Two Types of Censorship? An Assessment of the Informational Autocracy Thesis in the Online Space

December 9, 2024

Keywords: Censorship, Trajectory Clustering, Takedown Request, Internet Blockage, State Capacity
JEL Codes: D72, K24, L82, L86, P51

- three explanations: elite mass, capacity, market size

1 Introduction

The nature of authoritarian rule has evolved significantly in recent decades. While many 20th century dictators relied on overt repression and ideological indoctrination, a new model of "informational autocracy" has emerged that is better adapted to a world of open borders, international media, and knowledge-based economies (Guriev and Treisman, 2019). Rather than terrorizing their citizens, these modern autocrats seek to manipulate information and shape public perceptions to maintain power. One attempt to capture how social media may solidify authoritarian tendencies while maintaining the government's popularity is the concept of informational autocracies: rule primarily through the manipulation of information.

Recent events have highlighted stark differences in how autocratic regimes approach internet censorship. During the 2022 protests in China, authorities invested $6.6 billion in direct filtering infrastructure and completely blocked social media access (Fedasiuk, 2021). In contrast, during Russia's 2021 elections, the government primarily relied on legal takedown requests to platforms, making over 2,500 requests to remove content while maintaining a facade of open internet access.[1] These contrasting approaches reflect fundamentally different strategies for controlling online information, yet few studies have systematically analyzed how technological capacity shapes these choices.

This paper studies how regime type (Informational Autocracy vs Overt Dictatorship) affects strategic content removal around elections. The key institutional distinction lies in their approach to censorship: Overt Dictatorships employ direct, top-down filtering infrastructure (like China's Great Firewall established in 2000), while Informational Autocracies must rely on indirect methods through platform requests due to their more constrained institutional environment (like Russia's content removal system implemented in 2012). The staggered emergence of these two distinct censorship approaches between 2000-2022 - with ODs establishing comprehensive filtering systems in the early 2000s and IAs developing platform-based removal strategies in the 2010s - provides a promising source of quasi-experimental variation in censorship constraints that we can leverage to study differential takedown behavior around elections.

We collect data from multiple sources to analyze censorship behavior. For internet blocking patterns, we use the Open Observatory of Network Interference (OONI) dataset which provides detailed blocking rates for platforms like Facebook, WhatsApp, and Telegram from 2017-2022. For content removal patterns, we analyze Google's Transparency Reports which document government takedown requests. We supplement this with expert assessments from Freedom House and V-Dem, along with IT capacity measures from the ITU. Our empirical strategy employs both clustering analysis to identify regime types and panel regressions.

First, using unsupervised machine learning techniques on a comprehensive dataset combining measures of political killings, tertiary enrollment rates, Polity2 scores, elected

[1] Documented in our analysis of Google Transparency Report data (Figure), showing Russia's dramatic increase in takedown requests reaching over 2,500 by 2022.

leaders, political prisoners, and torture, we identify two distinct clusters of autocratic regimes. Our clustering results align remarkably well with Guriev and Treisman's (2019, 2020a,b) theoretical classification of Informational Autocracies and Overt Dictatorships. Countries like Russia, Venezuela, and Malaysia are consistently classified as IAs across both our empirical approach and their theoretical framework, while countries like China, Syria, and Angola are identified as ODs. This validation through multiple methods and datasets provides strong support for the fundamental distinction between these regime types.

1.1 Related Literature

This paper contributes to the literature on media censorship. [2] Previous literature has studied censorship of authoritarian regimes in the context of political advertising (Simonov and Rao, 2022; Beazer et al., 2022), strategic legislation (Lorentzen, 2014), newspaper and blogs (Esarey and Xiao, 2011), and VPNs (Chen and Yang, 2019). The authors find that the shift to digital media has brought up newer methods to control the narrative of citizens and limit political expression. Our paper complements this literature in three ways. First, unlike previous works that dichotomize the presence or absence of censorship, we differentiate between different types of censorship in authoritarian regimes, introducing a novel perspective to the discussion. Second, while most papers focus on the impacts of censorship, we investigate the determinants and the mechanisms of different censorship practices. Third, existing studies typically examine a single-country context, our analysis instead extends to a panel of countries, providing external validity.

Our paper also contributes to the literature on comparative studies of censorship. Zittrain et al. (2017); Goldsmith (2007); Hellmeier (2016); King et al. (2013); Bunn (2015); Shen and Truex (2021) provide the background on how different countries censor, distinguishing censorship in countries with lower opposition from censorship in countries with more stability across politics and economics. These studies find that there exists a constrain for authoritarian governments when shaping narratives within their countries either by direct Internet filtering or media capture. Stier (2015) further explores the

[2] See Prat and Strömberg (2013) for an excellent survey.

variations in media freedom across political regimes, emphasizing the impact of regime types on media policies. Similarly, Sinpeng (2020) examines the resilience of authoritarian states in Southeast Asia against online political opposition, highlighting the combination of political authoritarianism and increasing Internet controls. Our paper adds to this discussion and enhances the literature in three aspects. The existing literature is qualitative, whereas we use a data-driven approach where we use clustering techniques along with transparency reports and Internet blocking data that allow us to examine the different determinants of censorship. Secondly, current literature focuses on the differences in censorship between democracies and authoritarian regimes - where we focus on differences within authoritarian countries. Through the data-driven approach, we are also able to offer a causal linkage that explains the different types of censorship within authoritarian regimes.

This paper also adds to the literature on the effects of Information Technology (IT) capacity, especially pertaining to censorship. Previous literature looks at the differences in the capacity of governments to conduct censorship, with more competent governments using their IT infrastructure more effectively and less competent governments finding workarounds for their lack of IT capability (Ananyev et al., 2019; Chang and Lin, 2020; Land, 2019; Williams, 2015). Closely related is the seminal work of Egorov et al. (2009), which argues that resources often dictate the presence of censorship in authoritarian regimes. Our paper differs in the following two ways. First, Egorov et al. (2009) measures censorship as a scale in the Polity survey, while we measure censorship by governments' behavior of sending takedown requests and blocking the applications. Second, Egorov et al. (2009) measures capacity by oil production and reserve, while we measure a country's IT capacity directly. Overall, we attempt to provide a causal understanding of the IT capacity as a key determinant of a country's censorship style.

The remainder of the paper proceeds as follows. Section 3 describes the data sources. Section 4 discusses the clustering methodology and supporting case studies. Section 5 shows the predictive power of the two types of censorship. Section **??** investigates the determinants of the two types of censorship. Section 7 concludes.

2 Background and Related Literature

1 A Brief Review of the Informational Autocracy Thesis 2 review collateral censorship theory

2.1 Key Aspects of Informational Autocracies

The concept of informational autocracy, as developed by Guriev and Treisman (2019, 2020b,a, 2022), is characterized by four key aspects that distinguish it from traditional authoritarian regimes:

2.1.1 Political Violence

One of the most striking features of informational autocracies is their reduced reliance on overt political violence. Unlike traditional dictatorships that often rule through fear and repression, informational autocrats seek to minimize visible acts of state violence against citizens. This is not to say that these regimes never use force, but rather that they employ it more selectively and covertly (Guriev and Treisman, 2019). The reason for this shift is twofold. First, overt violence can undermine the regime's carefully cultivated image of competence and benevolence. Second, in an age of global media and instant communication, acts of state violence are more likely to be exposed and condemned internationally, potentially leading to sanctions or other forms of pressure (Guriev and Treisman, 2020b). Instead of mass repression, informational autocrats may target a small number of high-profile opponents, often using legal pretexts rather than outright violence. They may also outsource repression to non-state actors or use more subtle forms of coercion, such as economic pressure or surveillance (Guriev and Treisman, 2022).

2.1.2 Official Ideology

Another key difference is the de-emphasis of official ideology. While 20th century totalitarian regimes often imposed comprehensive ideologies that sought to reshape society, informational autocracies tend to be more pragmatic and less ideological (Guriev and Treisman, 2019). These regimes may still promote certain values or national ideas, but they generally avoid the kind of all-encompassing ideological projects seen in communist or fascist states. Instead, they focus on more immediate concerns such as economic growth, stability, and national pride (Guriev and Treisman, 2020b). This ideological flexibility allows informational autocrats to adapt more easily to changing circumstances and to appeal to a broader range of citizens. It also makes it easier for them to maintain relationships with Western democracies, as they can present themselves as pragmatic partners rather than ideological adversaries (Guriev and Treisman, 2022).

2.1.3 Elections

Informational autocracies almost universally hold elections, but these are typically neither free nor fair. The purpose of these elections is not to allow for genuine political competition, but rather to create a veneer of democratic legitimacy and to gauge public opinion (Guriev and Treisman, 2019). These regimes invest heavily in ensuring favorable electoral outcomes, not primarily through outright fraud (although this may occur), but through more subtle means. These can include:

- Controlling the media environment to favor the incumbent
- Using state resources to campaign
- Harassing or disqualifying opposition candidates
- Manipulating electoral rules
- Co-opting potential opponents

The goal is to win elections convincingly enough to cement the regime's legitimacy, but not so overwhelmingly as to destroy the illusion of competition (Guriev and Treisman, 2020b).

2.1.4 Size of the Elite

The final key aspect is the size and role of the elite. In informational autocracies, the elite – defined as those with access to independent sources of information about the regime's true nature and performance – plays a crucial role (Guriev and Treisman, 2019). These regimes seek to keep the informed elite relatively small and manageable. This allows them to focus their efforts on co-opting or censoring a limited number of potential critics, rather than having to control a large, well-informed population (Guriev and Treisman, 2020b). At the same time, the elite needs to be large enough to run a modern economy and state apparatus. This creates a delicate balancing act for informational autocrats. They must provide enough freedom and opportunity to foster a competent elite, while also preventing this elite from becoming too large or independent (Guriev and Treisman, 2022). The regime's relationship with the elite is often based on a combination of co-optation (through economic benefits or political inclusion) and intimidation (through selective prosecution or other forms of pressure). The goal is to ensure that the elite either actively supports the regime or at least refrains from publicly criticizing it (Guriev and Treisman, 2020a).

These four aspects – limited political violence, pragmatic ideology, managed elections, and a controlled elite – work together to create a system of rule that is more subtle and potentially more durable than traditional authoritarianism. By maintaining a façade of democracy and competent governance, informational autocracies can achieve higher levels of genuine popularity and perceived legitimacy than their more repressive counterparts (Guriev and Treisman, 2020a).

Rewrite from here

The model of overt dictatorship was based on fear. Many rulers terrorized their citizens, killing or imprisoning thousands and deliberately publicizing their brutality to

deter opposition. Totalitarians such as Hitler, Stalin, and Mao combined repression with indoctrination into ideologies that demanded devotion to the state. They often placed barriers between their citizens and the rest of the world with overt censorship, travel restrictions, and limits on international trade.

However, in recent years, a less bloody and ideological form of authoritarianism has been spreading. From Hugo Chávez's Venezuela to Vladimir Putin's Russia, illiberal leaders have managed to concentrate power without cutting their coun- tries off from global markets, imposing exotic social philosophies, or resorting to mass murder. Many of these new-style autocrats have come to office in elections and managed to preserve a democratic facade while covertly subverting political institu- tions. Rather than jailing thousands, they target opposition activists, harassing and humiliating them, accusing them of fabricated crimes, and encouraging them to emigrate. When these autocrats kill, they seek to conceal their responsibility.

The key to such regimes, we argue, is the manipulation of information. Rather than terrorizing or indoctrinating the population, rulers survive by leading citizens to believe—rationally but incorrectly—that they are competent and public-spirited. Having won popularity, dictators score points both at home and abroad by mimicking democracy. Violent repression, rather than helping, would be counter- productive because it would undercut the image of able governance that leaders seek to cultivate.

Using newly collected data, we show that recent autocrats employ violent repression and impose official ideologies far less often than their predecessors did. They also appear more prone to conceal rather than to publicize cases of state brutality. By analyzing texts of leaders' speeches, we show that "informational autocrats" favor a rhetoric of economic performance and provision of public services that resembles that of democratic leaders far more than it does the discourse of threats and fear embraced by old-style dictators. Authoritarian leaders are increasingly mimicking democracy by holding elections and, where necessary, falsifying the results.

A key element in our theory of informational autocracy is the gap in political knowledge between the "informed elite" and the general public. While the elite accurately observes the limitations of an incompetent incumbent, the public is susceptible to the ruler's

propaganda. Using individual-level data from the Gallup World Poll, we show that such a gap does indeed exist in many authoritarian states today. Unlike in democracies, where the highly educated are more likely than others to approve of their government, in authoritarian states the highly educated tend to be more critical. The highly educated are also more aware of media censorship than their less-schooled compatriots. Where most previous models have assumed that formal political institutions constrain such leaders, we place the emphasis on a knowledgeable elite with access to independent media.

The reasons for this shift in the strategies of authoritarian leaders are complex. We emphasize the role of economic modernization, and in particular the spread of higher education, which makes it harder to control the public by means of crude repression. Education levels have soared in many nondemocracies, and the increase correlates with the fall in violence. But other factors likely contribute. International linkages, the global human rights movement, and new information technologies have raised the cost of visible repression. Such technologies also make it easier for regime opponents to coordinate, although they simultaneously offer new opportunities for surveillance and propaganda. The decline in the appeal of authoritarian ideologies since the end of the Cold War may also have weakened old models of autocracy.

3 Data Description

This section presents the various data sources utilized in our analysis of digital censorship and control mechanisms. Our research incorporates a diverse range of datasets, categorized into four main types: expert surveys, online behavior monitoring, legal documents, and supplementary data. Collectively, these resources provide a multidimensional perspective on censorship practices, encompassing aspects such as government conduct, technological capacity, users' rights, and the nature of restricted content.

To provide historical context and capture more extreme forms of authoritarian control, we incorporate the Authoritarian Control Techniques Database. This dataset, derived from Guriev and Treisman's 2019 study "Informational Autocrats," offers crucial information on political killings, the percentage of elected leaders, political prisoners, and the

prevalence of torture. By including these metrics, we aim to situate digital censorship practices within a broader framework of authoritarian governance strategies, allowing for a more comprehensive understanding of how digital control relates to other forms of state repression. These authoritarian control variables - specifically political killings, elected leadership, political imprisonment, and torture usage - serve as key inputs for our cluster analysis to differentiate between informational autocracies (IA) and overt dictatorships (OD). Following Guriev and Treisman's theoretical framework, we use these measures to identify regimes that rely primarily on information control versus those that maintain power through more overt repression.

Our analysis measures collateral censorship through content removal requests and direct censorship through platform blocking. The first measure, capturing collateral censorship, comes from the Google Transparency Report (2011-2022). This dataset documents government requests to remove content from Google's platforms, providing comprehensive coverage of formal content control attempts. Each observation includes both the number of requests and the quantity of items targeted for removal, allowing us to distinguish between targeted interventions and broad censorship campaigns. Our sample includes X observations across Y countries. Our second measure, capturing direct censorship, comes from the Open Observatory of Network Interference (OONI) dataset (2017-2022). This dataset documents platform-level blocking through automated network measurement tests, providing systematic evidence of service interruption. Each observation records the results of connection attempts to major digital platforms, allowing us to identify both selective blocking of specific services and broader internet control campaigns. Our sample includes 9,277 observations testing the accessibility of key digital services: Facebook, Telegram, WhatsApp, and VPN services.

Following (Rao, 2021), the main independent variable, share of term left, is constructed using data from the Database of Political Institutions 2017, supplemented with constitutional term length information from the Comparative Constitution Project Data, Political Handbook of the World Online Edition, and the Inter-Parliamentary Union's PARLINE database. This measure represents the proportion of a leader's constitutional term remaining at each point in time, calculated as $\frac{(\text{end date} - \text{current date})}{\text{term length}}$. For example, a leader one year into a four-year term would have 0.75 of their term remaining. Our sec-

ond timing variable, time until next election, draws from the Comparative Constitutions Project's "Characteristics of National Constitutions, Version 4.0" dataset, which provides detailed information on constitutional frameworks governing electoral cycles ('hosterm'). This allows us to precisely measure the temporal distance to upcoming elections across different constitutional structures.

To look at variation in legal frameworks and internet governance capabilities across countries, we incorporate several complementary datasets. The World Intermediary Liability Map (WILMap), maintained by Stanford Law School's Center for Internet and Society, provides detailed information on internet regulations and intermediary liability regimes worldwide. This dataset helps us account for cross-national differences in the legal infrastructure governing content removal and platform operations. We supplement this with the V-Dem Digital Society Survey (2011-2022), which covers 179 countries and provides their Regimes of the World (ROW) classification, enabling systematic comparison across regime types. For technological capacity measures, we draw from the World Telecommunication/ICT Indicators Database (WTI) 2023, using 'Employees in IT' and 'Investment in IT' as proxies for digital control capabilities. Additional control variables include GDP per capita, internet penetration, and tertiary education enrollment from the World Bank Development Indicators. Finally, we incorporate standardized assessments of filtering practices from the OpenNet Initiative (ONI) and data on digital influence campaigns from the Empirical Studies of Conflict Database (ESOC 4.0), which covers 127 campaigns across 38 countries.

Table 1 provides summary statistics for our key variables across two-year intervals from 2011 to 2020. Content removal activity shows a dramatic increase over this period, with Google takedown requests rising from an average of 40 requests in 2011-2012 to 1,633 requests in 2019-2020. This trend is even more pronounced in the number of items requested for removal, which increased from 638 in 2011-2012 to 19,810 in 2019-2020, reflecting intensifying government efforts to control online content. The sharp increase is particularly notable after 2015, coinciding with growing state capacity for content regulation and filtering, as indicated by the rising filtering capacity index (from 0.997 to 1.300) and content regulation capacity index (from 1.018 to 1.085).

The control variables demonstrate important trends in technological and political development across our sample period. IT sector development shows substantial growth, with IT employment increasing from 40,296 thousand to 112,130 thousand workers. However, IT investment exhibits more volatility, ranging from 3,012M USD to 4,567M USD across the period. Political characteristics remain relatively stable, with Polity2 scores showing a slight upward trend from 4.03 to 4.24 (through 2016), while political killings increased from 15,936 to 19,338. The tertiary enrollment rate also shows steady growth from 37.74% to 42.65%, indicating expanding higher education access across our sample countries. These trends suggest the importance of controlling for both technological and institutional development in our analysis of content removal patterns.

Variable/Interval	2011-2012	2013-2014	2015-2016	2017-2018	2019-2020
FOTN Total	38.5	39.92	39.39	38.76	37.77
A: Obstacles to Access	12.65	13.39	13.93	14.24	14.33
B: Limits on Content	19.96	21.31	21.04	20.62	20.37
C: Violations of User Rights	20.09	19.84	18.79	17.83	17.06
Filtering Capacity	0.997	1.003	1.175	1.267	1.300
Shutdown Capacity	0.850	0.754	0.839	0.923	0.989
Cyber Security Capacity	0.530	0.400	0.445	0.506	0.479
Content Regulation Capacity	1.018	0.903	1.048	1.109	1.085
Google Takedown Requests	40	201	570	1506	1633
Google Item Requests	638	706	1365	12,504	19,810
IT Employees (in thousands)	40,296	88,578	86,559	106,830	112,130
IT Investment (in million USD)	4,447M	3,012M	4,567M	4,038M	3,678M
Polity2	4.03	4.16	4.24	-	-
Political Killings	15936.63	17218.43	19338.74	-	-
Tertiary Enrollment	37.74	41.96	42.65	-	-
OONI Data (avg confirmed + anomaly)					
Facebook	-	-	0.455	6.49	17.4
WhatsApp	-	-	1.06	2.87	3.53
Telegram	-	-	-	15.0	20.3
VPN	-	-	-	-	3.19

Note: All statistics represent averages over two-year intervals. IT statistics are based on available data from ITU. OONI data represents the average of confirmed and anomaly counts for the specified platforms across two-year intervals. ”-” indicates data not available for the interval.

Table 1: Combined Summary Statistics

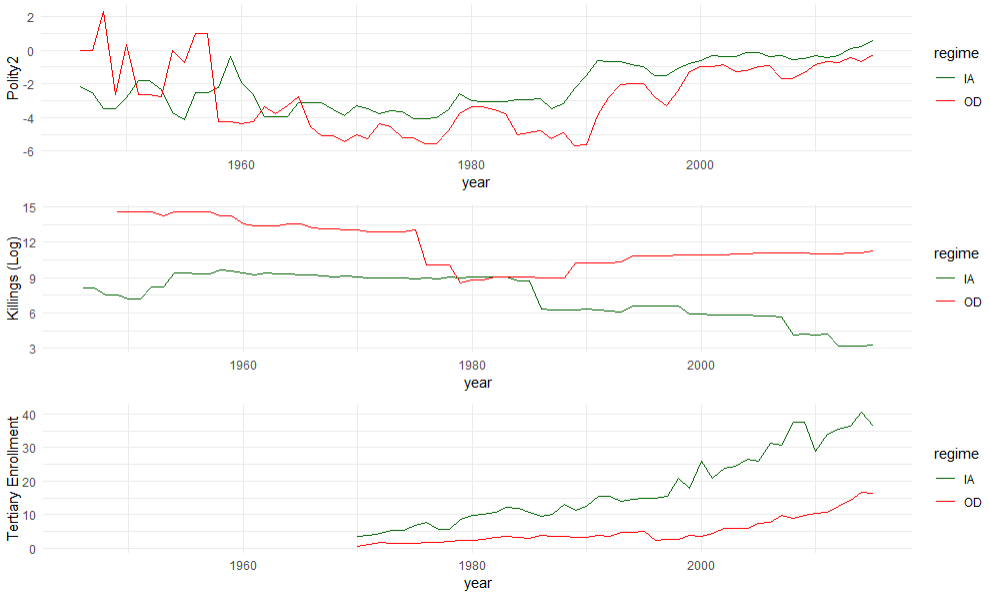


Figure 1: Authoritarian Control Techniques variables

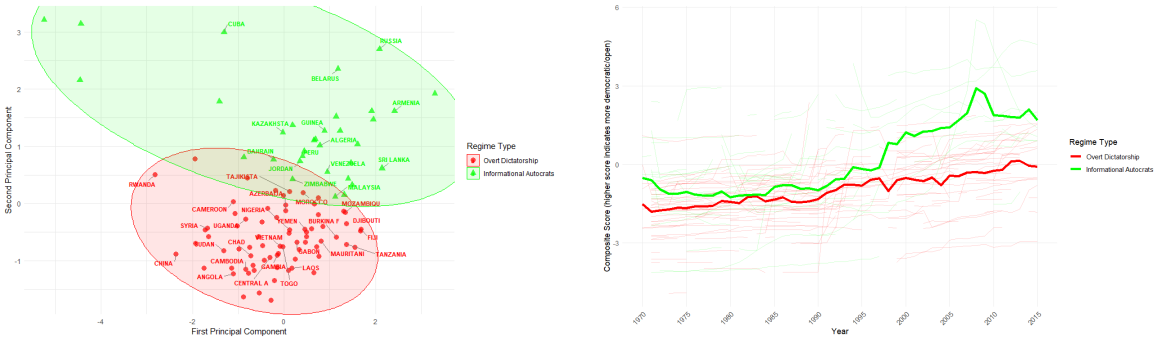
4 Classifying the Overt Dictatorship and the Informational Autocracy

4.1 Methodology

We employ a clustering approach to identify distinct patterns in autocratic control techniques across countries. Our analysis combines Principal Component Analysis (PCA) with K-means clustering, focusing on three key dimensions: political violence (average killings, political prisoners, torture prevalence), electoral institutions (Polity2 scores, percentage of elected leaders), and elite engagement (tertiary enrollment rates).

To address dimensionality concerns, we first apply PCA to these six variables. PCA transforms these correlated variables into uncorrelated principal components while preserving the underlying variance structure. We then apply K-means clustering (k=2) to the first two principal components, which capture the majority of variation in our data. This allows us to empirically identify two distinct groups of autocratic regimes based on their control strategies.

The resulting classification, visualized in Figure 1, reveals systematic differences in how different autocracies combine various control mechanisms. This data-driven approach provides an empirical foundation for examining how these distinct regime types differ in their approach to online content control.



(a) Cluster Plot (b) Trajectory Comparision

Figure 2: Clustering Results

4.2 Clustering Results

Figure 2(a) presents the outcomes of our PCA-K-means clustering. It depicts countries positioned along two principal components, with the x-axis representing the first principal component and the y-axis the second. This visualization reveals two distinct clusters: Informational Autocrats, denoted by green triangles, and Overt Dictatorships, represented by red circles. The Informational Autocrats cluster exhibits a broader distribution along both principal components, particularly extending higher on the second component. In contrast, the Overt Dictatorship cluster appears more compact and is situated lower on both components. This distribution suggests that Informational Autocrats display a wider range of characteristics, potentially indicating more nuanced or varied approaches to governance within this category.

Notably, the plot identifies specific countries within each cluster. For instance, Russia and Cuba are classified as Informational Autocrats, while China and Syria fall within the Overt Dictatorship category. The positioning of countries like Azerbaijan and Tajikistan in the overlap between clusters is particularly intriguing, suggesting these regimes

may exhibit characteristics of both autocratic types. This overlap underscores the complex nature of autocratic governance and the potential for hybrid forms that defy simple categorization.

Figure 2(b) illustrates the temporal trajectories of these two clusters from 1970 to 2015. The y-axis represents a composite score derived from our variables, with higher scores indicating more democratic or open characteristics. [3] Both clusters demonstrate an overall upward trend over the 45-year period, suggesting a general movement towards more open governance across autocratic regimes. However, the trajectories of the two clusters differ markedly in both their relative positions and dynamics. Informational Autocrats, represented by the green line, consistently maintain a higher composite score throughout the period. This suggests that these regimes tend to exhibit more democratic or open characteristics compared to Overt Dictatorships. The trajectory of Informational Autocrats also shows greater volatility, particularly post-2000, with pronounced fluctuations in the composite score. This volatility might reflect the challenges these regimes face in balancing control with the appearance of openness, or it could indicate more responsiveness to global events and pressures.

In contrast, the trajectory of Overt Dictatorships, depicted by the red line, shows a more gradual and steady increase over time. While still trending upwards, indicating some movement towards more open characteristics, this group maintains a consistently lower composite score compared to Informational Autocrats. The relative stability of this trajectory might suggest more resistance to change or a more consistent approach to governance among Overt Dictatorships. A particularly noteworthy feature of these trajectories is the divergence that begins around 1990. From this point, the composite score for Informational Autocrats increases more steeply than that of Overt Dictatorships. This divergence coincides with significant global events, including the end of the Cold War and the acceleration of globalization, which may have differentially impacted these two types of autocratic regimes.

[3] This composite score enables a quantitative comparison of regime characteristics over time, capturing multiple dimensions of autocratic governance in a single metric.

5 Direct Censorship versus Collateral Censorship

1 comparison of legislation: IA has more laws. Q. are they on intermediary liability? also country-specific case study goes here. 2 comparison of blocking behavior: OD blocks 3 comparison of legal takedowns 4 consequence of the censorship strategy: ONI - show OD is more successful in blocking all types content, FOTN - show that IA has better reputation internationally, V-Dem capacity - show that OD is able to develop capacity more because of no institutional constraint

5.1 Comparison of Legislation

Figure 3 illustrates the distribution of internet intermediary liability frameworks across Informational Autocracies (IA) and Overt Dictatorships (OD), drawing on data from Stanford Law School's World Intermediary Liability Map (WILMap). The data categorizes legal developments into four types: Decisions, Law, Other, and Pending Proposal, providing a comprehensive view of how different regime types formalize their digital control mechanisms. Most notably, IAs demonstrate approximately twice the frequency of formal legal "Decisions" compared to ODs (7.5 versus 3.0 average entries), and maintain higher frequencies across all regulatory categories.

This pattern aligns with theoretical expectations about regime differences in digital control strategies. While IAs show a clear preference for legalistic approaches to content regulation, evidenced by their higher counts of both "Decisions" and "Law" entries, ODs display a more modest formal regulatory footprint. However, these differences should not be interpreted as variations in control intensity, but rather as distinct approaches to formalizing digital governance. The lower frequency of formal mechanisms in ODs likely reflects a preference for direct, less documented forms of control rather than an absence of regulation.

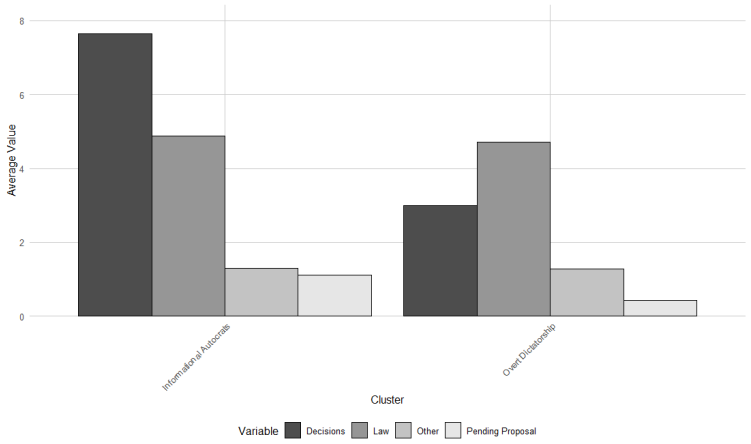


Figure 3: Comparison of Law - Wilmap Stanford Analysis

5.1.1 Type I Censorship: Overt Dictatorship

Countries in this group pervasively use the Internet as a means of state control over society. They emphasize sovereignty over the Internet and build closed borders for their Internet users. They consistently screen content of all types, and rarely is there an explanation of the reasons for a block. This group includes China, Iran, Saudi Arabia, Bahrain, etc.

China. China perceives the Internet as both an engine for development and a potential threat to governance. The rapid growth of its online community, now surpassing a billion users, reflects its deep integration into society. However, this expansion is matched by the government's efforts to mitigate risks through extensive information control and state oversight.

Chinese government's conception of the Internet is strikingly different from that in democratic countries. Internet censorship in China is based on the concept of digital sovereignty (*wangluo zhuquan*), which frequently serves as the justification for the government to wall off Internet access within the country. Stringent laws governing speech and telecommunication in China define the legal landscape. These include the "Measures for the Administration of Internet Information Services" and the "Provisions on the Administration of Internet News Information Services." These regulations impose tight controls on content providers, mandating the censorship of content deemed harmful to state se-

curity or social harmony. The framework establishes a restrictive environment, aiming to curtail the dissemination of sensitive information while maintaining state authority.

In 2020, the Chinese government and its affiliated entities allocated over $6.6 billion towards Internet censorship and surveillance, underscoring the priority of digital control in the political agenda (Fedasiuk, 2021). This financial commitment supported a variety of initiatives aimed at building a state-run infrastructure and workforce to monitor the country's 900 million Internet users. Central to these efforts was the Great Firewall built by the Cyberspace Affairs Commissions (CACs) and Public Security Bureaus (PSBs). Equipped with an unprecedented labor force, the Great Firewall can completely block access to a set of sensitive words and videos or exclude a set of users and regions from the Internet. For example, China conducts targeted "cleanup" campaigns to swiftly remove or block information related to sensitive subjects such as Tibet and Xinjiang, reflecting its proactive stance on censorship.

Comparatively, China's censorship model is markedly more direct and resource-intensive than those employed by countries with less financial commitment to information control. Unlike nations that may rely on subtler forms of media manipulation due to budgetary constraints, China invests heavily in a broad spectrum of censorship technologies. First, sophisticated keyword filtering, IP blocking, and software like Green Dam Youth Escort, specifically designed to monitor and restrict content (Qiang, 2011). Second, China bans VPNs together with the circumvention of it. Last but not least, China blocks multi-national social media platforms and substitutes them with state-supported ones. For instance, Google is substituted with Baidu, WhatsApp is substituted with WeChat, and Twitter is substituted with Weibo. The government also enforces real-name registration on these platforms, effectively limiting anonymous use of the Internet.

Iran. Iran perceives the Internet as both a potential threat to the ideological and moral fabric of the Islamic Republic and a technological advancement to be harnessed for economic growth. Despite recognizing its utility for fostering innovation and economic development, the government has implemented one of the world's most extensive technical filtering systems. The Internet's initial period of relative freedom has given way to a highly regulated digital sphere, overseen by a complex regulatory framework involving

multiple government agencies, including the Revolutionary Guard.

In Iran, the legal and regulatory framework for controlling speech online is extensive, rooted in constitutional mandates against anti-Islamic practices and broadened by decrees from the Supreme Council of the Cultural Revolution. The Press Law of 1986, with its amendments, requires online publications to obtain licenses, subjecting them to the same strictures as traditional media. The Cybercrimes Bill of 2008 obliges ISPs to block "forbidden" content and report violations, enhancing the state's control over the Internet. These laws, alongside directives from various government agencies, create a comprehensive legal structure for Internet censorship in Iran.

Iran's government has invested heavily in limiting its citizens' access to the global Internet, allocating at least $4.5 billion towards the development of a domestic intranet known as the National Information Network (NIN) [4]. This initiative, launched as early as 2005, aims to confine data requests within national borders, enabling stricter censorship and control over online content. The move to a domestic intranet reflects the regime's broader strategy of Internet restriction, which includes blocking access to thousands of websites.

Iran, while also having a pervasive censorship regime, tends to focus more on controlling and monitoring content that challenges the political and religious status quo. The Islamic Republic of Iran has expanded its technical filtering system, which is one of the most extensive globally. Iran has created a centralized system that complements the filtering conducted at the ISP level and has developed its own technology for identifying and blocking objectionable websites, thus reducing reliance on Western technologies. In addition to filtering, Iran employs legal actions and extensive surveillance to deter and control dissenting voices online. For instance, during the contentious 2009 presidential elections, political websites were specifically targeted for blocking. Iranian regulatory agencies have expanded, with the Revolutionary Guard playing an active role in enforcing content standards. This has contributed to an online environment that fosters self-censorship and discourages dissent.

[4] See Iran's Regime spends billions to limit citizens' Internet access

5.1.2 Type II Censorship: Informational Autocrats

Countries in this group actively compete and dominate in their cyberspace with pro-government political messages and requests. Instead of enforcing pervasive control on Internet access, governments in this group prefer to employ second- and third-generation techniques such as legal and technical instruments and state-sponsored influence campaigns to shape the information environment and stifle dissent and opposition. This group includes Russia, Thailand, Turkey, Egypt, etc.

Russia. Russia's approach to the Internet reflects a nuanced strategy beyond direct censorship, focusing instead on shaping the information space with pro-government messages. This strategy aligns with official doctrines like the information security doctrine, suggesting a broad, strategic view of cyberspace as a domain for state influence.

The legal landscape in Russia offers a complex mix of freedoms and controls. The Constitution of the Russian Federation guarantees free speech and privacy rights, but laws like the Law on Communications and the Law on Personal Data introduce nuanced restrictions, especially concerning state security and personal data processing. The Law on Information, Information Technologies, and Protection of Information, coupled with specific presidential decrees, establishes a framework that indirectly subjects Internet content to oversight, reflecting an overarching desire to monitor and potentially control the digital dialogue.

In Russia, the interplay between military expansion and media influence is evident in the substantial funds allocated for state propaganda. In 2022, the Russian Federation significantly overshot its budget for mass media, spending around 143 billion RUB (1.9 billion USD), with projections already setting 2023's propaganda budget at 1.6 billion USD [5]. This funding primarily supports pro-Kremlin narratives through major agencies such as VGTRK, RT, and Rossiya Segodnya, emphasizing internal dissemination and international outreach. Additionally, the defense sector's media arm, Zvezda, received nearly double its previous funding, reflecting a broader strategy to bolster military and media capabilities concurrently.

[5] See Kremlin spent 1.9 billion USD on propaganda last year, the budget exceeded by a quarter

Russia's approach to managing its information space subtly contrasts with pervasive control regimes, relying less on overt censorship and more on influence operation (IO). Unlike direct methods such as comprehensive filtering and blocking used in PC regimes, Russia employs a mix of legal intimidation, cyber capabilities for political influence, and pro-government propaganda to shape online discourse. For instance, Russian authorities legally pressure ISPs to self-censor under anti-extremism laws, while suspected state-sponsored cyberattacks against other nations and the nurturing of a pro-Kremlin blogger network exemplify their IO tactics. These methods enable the Russian state to maintain an appearance of an open Internet while covertly steering public opinion and suppressing dissent.

Turkey. Turkey's approach to the Internet reflects its sensitivity towards defamation and inappropriate content. This sensitivity has led to the closure of both local and international websites. The government views the Internet as a crucial sphere for regulation, balancing its EU aspirations with a tight grip on online content deemed inappropriate.

Turkey's legal landscape for speech and telecommunications is shaped significantly by its ambition to join the European Union, prompting substantial legal reforms. However, the country's Penal Code restricts freedoms by criminalizing speech that insults the Turkish identity or government institutions. The Internet's regulatory framework is further defined by the Law No. 5651, establishing the legal grounds for filtering and blocking mechanisms against illegal online information. Despite liberalizing the telecommunications market in 2005, Turk Telekom retains a dominant position, highlighting a partial monopoly in fixed-line services and broadband Internet operation.

Turkey has an annual budget of around 680 million lira ($38 million), for what they term the Directorate of Communications. This Directorate consists of over 90 offices in Turkey and around the world that carries out communication campaigns and restricts content available online as the news cycle progresses [6].

Turkey's approach to Internet governance embodies Information Operation (IO) strategies. While PC is characterized by heavy filtering and outright blocking of a wide spectrum of content, Turkey's strategy involves targeted content removal based on spe-

[6] See Insiders reveal how Erdogan tamed Turkey's newsrooms

cific legal pretexts. For example, Turkey has repeatedly employed Law No. 5651 to justify the temporary blocking of entire platforms like YouTube and Wordpress. This selective method indicates a preference for intermittent control tailored to immediate state concerns rather than a continuous, broad censorship apparatus typical of PC.

5.2 Comparison of Internet Blocking Behavior

To analyze differences in digital control strategies, we construct a 'blocking rate' metric using data from the Open Observatory of Network Interference (OONI). This rate is calculated as:

$$\frac{\text{anomaly\_count} + \text{confirmed\_count}}{\text{measurement\_count} - \text{failure\_count}}$$

where anomalies represent patterns consistent with intentional blocking but not definitively confirmed, confirmed counts indicate verified instances of state-mandated blocks, measurement count captures all connection attempts, and failure count represents connection failures due to technical issues rather than deliberate blocking. This formulation helps distinguish intentional blocking from network irregularities, providing a more accurate measure of state-directed censorship.

Figure 4 reveals a stark contrast in internet control approaches between Informational Autocracies (IA) and Overt Dictatorships (OD) from 2018 to 2022. ODs maintain approximately double the blocking rate of IAs (averaging 0.08 versus 0.04), with particularly pronounced spikes in early 2018 and mid-2020. This pattern holds across various digital platforms, as shown in Figure 5, though with notable variations. WhatsApp exhibits the largest disparity (OD rate of 0.10 versus IA rate of 0.02), followed by Facebook (0.09 versus 0.03), while VPN blocking shows more modest differences (0.07 versus 0.05), suggesting both regime types actively target encryption circumvention tools.

The statistical significance of these differences is confirmed by t-tests presented in Figure 6, with an overall p-value of $< 2e - 16$. The boxplots demonstrate systematically

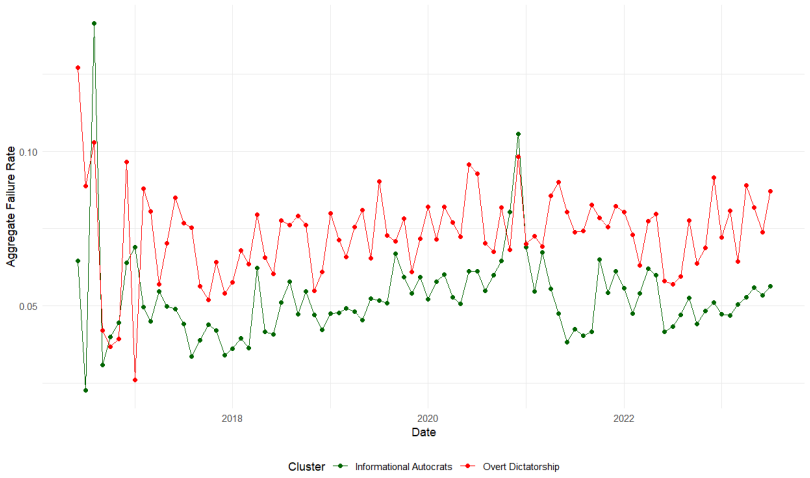
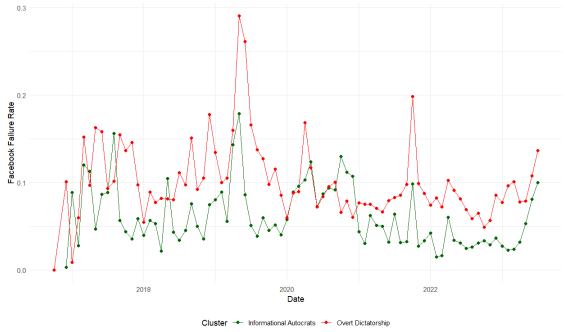
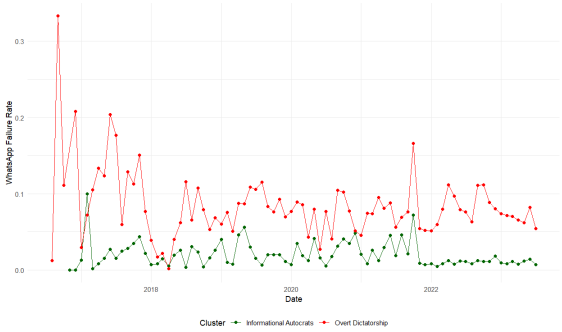


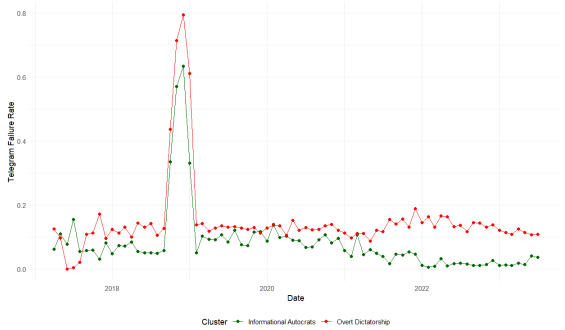
Figure 4: OONI - Aggregate Shutdown by Clusters



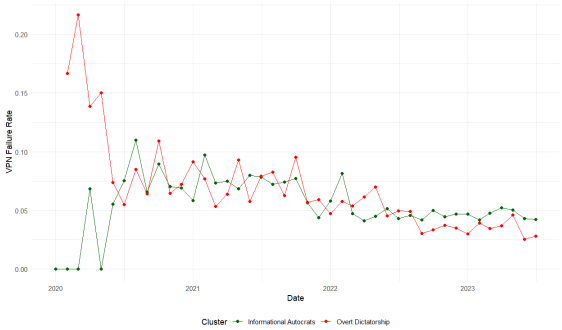
(a) Shutdowns for Facebook



(b) Shutdowns for WhatsApp



(c) Shutdowns for Telegram



(d) Shutdowns for VPN

Figure 5: OONI - Internet Shutdowns by Clusters across Applications

higher median blocking rates for ODs across all services, with particularly narrow confidence intervals for aggregate measures and Facebook blocking. Platform-specific tests reveal the strongest statistical differences in messaging applications (WhatsApp: $p < 0.001$) and social media (Facebook: $p < 0.001$), while VPN blocking shows a weaker, though still

significant, difference ($p < 0.01$).

These findings align with theoretical expectations about regime differences in digital control strategies. ODs' higher blocking rates across communication platforms suggest a willingness to accept the economic and social costs of overt censorship. In contrast, IAs' systematically lower blocking rates, particularly for social media, indicate a preference for more subtle control methods, potentially including targeted content removal, algorithmic manipulation, or legal pressures - strategies that maintain a facade of digital openness while still achieving information control objectives. The convergence in VPN blocking rates suggests that when faced with direct challenges to their control capacity, both regime types resort to similar technical countermeasures.

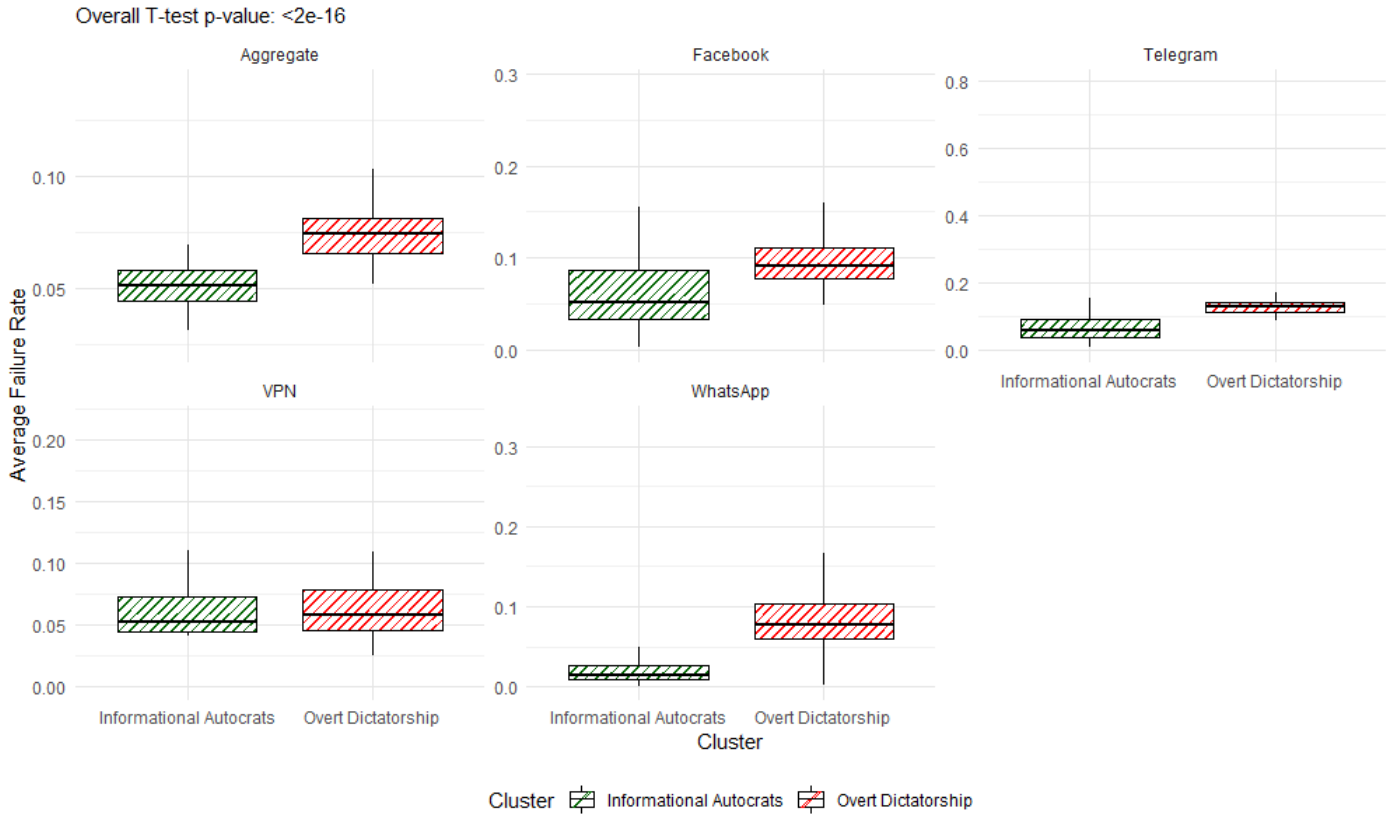
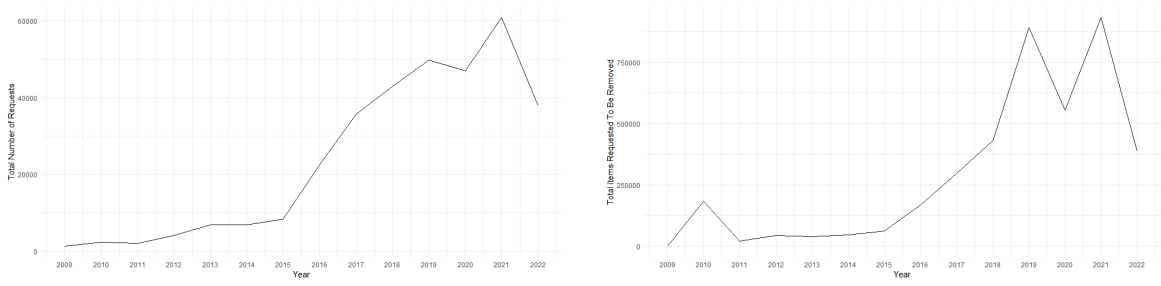


Figure 6: OONI T-Tests

5.3 Comparison of Legal Takedowns

Figures 7(a) and 7(b) reveal a dramatic increase in content removal activities from 2009 to 2022. The total number of requests shows particularly strong growth after 2015, rising from under 10,000 to over 60,000 by 2021, followed by a slight decline in 2022. This trend is even more pronounced in the volume of items requested for removal, which exhibits exponential growth post-2015, approaching 1 million items by 2021. This surge suggests a broader shift toward formal content moderation mechanisms across all regime types, though the underlying motivations vary significantly.



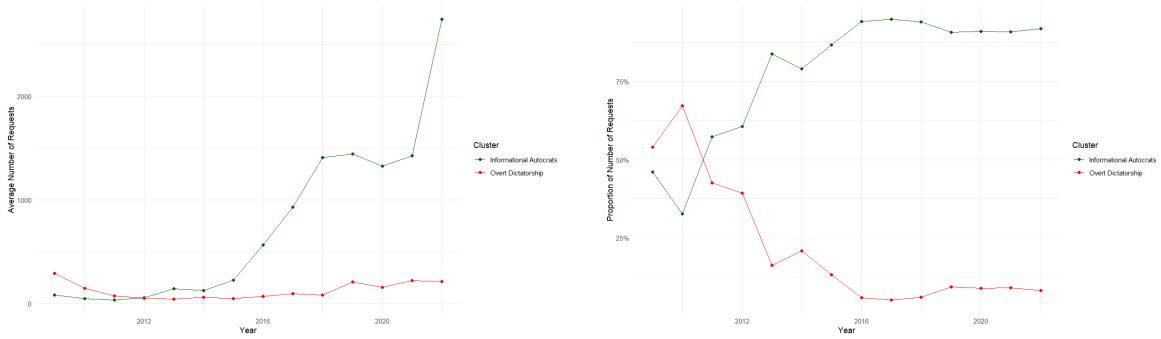
(a) Total Number of Requests over time

(b) Total Items Requested to be removed over time

Figure 7: Google Transparency Report - Total Requests and Removals over Time

When disaggregated by regime type, the data reveals striking differences in how Informational Autocracies (IA) and Overt Dictatorships (OD) approach content removal. Figures 8(a) and 8(b) demonstrate that IAs have dramatically intensified their use of formal removal requests since 2015, reaching over 2,500 average requests annually by 2022 and accounting for more than 90% of all requests. This sharp increase coincides with several high-profile political events and increased platform scrutiny of state-sponsored disinformation campaigns. In contrast, ODs maintain consistently low request levels, never exceeding 500 annually, suggesting a preference for direct control mechanisms over platform-mediated content removal.

The divergence in content targeting strategies is further illuminated in Figures 9(a) and 9(b). IAs not only make more requests but also demonstrate sophisticated bulk removal strategies, with requests peaking at over 20,000 items in 2021. The exponential growth in items targeted by IAs suggests increasingly systematic approaches to content



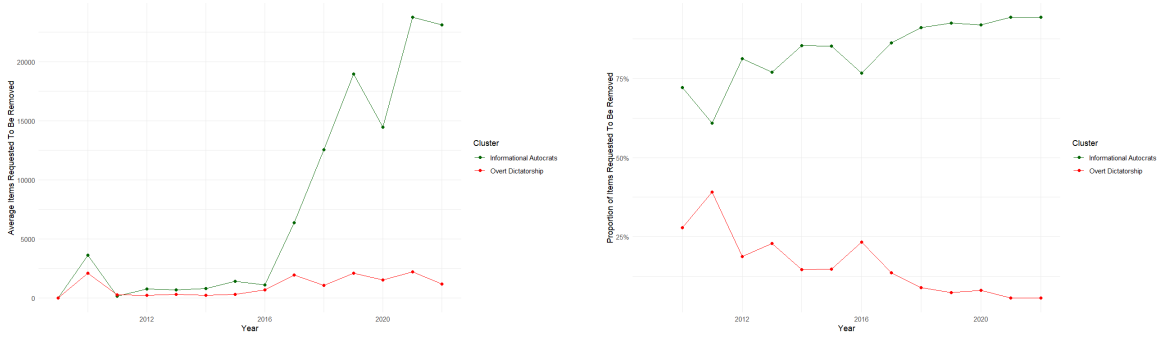
(a) Number of Requests by Clusters

(b) Proportion of Total Number of Requests by Clusters

Figure 8: Google Transparency Report - Total Number of Requests by Clusters

removal, possibly employing automated identification of "problematic" content. ODs, meanwhile, maintain relatively constant and modest removal requests, rarely exceeding 5,000 items and showing little variation over time, indicating a more selective or perhaps less sophisticated approach to platform-based content moderation.

The statistical robustness of these differences is confirmed in Figure 10. IAs show substantially higher medians for both metrics - approximately 4,000-4,500 items and 400-450 requests - compared to ODs' much lower medians of 500-1,000 items and 50-100 requests. The wider confidence intervals for IAs, particularly in items requested for removal, suggest considerable variation in content removal strategies among IA regimes, potentially reflecting different stages of developing their digital control capabilities or varying domestic political pressures.



(a) Items Requested to be removed by Clusters

(b) Proportion of Total Items Requested to be removed by Clusters

Figure 9: Google Transparency Report - Items Requested to be Removed by Clusters

These patterns reveal fundamentally different approaches to digital control. ODs' minimal engagement with formal removal mechanisms likely reflects their reliance on direct censorship infrastructure - including internet shutdowns, IP blocking, and DNS manipulation - making content removal requests largely redundant. In contrast, IAs' extensive and increasing use of formal channels suggests a more sophisticated strategy that leverages existing platform mechanisms to shape online discourse while maintaining plausible deniability. This approach, while potentially less immediately effective than direct censorship, offers IAs greater flexibility and lower reputational costs in managing online information, particularly in maintaining international legitimacy and business relationships.

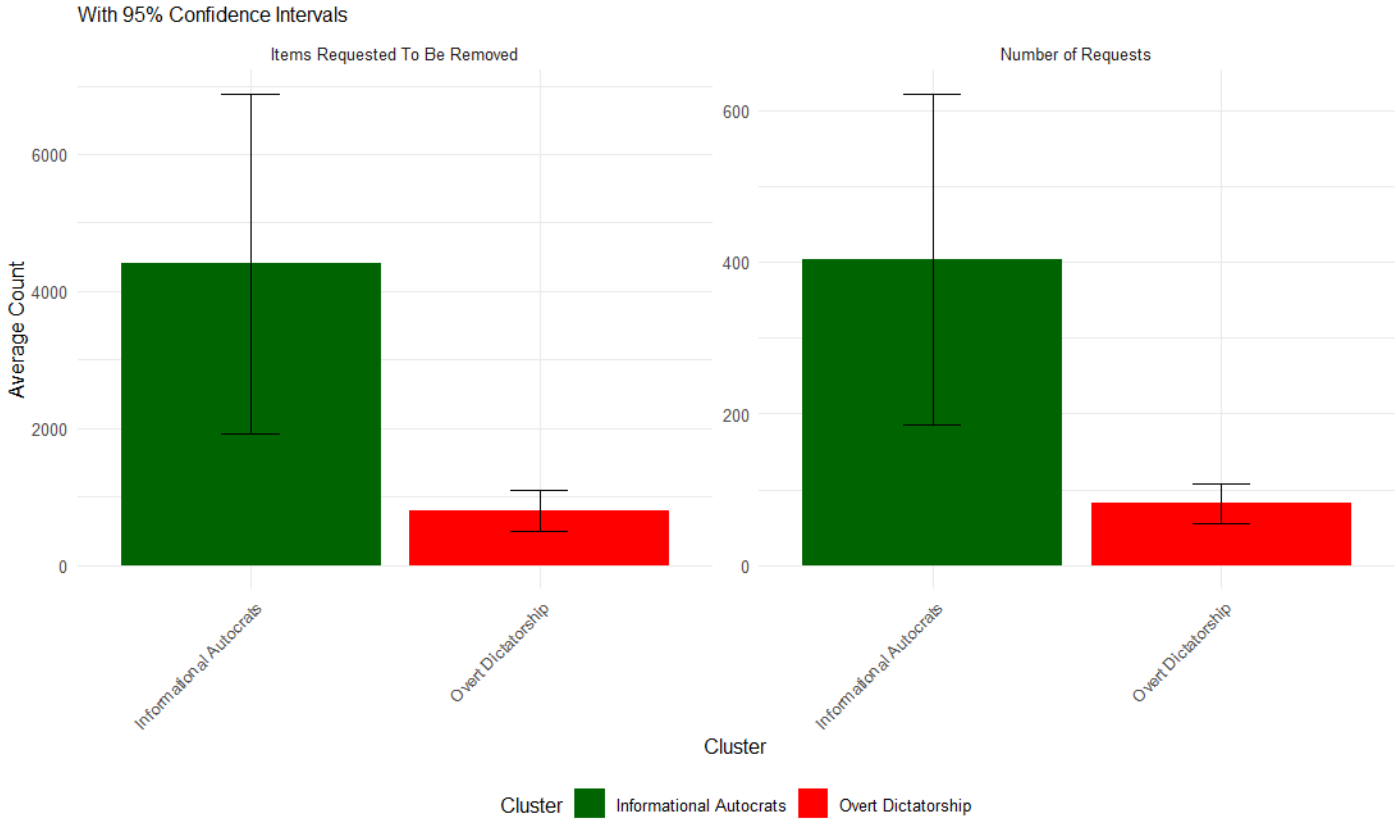


Figure 10: Comparison of Google Takedowns

Results in this Section provide a nuanced understanding of the Internet control strategies employed by the two types of censorship. The OD countries show their well-documented centralized control over online information. They do not use tools such as reporting to Internet service providers to filter the Internet in their countries largely be-

cause they have a more comprehensive, state-run mechanism in place. The IA countries are also successful in getting content removed, but in a very different way. The IA group has a strategic preference for utilizing formal channels of content regulation, aiming to shape the online narrative discreetly. This can be seen in their consistent engagement with formal takedown requests and state-sponsored anonymized attacks. Compared to the OD group, this approach is more subtle and indirect. The IA's approach also heavily relies on the existing infrastructure of digital platforms, which, to some extent, constrains the government's ability to remove content at will. There is a trade-off between the two censorship strategies: the OD approach is more effective but also more costly, while the IA approach might be less effective but lowers the cost.

5.4 Consequences of the Censorship Strategy

Why do the OD countries build their fitting system while IA countries cannot? We conjecture that IT capacity is the key explanatory variable. Building a state-run system requires substantial investment and skilled labor. Compared to China's $6.6 billion investment in controlling its billion Internet users, Russia and Turkey do not have the financial capacity to train an enormous monitoring workforce like China. Therefore, the IA countries employ a less resource-intensive but more strategic approach. They leverage the existing infrastructure built by private actors, most commonly Google and other tech companies, to remove massive amount of content with relatively less labor work. This approach might also improve politicians' reputations.

Figure 11 presents a comparative analysis of OpenNet Initiative (ONI) scores across different types of content for two distinct clusters of countries: Informational Autocracies (IA) and Overt Dictatorships (OD). The ONI scores, which range from 0 to 2, provide a quantitative measure of internet censorship intensity, with higher scores indicating more severe censorship.

The figure illustrates four categories of ONI scores: Conflict/Security, Political, Social, and Tools. For each category, the average scores for both IA (represented by green bars) and OD (represented by red bars) clusters are displayed side by side, allowing for

direct comparison. A striking pattern emerges across all four categories: the OD cluster consistently exhibits higher average ONI scores compared to the IA cluster. This pattern suggests a fundamental difference in the approach to internet control between these two regime types.

The ”Political” category shows the highest scores for both clusters, with ODs scoring approximately 1.75 and IAs scoring about 1.0. This indicates that political content is the most heavily censored across both regime types, but ODs engage in significantly more intense censorship of political information. The ”Social” category follows closely behind ”Political” in terms of censorship intensity. Again, ODs show markedly higher censorship levels (score 1.7) compared to IAs (score 0.9). The ”Tools” category, which refers to censorship of circumvention tools and technologies, shows moderate levels of censorship. ODs maintain a higher level of control (score 1.6) compared to IAs (score 0.5), indicating a more concerted effort by ODs to restrict access to tools that could bypass censorship. Interestingly, the ”Conflict/Security” category shows the lowest scores for both clusters, suggesting it’s the least censored type of content. However, the pattern of ODs (score 1.1) implementing more stringent controls than IAs (score 0.5) persists.

In summary, these figures paint a picture of two distinct approaches to information control. OD regimes employ a more heavy-handed, overtly restrictive approach across various platforms and content types. In contrast, IA regimes, while still engaging in significant levels of control, appear to be more selective and less overt in their censorship tactics. This aligns with the theoretical understanding of IA regimes as employing more sophisticated, less visible methods of maintaining information control while preserving a facade of openness.

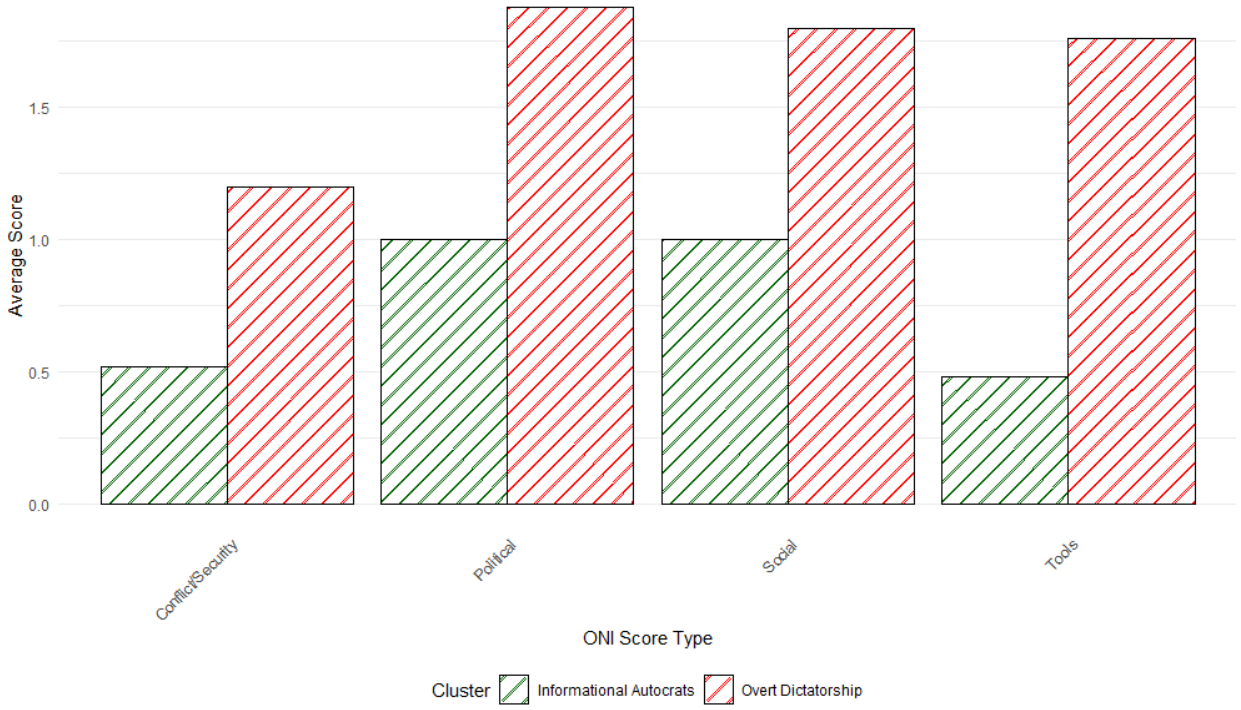


Figure 11: ONI Scores by Clusters across Different Types of Content

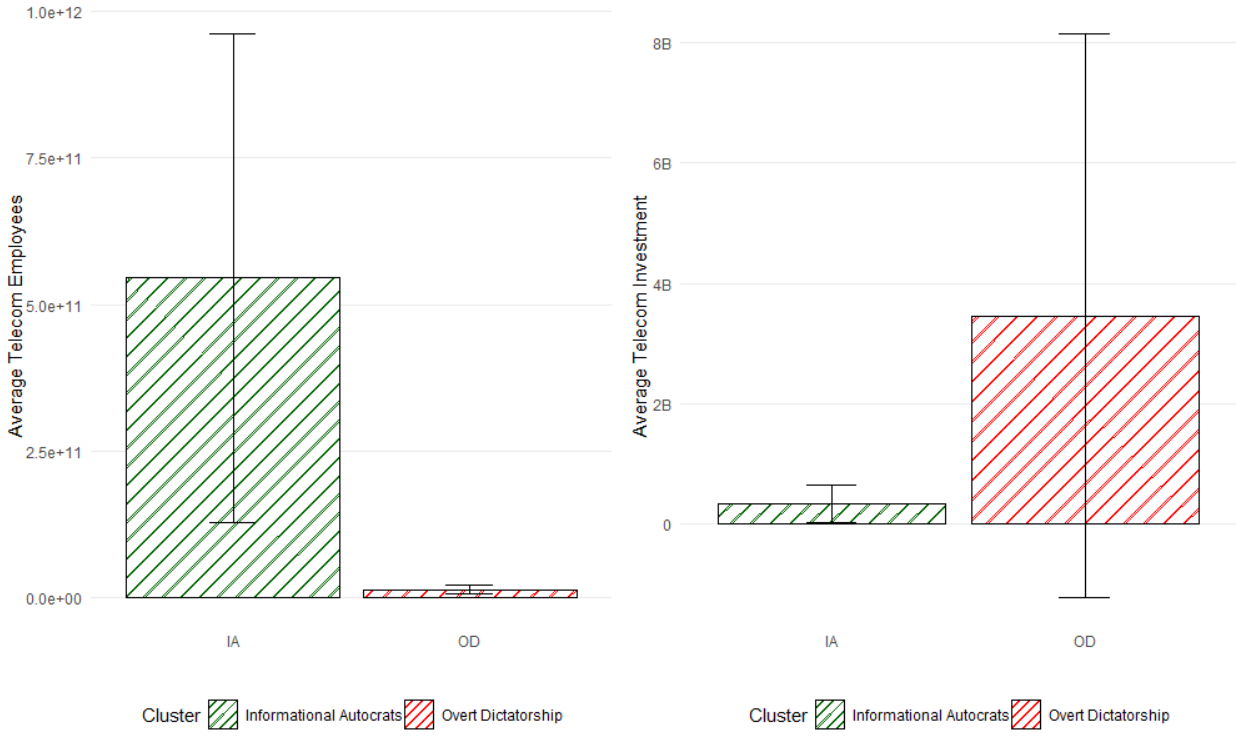


Figure 12: ITU Variables

6 Reputation and Collateral Censorship

IA : *ElectionTiming* 1 empirical strategy 2 empirical result 3 robustness: 1 how we measure election incentive, 2 court order, 3 democracy IA's fascination with the reputation game is also shown in influence campaign. ESOC graph here.

6.1 Empirical Strategy

This section outlines our empirical strategy, econometric specifications, and the identification assumptions required for consistent and unbiased estimates. Our primary goal is to identify the causal effects of electoral cycles on government content removal requests and court orders.

Following Rao (2021), we employ the share of term left as our main independent variable, measuring the proportion of a leader's current term that remains at any given time point. This measure normalizes electoral cycles across countries with different term lengths to a common scale ranging from 0 (end of term) to 1 (beginning of term). While we also examine time until next election as a robustness check, our primary specification using share of term left better captures the increasing pressures leaders face as they progress through their terms, regardless of varying institutional contexts or term durations.

A naive correlation between electoral timing and content removal requests would be confounded by several endogeneity concerns, including reverse causality (e.g., political instability affecting both content moderation and election timing) and omitted variable bias (e.g., unobserved institutional factors affecting both electoral cycles and censorship patterns). To obtain estimates with credible causal interpretation, we leverage the variation in electoral cycles across countries and over time.

Our identification relies on the assumption that, conditional on country characteristics, the timing within a leader's term provides quasi-random variation in censorship incentives. The key identifying assumption is that, absent electoral pressures, content removal requests would follow parallel trends across different phases of leaders' terms.

Our baseline specification is a two-way fixed effects (TWFE) model:

$$\text{Requests}_{it} = \alpha_i + \eta_t + \beta \cdot \text{TermLeft}_{it} + X_{it}\delta + \varepsilon_{it} \tag{1}$$

where Requests_{it} is the number of government content removal requests for country i in time t, α_i represents country fixed effects, and η_t captures time fixed effects. TermLeft_{it} measures the share of term remaining for the leader, our key variable of interest. X_{it} is a vector of time-varying country-level controls including GDP per capita, internet usage, and urban population share. Standard errors are clustered at the country level.

To test the validity of our identification strategy, we conduct a placebo test using court orders instead of government requests. Unlike government requests, which are directly influenced by political incentives, court orders primarily stem from private litigation and should be less susceptible to electoral pressures. The specification for court orders follows the same structure:

$$\text{CourtOrders}_{it} = \alpha_i + \eta_t + \beta \cdot \text{TermLeft}_{it} + X_{it}\delta + \varepsilon_{it} \tag{2}$$

Our results in Table 2 show that as leaders approach the end of their terms (i.e., lower share of term left), government content removal requests increase significantly. In contrast, Table 5 demonstrates that court orders do not show a consistent pattern with electoral cycles, supporting our hypothesis about the political nature of government requests.

6.2 Election Timing

This section presents our empirical findings on the relationship between electoral cycles, regime types, and content removal requests. We examine how the timing within a leader's term and the type of autocratic regime influence the number of content removal requests made to Google.

Tables 2 and 5 report our main results examining the relationship between electoral cycles and content removal patterns.

Table 2 presents estimates of the effect of electoral timing on government content removal requests. The baseline model in column (1) shows that the effect of share of term left is significantly negative. A one percentage point decrease in the share of term left is associated with approximately 961 additional content removal requests. This relationship remains robust and becomes stronger with the inclusion of country-level controls in column (2), where the effect increases to 1,109 additional requests. The most comprehensive specifications in columns (3) and (4), which include country and year fixed effects, show even larger magnitudes: a one percentage point decrease in share of term left is associated with between 1,299 and 1,450 additional requests. To put these estimates in context, moving from the beginning of a leader's term (share of term left = 1) to the end (share of term left = 0) would predict an increase of approximately 1,300 requests, representing a substantial increase in censorship activity as elections approach.

Table 5 serves as a placebo test, examining the relationship between electoral timing and court-ordered content removals. In contrast to government requests, court orders show a markedly different pattern. While the baseline specifications in columns (1) and (2) show small negative coefficients (-0.780 and -0.919 respectively), these effects become statistically insignificant once we include country and year fixed effects in columns (3) and (4). The magnitude of these coefficients is also substantially smaller than those for government requests. This pattern supports our hypothesis that electoral cycles primarily influence government-initiated censorship rather than broader content moderation patterns.

Both tables demonstrate the importance of controlling for economic and technological development. GDP per capita shows a consistently negative relationship with both types of removal requests, becoming statistically significant in the fixed effects specifications. This suggests that wealthier countries may have alternative mechanisms for content control or different approaches to online content moderation.

These results tell a coherent story about the relationship between electoral cycles and online censorship. Government requests for content removal show a clear electoral cycle,

with requests increasing significantly as leaders approach the end of their terms. The absence of such patterns in court orders suggests this is indeed a political phenomenon rather than a general trend in content moderation. These findings are consistent with our theoretical framework suggesting that incumbents strategically increase censorship efforts as elections approach.

	Number of Requests			
	(1)	(2)	(3)	(4)
Share of Term Left	-960.898*	-1108.967**	-1450.180**	-1298.569*
	(500.506)	(544.135)	(656.639)	(666.336)
GDP per capita (log)		-107.006	-1886.410*	-1999.058*
		(285.971)	(1087.791)	(1189.288)
Internet Users (% of pop.)		10.521	8.924	-5.402
		(9.850)	(19.767)	(33.485)
Urban Population (% of total)		2.792	46.657	-4.970
		(13.608)	(175.737)	(184.348)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	275	257	257	257
R^2	0.013	0.023	0.040	0.036

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Number of Requests with Share of Term Left

The IA group's influence campaign efforts not only target domestic citizens but also aim to swing public opinions at the international level. Figure 13 presents a network diagram that maps the connections between countries for state-sponsored digital influence activities. The node with the highest centrality in this network is Russia, a prominent case of the IA group. Russia has a history of being a primary attacker from which influence operations emanate. Russia's expansive network indicates a strategy of widespread targeting, including Ukraine, Australia, Italy, Canada, etc. On the other hand, China, as a prominent example of the OD group, is far less connected in the graph, linking only to a handful targets including Taiwan and the US. China's more focused network reflects the OD group's little interest in influence campaigns.

Figure 14 further categorizes the political intentions behind the influence efforts.

	Number of Requests			
	(1)	(2)	(3)	(4)
Time Until Next Election	-4.671	-6.250	-140.693	-150.905*
	(4.604)	(5.005)	(86.882)	(88.229)
GDP per capita (log)		-103.399	-1791.546	-1857.634
		(287.497)	(1094.349)	(1192.905)
Internet Users (% of pop.)		11.818	9.734	-6.267
		(9.888)	(19.894)	(33.670)
Urban Population (% of total)		1.283	9.490	-42.163
		(13.652)	(176.492)	(184.768)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	275	257	257	257
R^2	0.004	0.013	0.030	0.032

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Government Requests with Time Until Next Election

	Number of Court Orders			
	(1)	(2)	(3)	(4)
Share of Term Left	-42.337	-46.361	-54.379*	-58.948*
	(50.981)	(54.880)	(30.611)	(31.306)
GDP per capita (log)		-31.897	-147.768***	-161.211***
		(29.724)	(50.582)	(57.199)
Internet Users (% of pop.)		1.329	1.977*	1.333
		(1.068)	(1.003)	(1.766)
Urban Population (% of total)		1.160	11.154	9.344
		(1.426)	(8.694)	(9.363)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	264	247	247	247
R^2	0.003	0.017	0.111	0.067

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Court Orders with Share of Term Left

	Number of Court Orders			
	(1)	(2)	(3)	(4)
Time Until Next Election	-0.780	-0.919*	-5.007	-5.241
	(0.491)	(0.535)	(4.277)	(4.381)
GDP per capita (log)		-28.694	-146.094***	-160.289***
		(29.665)	(50.802)	(57.552)
Internet Users (% of pop.)		1.259	2.143**	1.547
		(1.061)	(1.001)	(1.771)
Urban Population (% of total)		1.288	9.772	8.279
		(1.420)	(8.711)	(9.421)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	264	247	247	247
R^2	0.010	0.026	0.102	0.056

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Court Orders with Time Until Next Election

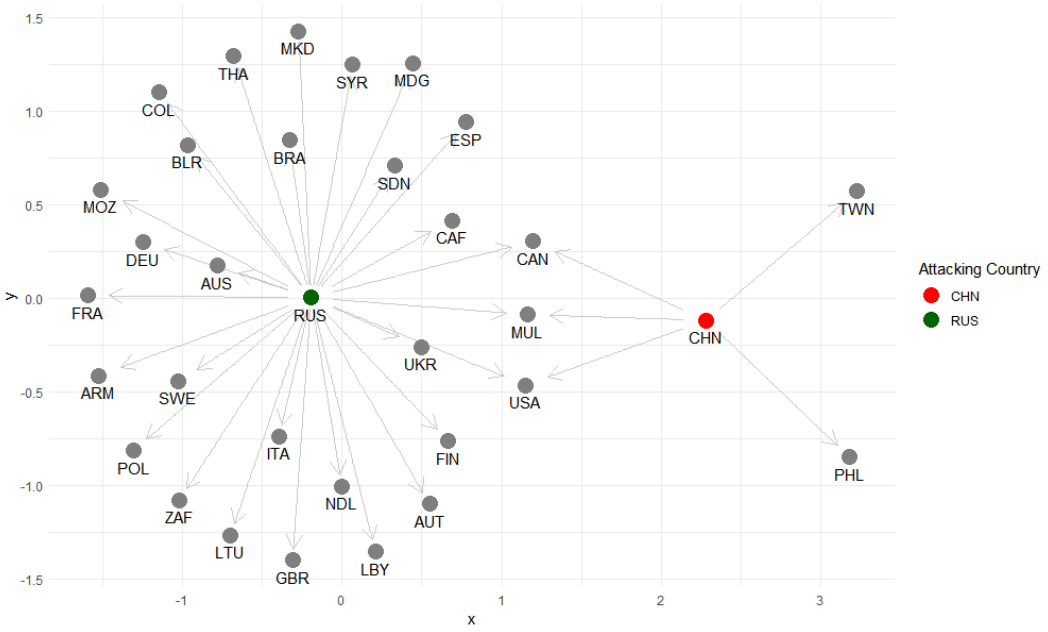


Figure 13: Network Diagram of ESOC Influence Data

This may include discrediting entities, spreading misinformation, supporting specific political entities in foreign elections, influencing policy decisions in areas like Syria or Ukraine, eroding trust in political systems, or shaping significant decisions like Brexit.

For all kinds of objectives, the IA group notably launch more influence campaigns than the PC group. The political goals of IA countries are diverse and spread out, showing up in every category. The influence activities of the OD group are mostly concentrated on discrediting an adversary or supporting an ally. Even in the categories of discredit or support, the number of influence campaigns by the IA countries is more than double that of the OD countries.

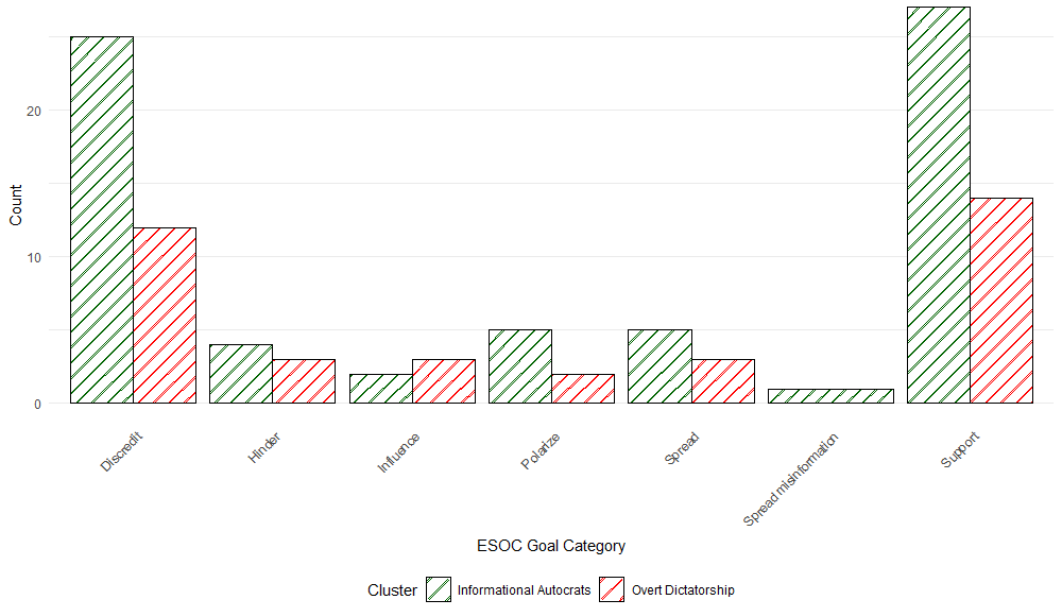


Figure 14: Influence Campaigns by Clusters across Different Political Goals

- direct censorship vs collateral censorship (see collateral censorship in my BTLJ paper)
- What are the governments complaining about? (gov requests data in Lumen)
- four consequence: international reputation and foreign operation, cover less types of content and lower capacity

7 Conclusion

On the policy front, our findings caution against overlooking media censorship detection. The critics of censorship practices often focus on regimes with overt control (PC)

and overlook those employing more subtle influence operations (IO). This distinction is vital as an increasing number of countries adopt these less visible forms of censorship. NGOs, multinational companies, and technology platforms must be vigilant in recognizing and responding to these IO strategies. The allure of IO for autocratic leaders lies in its low-cost efficiency and ability to exploit existing digital platforms to their advantage. Understanding the nuances between PC and IO approaches is essential for effectively addressing the challenges of digital censorship and advocating for freedom of expression across all regimes.

We conclude our paper by discussing a few limitations and promising extensions. One notable data limitation is our inability to capture all forms of digital influence, such as pro-government propaganda or the strategic use of trolls. This limits our understanding of the full spectrum of censorship and control tactics. Consequently, our findings should be generalized with caution, as they are most applicable in settings where content removal is the focus of Internet control. Content creation by the government that is not the focus of our analysis may play a significant role in other contexts. Future research can explore these under-studied aspects to provide a more comprehensive view of censorship strategies globally.

Looking ahead, the study of digital governance around the world presents fertile ground for future scholarly inquiry. The rich text data of takedown notices offers a promising avenue for text analysis techniques to dissect the objectives and effectiveness of influence campaigns. This paper contributes to a broader agenda aimed at understanding the cross-country differences in Internet control and the political economy behind it. Our related working paper, for example, seeks to explain why democratic countries remove an equal amount of content as their authoritarian counterparts.

References

Akdeniz, Yaman and Kerem Altiparmak, *Internet: restricted access: a critical assessment of Internet content regulation and censorship in Turkey*, Imaj Kitabevi & Imaj Yayinevi, 2008.

Akgül, Mustafa and Melih Kırlıdoğ, "Internet censorship in Turkey," *Internet Policy Review*, 2015, *4* (2), 1–22.

Ananyev, Maxim, Dimitrios Xefteris, Galina Zudenkova, and Maria Petrova, "Information and communication technologies, protests, and censorship," *Protests, and Censorship (August 20, 2019)*, 2019.

Badawy, Adam, Emilio Ferrara, and Kristina Lerman, "Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign," in "2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)" IEEE 2018, pp. 258–265.

Barquin, Sonia, Guillaume de Gantès, HV Vinayak, and Duhita Shrikhande, "Digital banking in Indonesia: Building loyalty and generating growth," *McKinsey & Company, February*, 2019, *6*.

Beazer, Quintin H, Charles D Crabtree, Christopher J Fariss, and Holger L Kern, "When do private actors engage in censorship? Evidence from a correspondence experiment with Russian private media firms," *British Journal of Political Science*, 2022, *52* (4), 1790–1809.

Besley, Timothy, *Principled agents?: The political economy of good government*, Oxford University Press on Demand, 2006.

Breindl, Yana and Bjoern Kuellmer, "Internet content regulation in France and Germany: Regulatory paths, actor constellations, and policies," *Journal of Information Technology & Politics*, 2013, *10* (4), 369–388.

Budnitsky, Stanislav and Lianrui Jia, "Branding Internet sovereignty: Digital media and the Chinese–Russian cyberalliance," *European Journal of Cultural Studies*, 2018, *21* (5), 594–613.

Bunn, Matthew, "Reimagining repression: New censorship theory and after," *History and Theory*, 2015, *54* (1), 25–44.

Chang, Chun-Chih and Thung-Hong Lin, "Autocracy login: internet censorship and civil society in the digital age," *Democratization*, 2020, *27* (5), 874–895.

Chen, Yuyu and David Y Yang, "The impact of media censorship: 1984 or brave new world?," *American Economic Review*, 2019, *109* (6), 2294–2332.

Chung, Jongpil, "Comparing online activities in China and South Korea: The internet and the political regime," *Asian Survey*, 2008, *48* (5), 727–751.

Dick, Archie L, Lilian I Oyieke, and Theo JD Bothma, "Are established democracies less vulnerable to Internet censorship than authoritarian regimes? The social media test," *FAIFE spotlight*, 2012.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin, "Why resource-poor dictators allow freer media: A theory and evidence from panel data," *American political science Review*, 2009, *103* (4), 645–668.

Emami, Karim, "Is it Necessary for Iran to Increase the Share of ICT Sector in GDP?," *Economics Research*, 2018, *18* (68), 45–74.

Esarey, Ashley and Qiang Xiao, "Digital communication and political change in China," *International Journal of Communication*, 2011, *5*, 22.

Faris, Robert and Nart Villeneuve, "Measuring global Internet filtering," *Access denied: The practice and policy of global Internet filtering*, 2008, *5.*

Fedasiuk, Ryan, "Buying silence: The price of Internet censorship in China," *China Brief*, 2021, *21* (1), 18–25.

Fish, Eric, "Is Internet censorship compatible with democracy? Legal restrictions of online speech in South Korea," *Asia-Pacific Journal on Human Rights and the Law*, 2009, *10* (2), 43–96.

Frydman, Benoît and Isabelle Rorive, "Regulating Internet content through intermediaries in Europe and the USA," *Zeitschrift für Rechtssoziologie*, 2002, *23* (1), 41–60.

Gehlbach, Scott and Konstantin Sonin, "Government control of the media," *Journal of public Economics*, 2014, *118*, 163–171.

Goldsmith, Jack, "Who controls the Internet? Illusions of a borderless world," *Strategic Direction*, 2007, *23* (11).

Guriev, Sergei and Daniel Treisman, "Informational autocrats," *Journal of economic perspectives*, 2019, *33* (4), 100–127.

_ and _, "The popularity of authoritarian leaders: A cross-national investigation," *World Politics*, 2020, *72* (4), 601–638.

_ and _, "A theory of informational autocracy," *Journal of public economics*, 2020, *186*, 104158.

_ and _, *Spin dictators: The changing face of tyranny in the 21st century*, Princeton University Press, 2022.

Hellmeier, Sebastian, "The dictator's digital toolkit: Explaining variation in internet filtering in authoritarian regimes," *Politics & Policy*, 2016, *44* (6), 1158–1191.

Hobbs, William R and Margaret E Roberts, "How sudden censorship can increase access to information," *American Political Science Review*, 2018, *112* (3), 621–636.

King, Gary, Benjamin Schneer, and Ariel White, "How the news media activate public expression and influence national agendas," *Science*, 2017, *358* (6364), 776–780.

_, Jennifer Pan, and Margaret E Roberts, "How censorship in China allows government criticism but silences collective expression," *American political science Review*, 2013, *107* (2), 326–343.

_, _, and _, "Reverse-engineering censorship in China: Randomized experimentation and participant observation," *Science*, 2014, *345* (6199), 1251722.

Kolozaridi, Polina and Dmitry Muravyov, "Contextualizing sovereignty: A critical review of competing explanations of the Internet governance in the (so-called) Russian case," *First Monday*, 2021.

Land, Molly K, "Against privatized censorship: Proposals for responsible delegation," *Va. J. Int'l L.*, 2019, *60*, 363.

Lorentzen, Peter, "China's strategic censorship," *American Journal of political science*, 2014, *58* (2), 402–414.

Mchangama, Jacob and Joelle Fiss, "The digital Berlin Wall: How Germany (accidentally) created a prototype for global online censorship," *Copenhagen: Justitia and Authors*, 2019.

Mueller, Milton L, *Networks and states: The global politics of Internet governance*, MIT press, 2010.

Pearce, Paul, Roya Ensafi, Frank Li, Nick Feamster, and Vern Paxson, "Augur: Internet-wide detection of connectivity disruptions," in "2017 IEEE Symposium on Security and Privacy (SP)" IEEE 2017, pp. 427–443.

Petrov, Nikolay, Maria Lipman, and Henry E Hale, "Three dilemmas of hybrid regime governance: Russia from Putin to Putin," *Post-Soviet Affairs*, 2014, *30* (1), 1–26.

Prat, Andrea and David Strömberg, "The political economy of mass media," *Advances in economics and econometrics*, 2013, *2*, 135.

Qiang, Xiao, "Liberation technology: the battle for the Chinese internet," *Journal of Democracy*, 2011, *22* (2), 47–61.

Rao, Weijia, "Domestic Politics and Settlement in Investor-State Arbitration," *The Journal of Legal Studies*, 2021, *50* (1), 145–185.

Roberts, Margaret, *Censored: distraction and diversion inside China's Great Firewall*, Princeton University Press, 2018.

Roberts, Margaret E., *Censored: Distraction and Diversion Inside China's Great Firewall*, Princeton University Press, 2018.

Schedler, Andreas, "Authoritarianism's last line of defense," *J. Democracy*, 2010, *21*, 69.

Shadmehr, Mehdi and Dan Bernhardt, "State censorship," *American Economic Journal: Microeconomics*, 2015, *7* (2), 280–307.

Shen, Xiaoxiao and Rory Truex, "In search of self-censorship," *British Journal of Political Science*, 2021, *51* (4), 1672–1684.

Simonov, Andrey and Justin Rao, "Demand for online news under government control: Evidence from russia," *Journal of Political Economy*, 2022, *130* (2), 259–309.

Sinpeng, Aim, "Digital media, political authoritarianism, and Internet controls in Southeast Asia," *Media, Culture & Society*, 2020, *42* (1), 25–39.

Sivetc, Liudmila, "Controlling free expression "by infrastructure" in the Russian Internet: The consequences of RuNet sovereignization," *First Monday*, 2021.

Stier, Sebastian, "Democracy, autocracy and the news: the impact of regime type on media freedom," *Democratization*, 2015, *22* (7), 1273–1295.

Tréguer, Félix, "From deep state illegality to law of the land: The case of internet surveillance in France," in "7th Biennial Surveillance & Society Conference (SSN 2016):" Power, performance and trust"" 2016.

Urman, Aleksandra and Mykola Makhortykh, "How transparent are transparency reports? Comparative analysis of transparency reporting across online platforms," *Telecommunications Policy*, 2023, p. 102477.

Warf, Barney, "Geographies of global Internet censorship," *GeoJournal*, 2011, *76*, 1–23.

Williams, Andrew, "A global index of information transparency and accountability," *Journal of Comparative Economics*, 2015, *43* (3), 804–824.

Wulf, Volker, Dave Randall, Konstantin Aal, and Markus Rohde, "The Personal is the Political: Internet Filtering and Counter Appropriation in the Islamic Republic of Iran," *Computer Supported Cooperative Work (CSCW)*, 2022, *31* (2), 373–409.

Yesil, Bilge and Efe Kerem Sozeri, "Online surveillance in Turkey: Legislation, technology and citizen involvement," *Surveillance & Society*, 2017, *15* (3/4), 543–549.

Zarras, Apostolis, "Leveraging Internet services to evade censorship," in "International Conference on Information Security" Springer 2016, pp. 253–270.

Zittrain, Jonathan L, Robert Faris, Helmi Noman, Justin Clark, Casey Tilton, and Ryan Morrison-Westphal, "The shifting landscape of global internet censorship," *Berkman Klein Center Research Publication*, 2017, (2017-4), 17–38.

A Appendix. Robustness of Clustering Results

Country Name	Classification in Guriev and Treisman (2019)	Classification in Guriev and Treisman (2020b)	Classification in Guriev and Treisman (2020a)	This paper
Venezuela	IA	IA	IA	IA
Russia	IA	IA	IA	IA
Peru	IA	–	–	IA
Malaysia	IA	–	–	IA
Hungary	IA	–	–	IA
Singapore	IA	IA	IA	IA
Ecuador	IA	IA	IA	IA
Armenia	–	IA	IA	IA
Fiji	–	IA	–	IA
Belarus	–	IA	IA	IA
Guinea	–	IA	IA	IA
Kazakhstan	–	IA	IA	IA
Algeria	–	IA	–	IA
Bahrain	–	IA	–	IA
Cuba	–	IA	–	IA
Jordan	–	IA	–	IA
China	OD	–	–	OD
Angola	–	OD	OD	OD
Central African Republic	–	OD	OD	OD
Chad	–	OD	OD	OD
Congo Brazzaville	–	OD	OD	OD
Congo Kinshasa	–	OD	OD	OD
Togo	–	OD	OD	OD
Uganda	–	OD	OD	OD
Cambodia	–	OD	OD	OD
Cameroon	–	OD	OD	OD
Bangladesh	–	OD	OD	OD
Nigeria	–	OD	OD	OD
Rwanda	–	–	OD	OD
Sudan	–	–	OD	OD
Syria	–	–	OD	OD

Table 6: Country Classification: IA vs OD

B Appendix. Sub-indices of Internet Control

C Appendix. ESOC Network Graph by Clusters

D Appendix. Additional Regression Results on IT Capacity

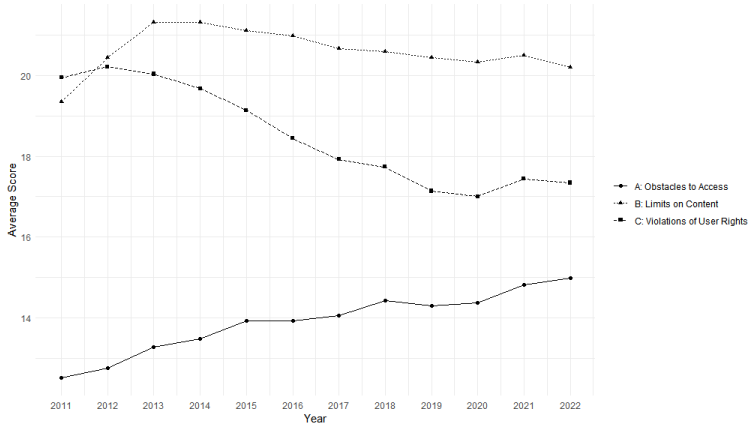
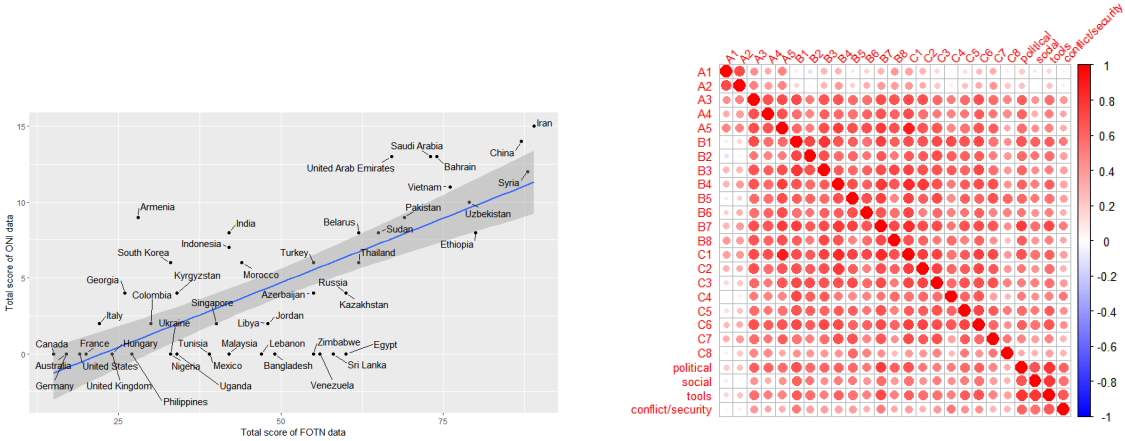


Figure 15: FOTN Subscores across time



(a) Scatter Plot

(b) Correlation Matrix

Figure 16: Relationship between ONI & FOTN data

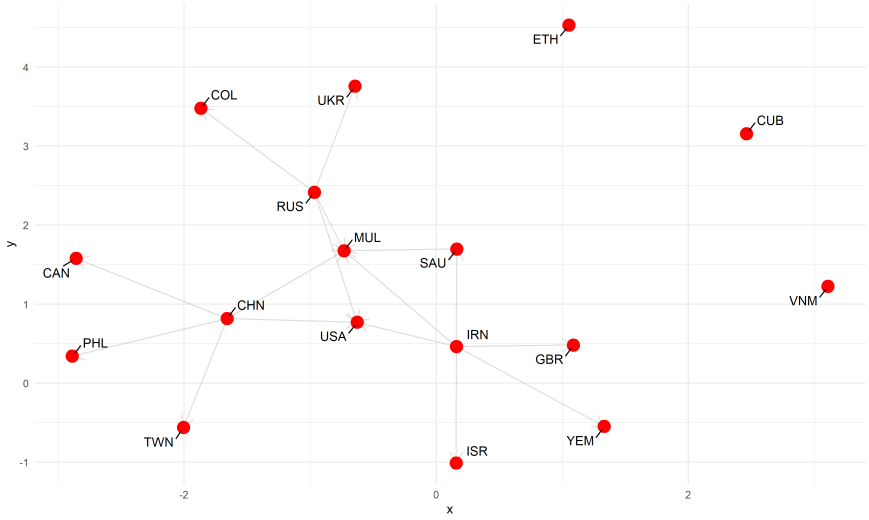


Figure 17: Network Diagram of OD

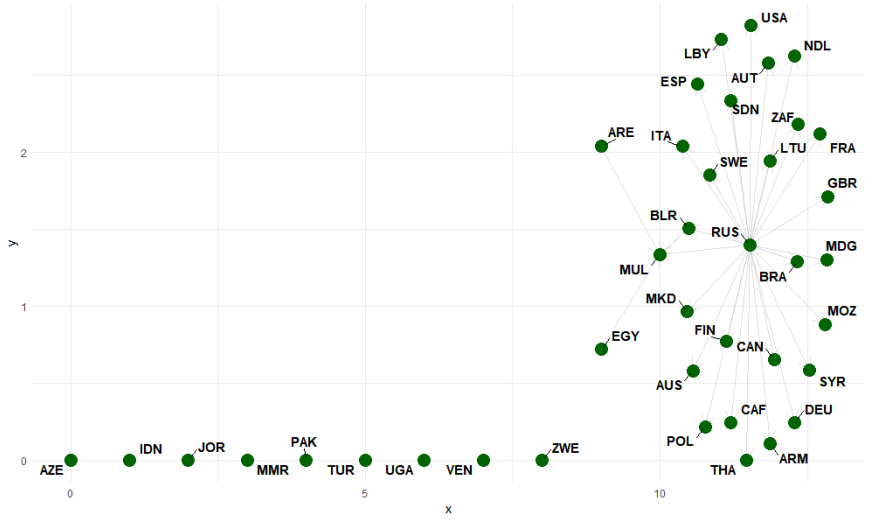


Figure 18: Network Diagram of IA

	(1) LPM	(2) LPM	(3) LPM	(4) LPM	(5) Logit	(6) Probit
Government Regulation Capacity	-0.09617***	-0.1323***	0.03735	0.03113	1.076	0.4491
	(0.02119)	(0.02618)	(0.06363)	(0.06529)	(2.101)	(1.219)
GDP per capita (log)		0.07331*	0.001498	0.01921	0.03108	0.04169
		(0.03601)	(0.05643)	(0.05806)	(1.838)	(0.9665)
Internet users (% of pop.)		0.00054	-0.003026*	-0.000374	-0.02808	-0.01013
		(0.00133)	(0.001336)	(0.002065)	(0.05266)	(0.02977)
Urban population (% of total)		-0.001205	0.01050	0.01874	0.08017	-0.07613
		(0.00164)	(0.01244)	(0.01348)	(0.8853)	(0.4564)
Country FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Observations	382	358	319	309	309	309
Adjusted R^2	0.0513	0.0718	0.776	0.7821		

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Effect of Government Regulation Capacity on the Probability of IO censorship

	(1)	(2)	(3)	(4)
Government Internet Shutdown Capacity	-998.679**	-1,290.818***	2,241.405*	528.361
	(417.718)	(462.660)	(1,274.324)	(1,367.870)
GDP per capita (log)		-598.821	-3,644.115*	-2,962.452
		(621.722)	(2,011.610)	(2,209.483)
Internet users (% of pop.)		42.155**	49.505	-76.559
		(21.247)	(38.868)	(53.098)
Urban population (% of total)		10.876	-59.853	-560.643
		(28.277)	(314.033)	(357.562)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	207	187	187	187
Adjusted R^2	0.022	0.043	0.470	0.488

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Effect of Government Internet Shutdown Capacity on the Number of Requests

	(1)	(2)	(3)	(4)
Government Cyber Security Capacity	1289.349***	1752.012***	-83.272	-1778.756
	(276.269)	(373.195)	(1631.262)	(1716.687)
GDP per capita (log)		-1338.916**	-3523.970*	-2598.155
		(628.878)	(2030.707)	(2198.263)
Internet users (% of pop.)		28.172	77.061**	-71.887
		(20.459)	(37.058)	(52.881)
Urban population (% of total)		1.721	-106.299	-652.442*
		(27.120)	(316.071)	(349.577)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	207	187	187	187
Adjusted R^2	0.092	0.110	0.459	0.491

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Effect of Government Cyber Security Capacity on the Number of Requests

	(1)	(2)	(3)	(4)
Government Internet Filtering Capacity	286.354	-438.869	286.354	-2,429.591
	(872.100)	(1,348.542)	(872.100)	(1,754.508)
Filtering Capacity Above Mean	2,120.513	1,748.914	2,120.513	144.240
	(2,028.506)	(2,334.326)	(2,028.506)	(2,346.130)
Filtering Capacity * Above Mean	-741.775	-193.289	-741.775	2,344.607
	(1,327.252)	(1,697.739)	(1,327.252)	(1,873.898)
GDP per capita (log)		-490.205		-2,388.068
		(636.444)		(2,198.765)
Internet users (% of pop.)		32.442		-104.181*
		(22.486)		(53.848)
Urban population (% of total)		6.623		-660.645*
		(29.356)		(347.021)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	207	187	207	187
Adjusted R^2	-0.0001	0.001	-0.0001	0.504

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Effect of Government Internet Filtering Capacity on the Number of Requests

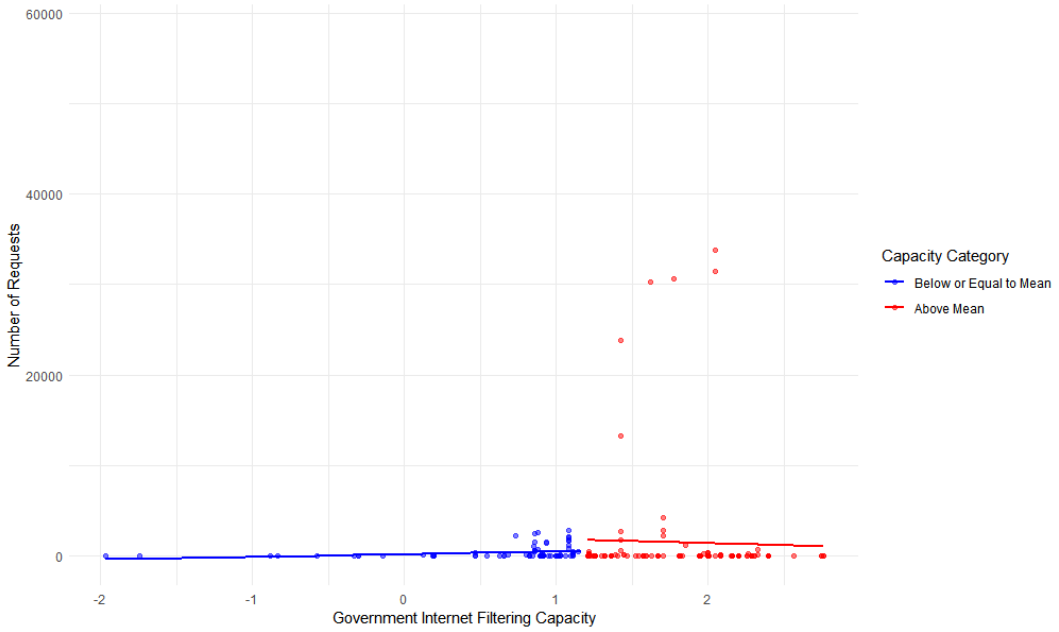


Figure 19: Piecewise Linear Relationship with Interaction Term

	(1)	(2)	(3)	(4)
Government Cyber Security Capacity	30.476	166.795	30.476	-1,200.582
	(980.165)	(1,298.553)	(980.165)	(3,462.508)
Cyber Security Capacity Above Mean	-1,825.220*	-3,092.135***	-1,825.220*	336.606
	(967.053)	(1,117.254)	(967.053)	(2,549.207)
Cyber Security Capacity * Above Mean	2,143.878**	3,001.776**	2,143.878**	-848.933
	(1,077.998)	(1,384.777)	(1,077.998)	(4,046.631)
GDP per capita (log)		-1,682.536***		-2,626.932
		(623.375)		(2,244.064)
Internet users (% of pop.)		39.953*		-73.259
		(20.274)		(55.397)
Urban population (% of total)		0.852		-671.777*
		(26.974)		(363.882)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	207	187	207	187
Adjusted R2	0.113	0.155	0.113	0.484

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Effect of Government Cyber Security Capacity on the Number of Requests

	(1)	(2)	(3)	(4)
Government Internet Shutdown Capacity	692.382	-2,686.258	-2,837.456	-3,136.705
	(5,497.859)	(5,355.164)	(15,504.180)	(15,433.130)
GDP per capita (log)		3,046.936	1,090.074	-5,052.389
		(7,106.099)	(19,722.140)	(20,315.110)
Internet users (% of pop.)		-72.040	-65.341	-702.224
		(311.112)	(720.483)	(916.708)
Urban population (% of total)		69.725	3,928.282	-397.516
		(323.523)	(6,976.025)	(8,145.733)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	215	196	196	196
Adjusted R2	-0.005	-0.017	0.248	0.265

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Effect of Government Internet Shutdown Capacity on Failure Counts for OONI

	(1)	(2)	(3)	(4)
Government Cyber Security Capacity	11,890.110***	20,196.340***	-22,652.830	-24,677.200
	(4,200.652)	(5,043.582)	(23,541.080)	(23,609.980)
GDP per capita (log)		-8,792.348	1,339.753	-5,184.907
		(7,392.136)	(19,642.160)	(20,231.530)
Internet users (% of pop.)		-68.104	26.008	-541.742
		(297.088)	(712.145)	(909.587)
Urban population (% of total)		-13.670	3,718.134	-140.512
		(305.297)	(6,959.569)	(8,113.056)
Country FE	No	No	Yes	Yes
Year FE	No	No	No	Yes
Observations	215	196	196	196
Adjusted R2	0.032	0.060	0.252	0.270

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Effect of Government Cyber Security Capacity on Failure Counts for OONI