistical models to identify historical patterns between economic, social, and political conditions and future shifts in RAI activity. Specifically, these models look for correlations between the value of our target variables (the number of articles reporting on Russian and Chinese influence events per 10,000 articles published) and past values of our predictor variables (lagged measures of economic activity, domestic political events, and RAI events) between h and 12 months prior to that given month (h is set to the number of months into the future that we are forecasting).[24] We measure model performance using the

[23]For each country, we utilize every variable available on TradingEconomics that meet two criteria: the variable is updated at least quarterly (frequency) and we observe at least as many unique values as there are years in the data (variation). Our most common economic variables are Utility price inflation, Foreign Exchange Reserves, Transportation CPI, Imports, Inflation Rate Month-over-Month, Food Price Inflation, Interest Rates, Exports, Business Confidence, and Crude Oil Production.

[24]We use the glmnet package in R to estimate regularized linear regression models with cross validation. Regularized models are designed to guard against overfitting when using a large number of predictor variables by identifying the subset of variables that are most consistently predictive across different subsets of the data and selecting only those variables to use in a predictive model. To tune the regularization penalty, we use leave-one-out cross validation (LOOCV) to select the optimal risk consistent lambda [@homrighausen2014leave]. In addition to h month lags of predictor variables, we also include 12-month lags to capture potential seasonality.

Quasi-out-of-sample R^2. This measure reports the amount of out-of-sample variation in the target variable that is explained by the model in the best performing cross-validation fold. In our other forecasting applications, we use a cutoff of 0.6 to identify high-performing models.

Figure 13 shows model performance by influencer country and forecast length. Each point corresponds to the performence of models for different countries and different forecast lengths (1-7 months into the future). We see a great deal of variation in performance. However, we see many points above the dashed grey line, indicating models that we consider to be high-performing. We also see that the average forecast performance value is significantly higher for Russian influence relative to Chinese influence. Figure 14 shows the variation in performance across countries when averaging across the target variable and forecast length. Again, we see that several countries have very high average performance values. These results suggest that the RAI data may be able to provide valuable insights into future levels of RAI activity. Ultimately, we will need to monitor the ability of these models to accurately provide advanced warning of events that occur in the real world. However, our early results are very encouraging.

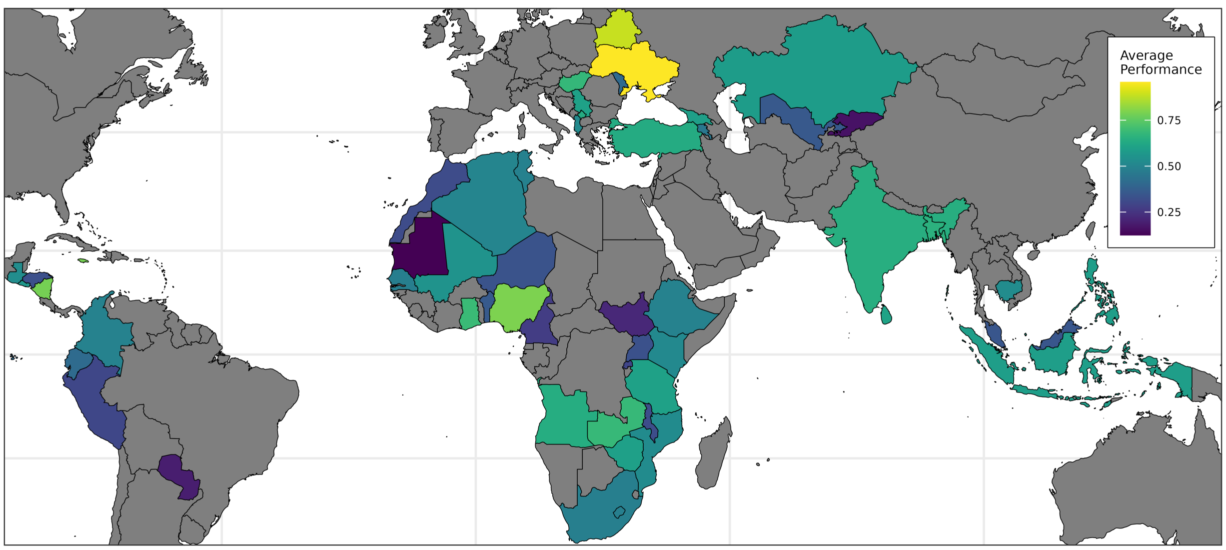

Figure 14: Forecast performance by target country. Colors capture average performance across forecast lengths.

Online Presentation of RAI Data and Forecasts

To facilitate use of these data and forecasts by both researchers and policymakers, we developed three interactive data dashboards. These dashboards are hosted on a dedicated page of the Machine Learning for Peace (MLP) webpage which can be accessed by clicking on the Authoritarian Influence Dashboards tab. Each of these three dashboards serves a unique purpose:

- The RAI Explorer allows users to visualize changes in the RAI data over time and across countries and at different levels of aggregation. This app includes 13 distinct figures and allows users to manipulate the time periods they want to see depicted.

- The RAI Data Dashboard allows users to examine each of the 22 RAI event types by countries and over time. This app allows users to manipulate the time period they want to see depicted and they countries they want to see comparisons between.
- The RAI Forecast allows users to see figures reporting the results of our forecasts of RAI activity up to 7-month into the future. Users can manipulate the specific country and influencer country they want to see forecasts for. Visualizations also include a heatmap that reports the predictor variables that have the greater influence on each model's decision making process.

For the life of the MLP project, the data and forecasts for every country will be updated every ninety days on a rolling basis. As the RAI data are updated, new data will be pushed to these apps, providing an easy way for users to track changing conditions.

Conclusion

In this paper, we introduce the Resurgent Authoritarian Influence (RAI) dataset, tracking the incidence of reporting on events indicative of 22 different types of Russian and Chinese influence. We also describe the construction of 21 index measures summarizing changes in RAI activity across time, space, and influencer country. Using these data, we present three main descriptive findings. First, Russian influence events have been more geographically concentrated in their geographic neighborhood, while China's have been less geographically proscribed. However, Russia's share of RAI influence has expanded geographically and significantly eroded Chinese dominance in recent years. Second, the highest levels of RAI activity occur in countries within Russia's immediate sphere of influence. Third, there has been a growing emphasis on bilateral Diplomacy over other modes of influence; this appears to be especially true for Russia. Third, we provide evidence for a dramatic increase in RAI activity in recent years. We also demonstrate the utility of these data for forecasting RAI activity; an exciting finding that could have great strategic value for policymakers looking to counter authoritarian influence.

We argue that the RAI dataset presents a valuable and timely resource. RAI provides unprecedented insight into the dynamics of global power politics, enabling a deeper understanding of how Russian and China employ strategies to shape the geopolitical landscape and how those strategies have changed over time. We hope that these data will enable scholars to make inferences about the underlying motives, mechanisms, and potential consequences of foreign authoritarian influence. We also believe that insights gleaned from this dataset can aid policymakers in crafting diplomatic and strategic approaches that align with ever-changing geopolitical realities, enhancing the ability to anticipate or mitigate the influence exerted by these countries. As global dynamics continue to evolve, this dataset stands as a powerful tool, fostering both scholarly discourse and effective policy formulation in an increasingly complex international arena.

References

Custer, S, A Dreher, TB Elston, A Fuchs, S Ghose, J Lin, A Malik, et al. 2021. "Tracking Chinese Development Finance: An Application of AidData's TUFF 2.0 Methodology." *Williamsburg, VA: AidData at William & Mary.*

Diamond, Larry. 2020. "Democratic Regression in Comparative Perspective: Scope, Methods, and Causes." *Democratization*, 1–21.

Dreher, Axel, Andreas Fuchs, Roland Hodler, Bradley C Parks, Paul A Raschky, and Michael J Tierney. 2019. "African Leaders and the Geography of China's Foreign Assistance." *Journal of Development Economics* 140: 44–71.

Dreher, Axel, Andreas Fuchs, Bradley Parks, Austin Strange, and Michael J Tierney. 2021. "Aid, China, and Growth: Evidence from a New Global Development Finance Dataset." *American Economic Journal: Economic Policy* 13 (2): 135–74.

Eady, Gregory, Tom Paskhalis, Jan Zilinsky, Richard Bonneau, Jonathan Nagler, and Joshua A Tucker. 2023. "Exposure to the Russian Internet Research Agency Foreign Influence Campaign on Twitter in the 2016 US Election and Its Relationship to Attitudes and Voting Behavior." *Nature Communications* 14 (1): 62.

Fryzlewicz, Piotr. 2014. "Wild Binary Segmentation for Multiple Change-Point Detection."

Gehring, Kai, Lennart C Kaplan, and Melvin HL Wong. 2019. "China and the World Bank: How Contrasting Development Approaches Affect the Stability of African States."

Isaksson, Ann-Sofie, and Andreas Kotsadam. 2018. "Chinese Aid and Local Corruption." *Journal of Public Economics* 159: 146–59.

Karlsen, Geir Hågen. 2019. "Divide and Rule: Ten Lessons about Russian Political Influence Activities in Europe." *Palgrave Communications* 5 (1): 1–14.

Kishi, Roudabeh, and Clionadh Raleigh. 2017. "Chinese Official Finance and State Repression in Africa."

Knutsen, Tora, and Andreas Kotsadam. 2020. "The Political Economy of Aid Allocation: Aid and Incumbency at the Local Level in Sub Saharan Africa." *World Development* 127: 104729.

Ma, Hong, and Yue Teng. 2018. "How Political Incentives Affect Chinese Outward Foreign Direct Investment: A UN Security Council Membership Perspective." *The World Economy* 41 (12): 3416–41.

McGlynn, J. 2023. "Why Russia Markets Itself as an Anti-Colonial Power to Africans." *Foreign Policy* 8.

Strange, Austin M, Axel Dreher, Andreas Fuchs, Bradley Parks, and Michael J Tierney. 2017. "Tracking Underreported Financial Flows: China's Development Finance and the Aid–Conflict Nexus Revisited." *Journal of Conflict Resolution* 61 (5): 935–63.