digital society project

Measuring Internet Politics:
Introducing the Digital Society Project
(DSP)

Valeriya Mechkova, Daniel Pemstein,
Brigitte Seim, Steven Wilson

Digital Society Project[1]

Working Paper #1

Measuring Internet Politics:

Digital Society Project (DSP)

Annual Report v3 2021

Valeriya Mechkova (University of Gothenburg)

Daniel Pemstein (North Dakota State University)

Brigitte Seim (University of North Carolina, Chapel Hill)

Steven Wilson (Brandeis University)[2]

Table of Contents

Introduction

As of 2019, the global number of internet users has surpassed 4 billion, or more than half of the total population, with the average internet-user spending around 6.5 hours per day using devices connected to the internet (Kemp 2019). Similarly, 3.5 billion people use some type of social media platform, an increase of one billion over the past year (Ibid). These statistics are remarkable, but how has this massive shift in access to digital media affected political behavior? Has the internet and social media helped citizens to organize themselves to hold governments more accountable, reach across past previous divides, and stimulate discussions? Or is the opposite true: has the internet created stronger polarization among groups, and given ill-minded governments a new, effective, way to control us, and target other states?

In this report, we overview the DSP dataset, the product of a global survey of hundreds of country and area experts, and preview key descriptive patterns from this data collection effort. The data covers virtually all countries in the world from 2000 to 2020 and measures a set of 35 indicators of polarization and politicization of social media, misinformation campaigns and coordinated information operations, and foreign influence in and monitoring of domestic politics. We expect that the data and the research produced by this project will be of great interest to both the academic and policy communities, at a time when understanding the political and social consequences of the internet is rapidly increasing.

Motivation

The primary goal of this project is to provide high-quality, publicly available data describing the intersection between politics and social media. While there is great demand for such data, reliable measures of key indicators, with wide global and temporal coverage, are largely unavailable. We anticipate that academics will use these data to understand how people use social media as a political tool and to explore how political institutions and social media usage interact. Policymakers will use these data to, among a host of applications, understand how, and where, to intervene to curb internet-driven political violence, reduce electoral manipulation, counter foreign information operations, and enhance governmental accountability.

There is a theoretical expectation that the rise of social media should alter politics by reducing the transaction costs that factor into solving collective action problems (Castells, 2009; Shirky, 2009). The

ability to communicate is an essential component of most elements of politics, and as such, we expect that the changes wrought by the rise of universal, instantaneous, and mobile mass communication should therefore affect a myriad of political outcomes. Larry Diamond dubbed such technology "liberation technology" and the *Journal of Democracy* has examined regularly the role of technology in increasing the ability of social movements to resist regimes, in addition to examining the responses of states to this technology (Diamond, 2010).

The low barrier of entry for the collection of social media data by scholars has led to a proliferation of small-n studies of the effects of social media on various variables in global contexts. To name just a few to hint at the variety: the use of ICTs to facilitate violence in Africa (Pierskalla & Hollenbach, 2013), their use in election monitoring in Nigeria (Bailard & Livingston, 2014), social media's role in Euromaidan (Wilson, 2017), its effect on political participation in the EU (Valeriani & Vaccari, 2015), its role in organization during the 2011 London riots (Baker, 2012; Denef, Bayerl, & Kaptein, 2013), and even the crowd-sourced coordination of fish prices (Aker & Mbiti, 2010).

Research on the effect of social media on social mobilization, especially in authoritarian contexts, has been particularly extensive (Anderson, 2011; Farrell, 2012; Tucker, 2013; Tucker, Barberá, & Metzger, 2013; Tufekci & Wilson, 2012; Tufecki 2017). The bulk of this work has been focused on social media's role in the Arab Spring (i.e., the so-called "Facebook Revolutions") (Alqudsi-ghabra, Al-bannai, & Al-bahrani, 2011; Hofheinz, 2005; Mackell, 2011; Murphy, 2006; Oghia & Indelicato, 2011; Sabadello, 2012; Stepanova, 2011; Zhuo et al., 2011) and earlier Color Revolutions (Bunce & Wolchik, 2010; Chowdhury, 2008; Dyczok, 2005; Goldstein, 2007; Kyj, 2006). While much of this work highlights social media's potential for citizen mobilization in closed regimes, authoritarian states with high technical capacity—notably China—are able to allow substantial political criticism on social media while stymying collective action (King, Pan & Roberts 2013).

In addition to the work focusing on whether and how social media empowers grassroots organizations, scholars increasingly acknowledge that social media also has a dark side. Multiple authors examine how digital communication platforms affect political violence (Bak, Sriyai & Meserve, 2018; Gohdes, 2015; Warren 2015). There is growing evidence that the internet has stabilized technically capable authoritarian regimes by enhancing their capacity to monitor populations and solve the dictator's information problem (Morozov, 2012). These include China's use of social media

monitoring to learn which policies and local officials are unpopular, Russia's domestic astroturfing efforts online, and the use of social media to help the government identify regime opponents in various Arab countries (Gohdes, Forthcoming; Gunitsky, 2015; Wilson, 2016). Indeed, authoritarian regimes have developed an ever-evolving menu of strategies for policing internet content and disrupting collective action (Deibert et al., 2008; Roberts, 2018). Yet, intriguingly, some of these monitoring mechanisms are increasingly being used in non-authoritarian contexts as a way to improve governance outcomes, by increasing the ability of governments to respond directly to the concerns of populations (Moreno, 2012).

A growing literature has also explored the negative implications of the internet for democracies. Evidence suggests that social media has helped destabilize new democracies by making short term collective action easy at the expense of building institutions (Faris and Etling, 2008). Others have focused on more specific problems that arise from social media, such as the danger of homophily (the self-sorting of individuals into sheltered groups of those with similar beliefs) (Garret & Resnick 2011; Gentzkow and Shapiro, 2010; Page, 2008; Pariser, 2012; Sunstein, 2009; Wojcieszak and Mutz, 2009), or implications of a digital divide domestically (Norris, 2001; Schlozman, Verba, and Brady, 2010). Scholars in this literature have argued that, even in democracies, internet censorship is politically motivated (Meserve & Pemstein, 2018), and that legal protections for civil liberties are often ineffectively extended to the digital realm (Gillespie, 2018; Zittrain, 2003; Adler, 2011; Meserve, 2018).

And, of course, the 2016 American presidential elections point to concerns about political and electoral cyber-security, and the weaponization of social media by foreign actors to interfere in democratic processes. As with the effect of social media on mobilization, the study of regime capacity for operating in this context is confined to small, single country case studies analyzing the capabilities of particular states (Geers, 2015; Hjortdal, 2011; Krekel, 2014; Mandiant, 2013; Phahlamohlaka, et al., 2011; Robinson, et al., 2013), or use broad instruments to approximate general measures of state capacity that cannot capture specific capabilities (Tembe, et al., 2014; Wilson, 2016).

The works discussed above provide substantively varied and theoretically rich perspectives on the effects of social media on politics. However, one drawback of this work is that it is almost exclusively composed of single country case studies, and in particular, cases that select upon the dependent variable of *something interesting* having happened. Despite its importance to understanding authoritarian

persistence and democratic politics in the contemporary world, an almost total lack of cross-national comparative data persists. While scholars would benefit from such data, this need is especially acute for policy-makers and firms who increasingly need to make decisions in light of global variation in digital politics. Global variation in the state's capacity to control and monitor its population's internet usage, or the extent to which individuals use social media to politically organize, is not unknowable. But these quantities are difficult to measure cross-nationally, because such information is the domain knowledge of individuals who are experts on particular countries.

Implementation of the Digital Society Survey

While a variety of strategies exist to collect such cross-national, and over-time, data, the Varieties of Democracy Project (V-Dem) (Coppedge, et al., 2018) provides a model that has met with substantial success. In particular, by leveraging a vast network of country and domain experts, V-Dem has compiled a database measuring democratic institutions over vast swaths of time and space, that has proven useful to a diverse array of academics and policymakers. The DSP builds on the V-Dem infrastructure, redirecting the efforts of its expert network to better understand global internet politics.

The Digital Society Survey is an expert-coded survey comprising thirty-five indicators. The survey captures the politicization of social media, misinformation campaigns and coordinated information operations, and foreign influence in and monitoring of domestic politics via the Internet. Other than a handful of multiple selection and free-response style questions, all questions ask respondents to rate aspects of internet politics using Likert scales. The survey includes a full set of anchoring vignettes to help anchor scales across experts, and respondents were able to answer questions in six languages (English, French, Spanish, Portuguese, Arabic, and Russian).

Initial data collection concluded in January of 2019 for 180 countries from 2000 to 2018, and the V-Dem data team processed these data using the standard V-Dem measurement modeling and quality control processes (Coppedge et al., 2018; Pemstein, 2018). The data were updated in 2020 and 2021, and v3 of the dataset is now available. The indicators cover five sub-domains, detailed below, along with the question text for each question (though not including the full set of Likert scale choices for each, given space constraints). The full codebook with description of each variable is available at: http://digitalsocietyproject.org/data/

Coordinated Information Operations

Social media is increasingly used as a tool of coordinated information operations. These operations can be used by either foreign powers with a vested interest in the political trajectory of the country, or by domestic actors with an incentive to skew information available to the public. These actors use the reach of social media and tools such as "troll armies" to generate and disseminate particular viewpoints or fake news. This portion of the survey captures the involvement of foreign actors in domestic politics via Internet technologies, and the presence and characteristics of either foreign or domestic coordinated information operations. In addition, it captures the capacity of regimes for using such techniques both domestically and abroad.

Indicator	Question Text
Government dissemination of false information domestic	How often do the government and its agents use social media to disseminate misleading viewpoints or false information to influence its own population?
Government dissemination of false information abroad	How often do the government and its agents use social media to disseminate misleading viewpoints or false information to influence citizens of other countries abroad?
Party dissemination of false information domestic	How often do major political parties and candidates for office use social media to disseminate misleading viewpoints or false information to influence their own population?
Party dissemination of false information abroad	How often do major political parties and candidates for office use social media to disseminate misleading viewpoints or false information to influence citizens of other countries abroad?
Foreign governments dissemination of false information	How routinely do foreign governments and their agents use social media to disseminate misleading viewpoints or false information to influence domestic politics in this country?
Foreign governments ads	How routinely do foreign governments and their agents use paid advertisements on social media in order to disseminate misleading viewpoints or false information to influence domestic politics in this country?

Digital Media Freedom

These questions will mode (e.g., filtering, active takedowns, limitation of access), actor (e.g., government, non-state actors), and extent of censorship.

Indicator	Question Text
Government Internet filtering capacity	Independent of whether it actually does so in practice, does the government have the technical capacity to censor information (text, audio, images, or video) on the Internet by filtering (blocking access to certain websites) if it decided to?
Government Internet filtering in practice	How frequently does the government censor political information (text, audio, images, or video) on the Internet by filtering (blocking access to certain websites)?
Government Internet shut down capacity	Independent of whether it actually does so in practice, does the government have the technical capacity to actively shut down domestic access to the Internet if it decided to?
Government Internet shut down in practice	How often does the government shut down domestic access to the Internet?
Government social media shut down in practice	How often does the government shut down access to social media platforms?
Government social media alternatives	How prevalent is the usage of social media platforms that are wholly controlled by either the government or its agents in this country?
Government social media monitoring	How comprehensive is the surveillance of political content in social media by the government or its agents?
Government social media censorship in practice	To what degree does the government censor political content (i.e., deleting or filtering specific posts for political reasons) on social media in practice?
Government cyber security capacity	Does the government have sufficiently technologically skilled staff and resources to mitigate harm from cyber-security threats?
Political parties cyber security capacity	Do the major political parties have sufficiently technologically skilled staff and resources to mitigate harm from cyber security threats?

Online Media Polarization

This portion of the survey provides indicators of the level of polarization in discourse in both online and traditional media, probing the extent to which media environments are fractionalized, the extent to which citizens obtain political information from polarized sources, and the extent to which media markets serve particular ideological niches.

Indicator	Question Text
Online media existence	Do people consume domestic online media?
Online media perspectives	Do the major domestic online media outlets represent a wide range of political perspectives?
Online media fractionalization	Do the major domestic online media outlets give a similar presentation of major (political) news?

Social Cleavages

This portion of survey examines the extent to which social cleavages proliferate, are activated, and engender ongoing conflict within states. This exploration includes several questions specific to online, social media discourse, as well as more indirectly related measures of cleaves in society more generally.

Indicator	Question Text
Online harassment of groups	Which groups are targets of hate speech or harassment in online media? (Multiple selection of 10 groups, with free-text entry for other)
Use of social media to organize offline violence	How often do people use social media to organize offline violence?
Average people's use of social media to organize offline action	How often do average people use social media to organize offline political action of any kind?

Elites' use of social media to organize offline action	How often do domestic elites use social media to organize offline political action of any kind?
Party/candidate use of social media in campaigns	To what extent do major political parties and candidates use social media during electoral campaigns to communicate with constituents?
Arrests for political content	If a citizen posts political content online that would run counter to the government and its policies, what is the likelihood that citizen is arrested?
Types of organization through social media	What types of offline political action is most commonly mobilized on social media? (Multiple section of 9 actions, with free-text entry for other)
Polarization of society	How would you characterize the differences of opinions on major political issues in this society?
Political parties hate speech	How often do major political parties use hate speech as part of their rhetoric?

State Internet Regulation Capacity and Approach

States vary dramatically in their capacity to regulate online content. This portion of the survey examines the extent to which the state has the capacity to regulate online content, and the model that the state uses to regulate online content. In particular we ask questions about the extent to which laws allow states to remove content, privacy and data protections provided by law, the extent to which actors can leverage copyright and defamation law to force the removal of online content, and de-facto levels of state intervention in online media.

Indicator	Question Text
Internet legal regulation content	What type of content is covered in the legal framework to regulate Internet?
Privacy protection by law exists	Does a legal framework to protect Internet users' privacy and their data exist?
Privacy protection by law content	What does the legal framework to protect Internet users' privacy and their data stipulate?

Government capacity to regulate online content	Does the government have sufficient staff and resources to regulate Internet content in accordance with existing law?
Government online content regulation approach	Does the government use its own resources and institutions to monitor and regulate online content or does it distribute this regulatory burden to private actors such as Internet service providers?
Defamation protection	Does the legal framework provide protection against defamatory online content, or hate speech?
Abuse of defamation and copyright law by elites	To what extent do elites abuse the legal system (e.g., defamation and copyright law) to censor political speech online?

Data Collection

To generate the data for the DSP survey, we rely on the expertise and infrastructure of the Varieties of Democracy Project. V-Dem's data team has collected and processed the survey using the standard V-Dem measurement modeling and quality control processes (Coppedge et al., 2018; Pemstein, 2018).

Currently the V-Dem project stands as a world-leading research hub for analyzing and producing high quality data on democracy. The data base is the largest of its kind, covering virtually all countries in the world. The latest version of the data set (release 11) includes 180 countries from 1900 to 2021, and consists of more than 350 indicators on various aspects of democracy.[3]

---

[3] V-Dem infrastructure, data collection, research, collaboration and outreach is/has been funded by a collection of research foundations and international sources including the European Commission/DEVCO, the World Bank, Ministry of Foreign Affairs, Sweden, Danish International Development Agency, Canadian International Development Agency; the Research Council of Norway/NORAD, the Mo Ibrahim Foundation, the B-Team, International IDEA, The European Research Council, the Research Councils of Sweden, Norway, and Denmark, the Riksbankens Jubileumsfond, and the M&M Wallenberg and the K&A Wallenberg foundations. Co-funding has been provided by the Vice Chancellor, the Dean of Social Sciences, and Department of Political Science at University of Gothenburg (UGOT).

digital
society
project

To create that data set, V-Dem has built extensive infrastructure which is specifically designed to collect data on concepts that are difficult to measure, by minimizing as much as possible the bias and error connected with this process. V-Dem's network of experts consists of over 3,200 local and cross-national scholars from more than 180 countries (Mechkova & Sigman, 2018). One of V-Dem's biggest advantages is the way in which V-Dem processes and aggregates the expert-collected data, in order to produce valid and reliable estimates from multiple experts (Coppedge et al., 2018). These experts are typically academics originally from or with extensive experience in the country they are coding (Mechkova & Sigman, 2018). Usually, five independent country experts provide answers to all evaluative indicators, which allows for inter-coder reliability tests and detection of systematic biases.

Biases could come from several sources. First, judgments may differ across experts and cases. In particular, because experts come from different contexts and are not able to communicate with each other, they may apply different standards when answering questions. Second, even equally knowledgeable experts may have different perceptions and disagree with another because of random errors. Therefore, it is imperative to capture and report potential measurement error. To address these issues, V-Dem uses both cutting-edge theory and methods. Pemstein et al. (2018) have developed a custom-made Bayesian Item-Response Theory (IRT) measurement model. This model allows for experts to vary both in reliability—the rate at which they make stochastic errors—and rating thresholds—systematic bias in how they map their perceptions about the world into answers to Likert-scale questions. V-Dem pairs this IRT framework with anchoring vignettes (Pemstein, Seim & Lindberg 2016), which use hypothetical examples to effectively learn how experts' rating thresholds vary. To further enhance cross-expert, and country, comparability, many coders also asked to rate other countries than their original case, providing information about how experts' rating thresholds vary. This modeling framework allows V-Dem both to rationally incorporate information from heterogenous experts, and to quantify the amount of certainty in the resulting data. In particular, V-Dem data are accompanied by confidence intervals that reflect inter-expert (dis)agreement, the amount of information available for each observation (country-year-question), and variation in the reliability of the experts who rate particular cases.

Findings from the v3 (2021) DSP Dataset

Online Mobilization

Several DSP indicators capture dimensions of how the Internet and social media are being used to solve collective action problems, both for good and for ill. In "What types of offline political action are most commonly mobilized on social media?" we find that online organization of offline political action seems to be extremely widespread, with the most common offline political action organized online being street protest (organized online in 77% of countries), petition signing (68%), voter turnout efforts (62%), and strikes/labor actions (44%). There are definite geographic patterns in this data. Whereas organizing petition signing online is extremely common in Western Europe and North America (where more than 80% of countries experienced this in 2020), it is extremely uncommon in sub-Saharan Africa (where only approximately 50% of countries experienced this in 2020). In contrast, it is common to organize strikes or other labor action online in Latin America (where nearly 60% of countries experienced this in 2020), but this is uncommon in Eastern Europe and Central Asia (where less than 30% of countries experienced this in 2020).

Significant violent action is also mobilized online, though these offline actions are more rarely organized online. Terrorism and vigilantism (mob lynching, stalking, harassment) are organized online in less than 10% of countries. Interestingly, these two behaviors are only moderately positively correlated ($r = 0.472, p < 0.000$), indicating that different varieties of violence are organized in different country contexts. In addition, the use of social media in organizing ethnic cleansing or genocide is reported by at least a third of the expert coders in seven countries: Afghanistan, Burma, Ethiopia, India, South Sudan, Sudan, and Yemen.

In Figure 1 and Figure 2, we show the 2020 breakdown across different offline activities that are organized online. Figure 1 depicts the pattern across the world, and Figure 2 depicts the pattern by region.
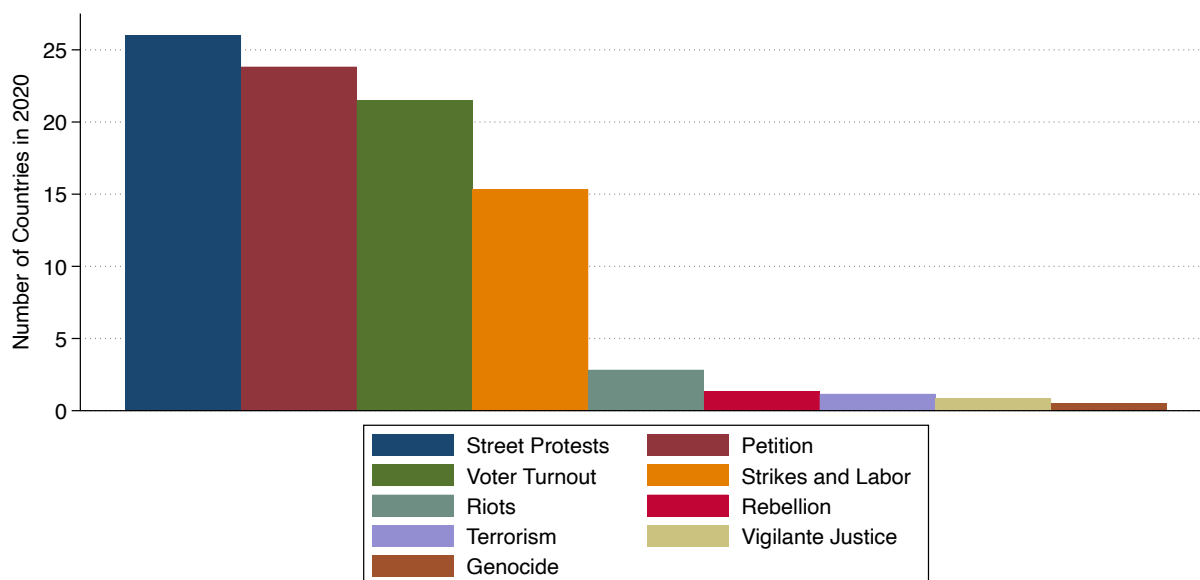
Figure 1. Breakdown in Offline Activities Organized Online across World in 2020
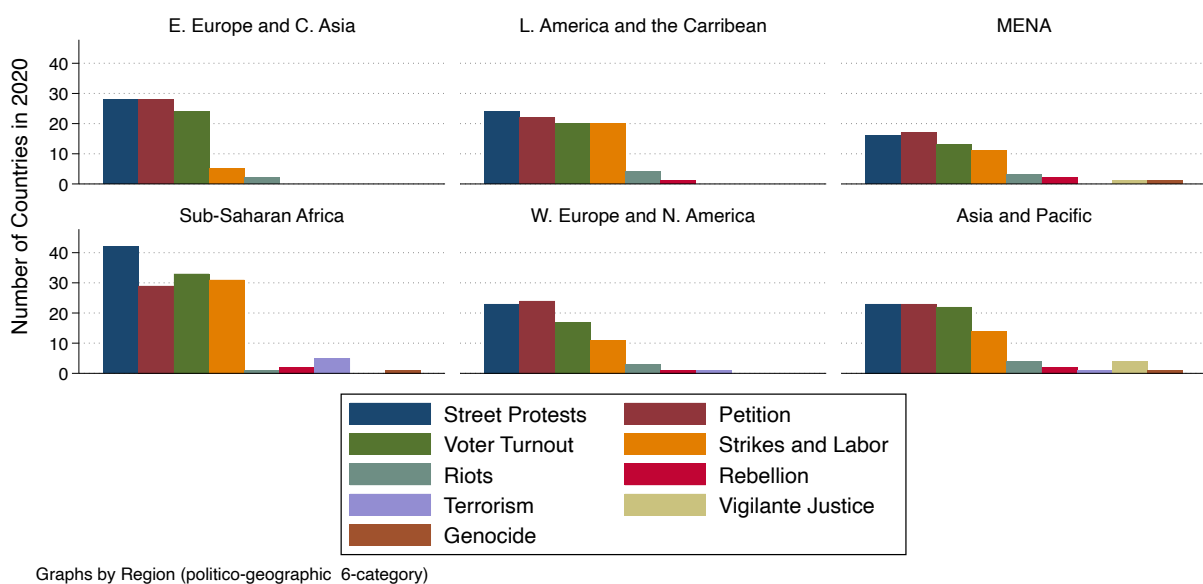

Graphs by Region (politico-geographic 6-category)

Figure 2. Breakdown in Offline Activities Organized Online by Region in 2020

The indicator "How often do people use social media to organize offline violence?" provides some additional perspective, on a three-point Likert scale. In 2020, ten countries – Afghanistan, Central African Republic, Guinea, India, Iraq, Ivory Coast, Libya, Nicaragua, Syria, Turkey – are ranked at the level of "Frequently: There are numerous cases in which people have used social media to organize offline violence." In v1 of the dataset (2019), only five countries were in this category. One hundred

and twenty-two (122) countries ranked in the "sometimes" category in 2020, and 47 countries were ranked in the "never" category in 2020.

In addition, we distinguish which segment of the population is organizing "offline political action of any kind" with social media in a pair of indicators that separately capture whether average people or domestic elites are doing so. Figure 3 shows the relationship between these two indicators. The dark blue lines represent the average scores for 2020 for each of the two indicators. While there is a positive relationship between the two ($r = 0.629$, $p < 0.000$), online mobilization in a set of countries is skewed towards being *either* elite-perpetrated or population-perpetrated. Note that one of the country cases weighted significantly towards elites is Singapore (SGP), and one of the countries weighted significantly towards average people is Belarus (BLR).

Figure 3: Use of Social Media to Mobilize Offline Action (Elites vs. Average People), 2020

Digital Media Freedom

In Figure 4, we examine one of the common tactics to censor political information on the Internet: internet filtering (blocking access to certain websites). We see that there is great variation in the frequency with which governments engage in internet filtering. The countries with the worst record on this indicator are United Arab Emirates, North Korea, Syria, Cuba, and Turkmenistan. On the other side of the spectrum are Denmark, Norway, Luxembourg, Lithuania, and Belgium.
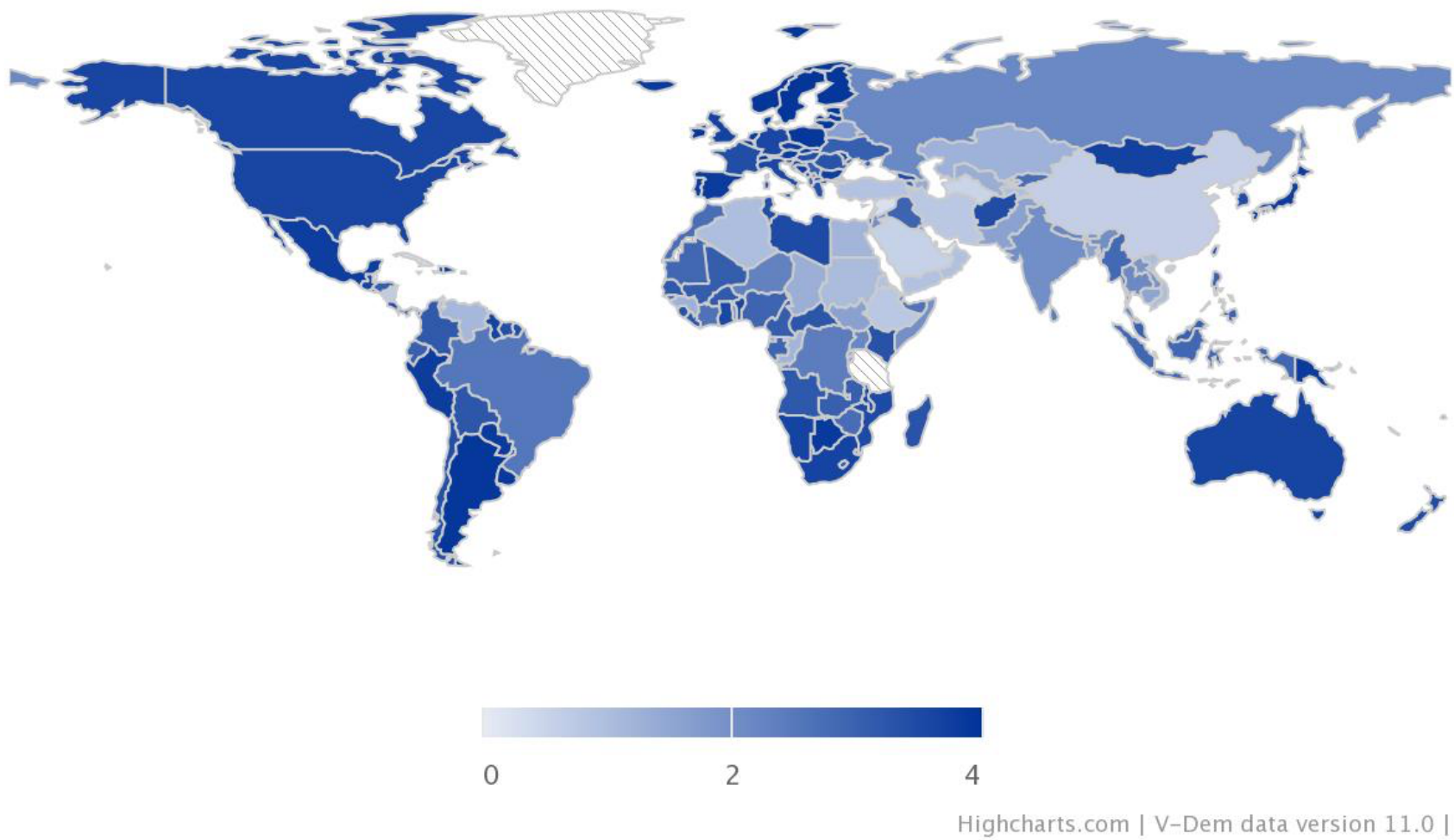
Figure 4. Government Internet Filtering in Practice, 2020

Highcharts.com | V–Dem data version 11.0 |

Figure 5 compares how often governments filter internet content to two other popular tactics: internet shutdown and social media censorship in practice. We see on the graphs that the average levels have not changed much between 2000 and 2020. However, we also see that governments tend to use total Internet shutdown less than the other two tactics, perhaps because of the technical difficulty to totally shutdown the Internet. In comparison, filtering specific content is a much more popular tactic.
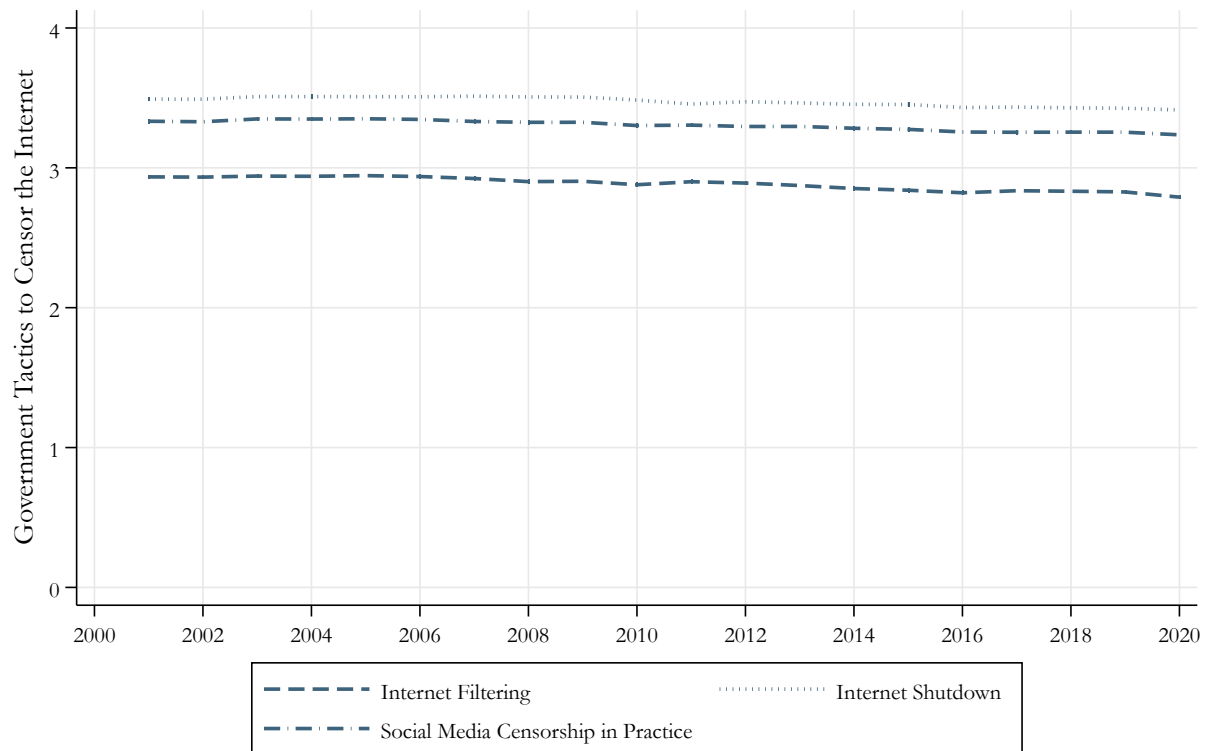

Figure 5. Government Tactics to Censor the Internet

Coordinated Information Operations

Next, we examine the extent to which information operations using online media are present in countries around the world, both by domestic and foreign actors. One concern about this set of questions was whether domestic online media was widespread enough in most countries to represent a concern. That is, in less technically inclined countries, it was plausible that the content of online media is a foreign import due to domestic technical capacity limitations.

As such, we included an indicator in the Online Media Polarization section of the survey to capture the extent to which domestically sourced online media is consumed in each country ("Do people consume domestic online media?"). This indicator is ordinalized into four categories based on the original Likert scale of the question: not at all, limited, relatively extensive, and extensive. As of 2020, not a single country falls into the lowest category, while only 15 fall into the "limited" category (down from 23 in this category in 2018). All other countries have domestic online media consumption that is "relatively extensive" (which is described as "domestic online media consumption is common") or "extensive" ("almost everyone consumes domestic online media"). Since domestic online media consumption is nearly universally high, questions about the distribution of false information in that sphere are particularly salient.

We have several indicators that map the degree to which false information operations exist on social media in each country. First, we ask the degree to which "the government and its agents use social media to disseminate misleading viewpoints or false information to influence its own population." Second, we also measure the degree to which governments use social media to spread false information to "to influence citizens of other countries abroad." These two indicators correlate highly ($r = 0.918$, $p < 0.000$), indicating that the countries employing false information campaigns are doing so to influence both their own populations and those of other countries. Table 1 lists the governments most likely to disseminate false information domestically and abroad. Among the worst offenders on both indicators are Syria, North Korea, Venezuela, China, Russia, South Sudan, and Yemen.

Table 1. Countries Disseminating False Information Domestically and Abroad

Domestically		Abroad	
Country	Score	Country	Score
Syria	0.181	Syria	0.377
North Korea	0.215	North Korea	0.285
Venezuela	0.218	Venezuela	0.272
Cuba	0.246	China	0.433
China	0.431	Russia	0.933
Somalia	0.474	South Sudan	0.813
Russia	0.502	Yemen	0.425

South Sudan	0.513	Azerbaijan	0.829
Yemen	0.564	Eritrea	0.897
Belarus	0.605	Hong Kong	0.931

We also measure the degree to which *foreign* governments use social media to spread false information to influence domestic politics in the country. In Figure 6 we compare this indicator to the one capturing the domestic government's tendency to disseminate false information. The dark blue lines represent the average scores for 2020 for each of the two indicators. We see that the countries being affected the most by foreign governments' dissemination of false information but doing so the least in their own countries are Latvia and Taiwan, followed by West Bank and Kosovo. On the other side of the spectrum are countries which frequently disseminate false information to their own populations, but are relatively free from foreign interference. These are Cuba, Russia, Turkmenistan, and Mexico.
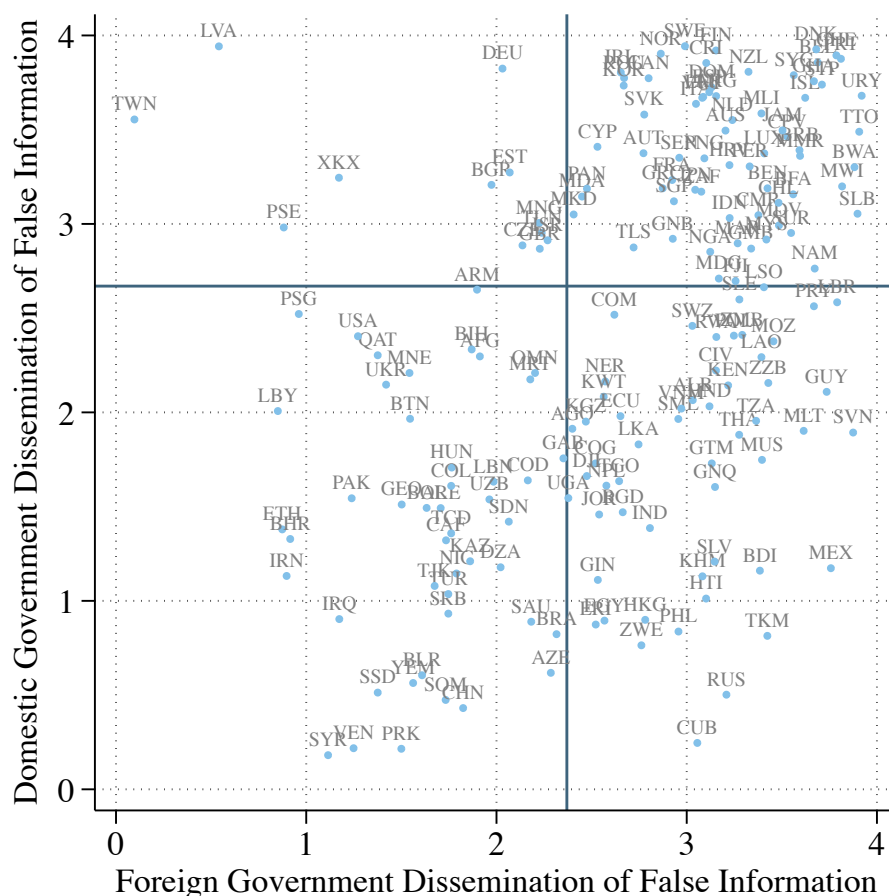
Figure 6. Foreign and Domestic Government Dissemination of False Information, 2020

Finally, we also ask "How routinely do foreign governments and their agents use paid advertisements on social media in order to disseminate misleading viewpoints or false information to influence domestic politics in this country?" In 2020, no countries qualify for the "worst" category in which such action is described as occurring "extremely often" and with regard to "all key political issues," whereas three countries qualified for this category in 2018. Similarly, whereas in 2018 only 36 countries achieved the "best" rating, in which such action is described as occurring "never or almost never," 53 countries achieved this rating in 2020.

Social Cleavages

With the rise of Internet usage around the world, the new communications platforms have been colonized by offline hatred. We measure this in part with the multiple selection question "Which groups are targets of hate speech or harassment in online media?". In 2020, the most common specific

targets were LGBTQ groups and individuals (in 59% of countries), specific ethnic groups (50%), specific religious groups (39%), and women (36%). While alarmingly high, we note that these statistics have declined significantly since the v1 dataset release in 2021, from 76%, 66%, 58%, and 51%, respectively. In 21% of the countries for which we have data in the v3 release, the expert consensus holds that no specific groups are targeted by hate speech or harassment online. This lack of online harassment, however, is negatively correlated with consumption of domestic online media ($r = -0.213$, $p < 0.004$), implying that the growth of the online space generally brings the growth of online hate speech and harassment. In Figure 7 and Figure 8, we show the 2020 breakdown across different targeted groups. Figure 7 depicts the pattern across the world, and Figure 8 depicts the pattern by region.
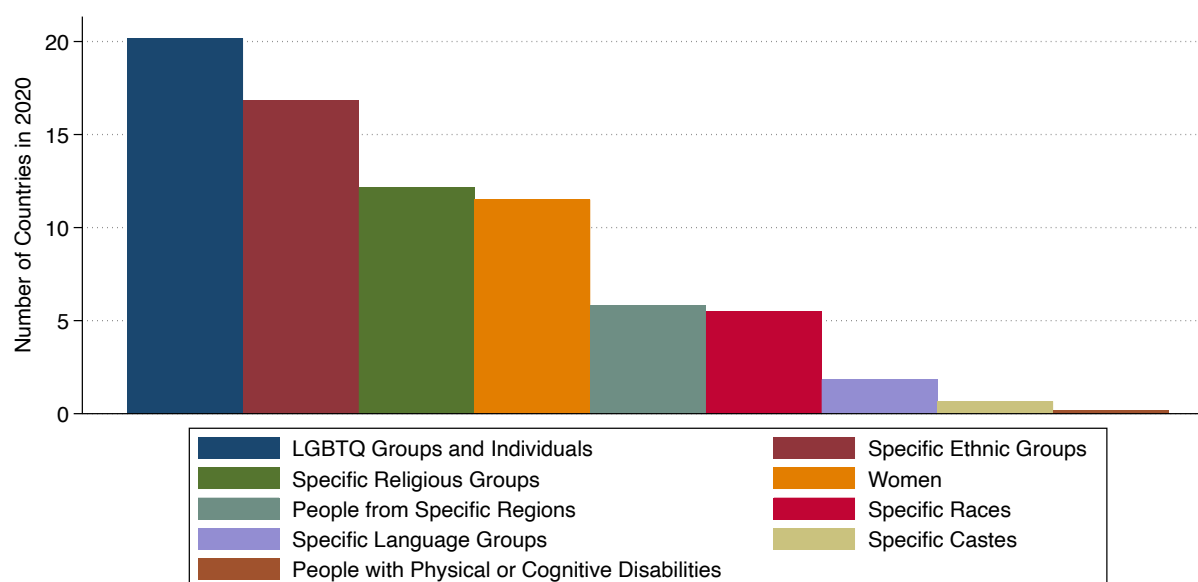

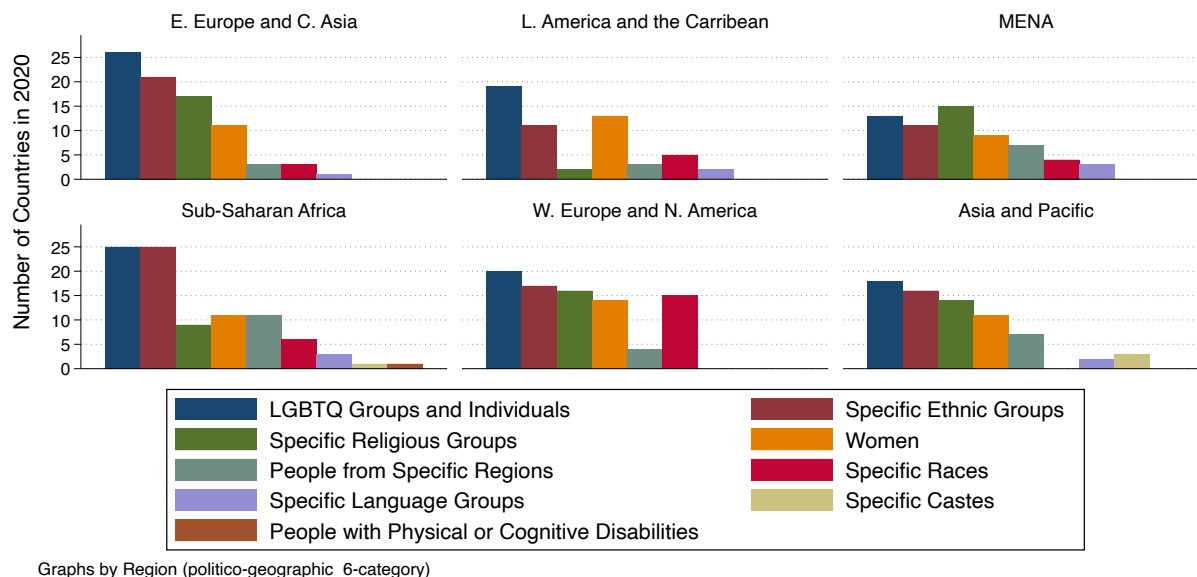Figure 7. Breakdown in Groups Harassed Online across World in 2020

Figure 8. Breakdown in Groups Harassed Online by Region in 2020

In Figure 9, we present a scatterplot of two indicators: arrests for political content; and government social media monitoring. The dark blue lines represent the average scores for 2020 for each of the two indicators. We see that these two indicators are highly correlated ($r = 0.823$, $p < 0.000$), and countries that are relatively democratic on both dimensions cluster at the upper-right corner of the figure. The implication of this figure is that governments that closely monitor social media are quite likely to follow-up on the information they find to arrest citizens. In 43% of the countries, if a citizen posts political content online that would run counter to the government and its policies, they would be "likely" or "extremely likely" to get arrested.
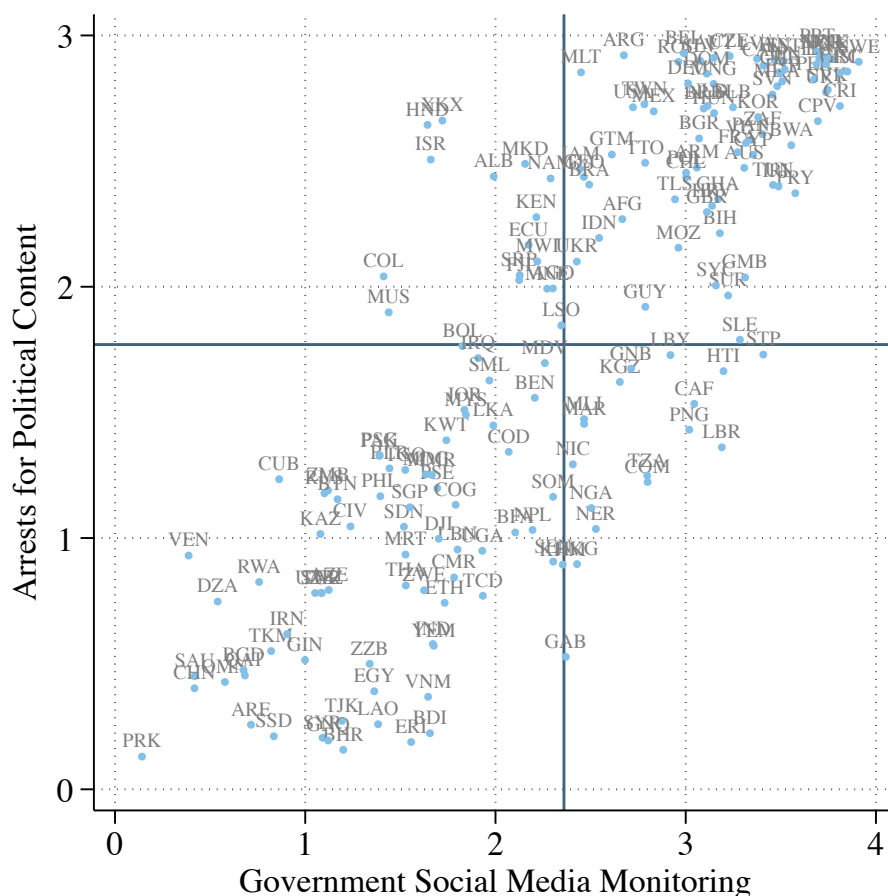
Figure 9. Arrests for political content and Government monitoring, 2020

Online Media Polarization

We next examine one of the indicators in the survey section Online Media Polarization that measures the extent to which online media outlets give a similar presentation of major (political) news. Figure 10 shows the average scores on this measure from 2000 to 2020 aggregated by region. We see that the least fractionalized countries tend to concentrate in Western Europe and North America (blue dashed line), while the sub-Saharan African countries appear to be the most fractionalized. It is also clear that the MENA region has improved on this indicator over time and the Eastern Europe and Central Asia region has declined.
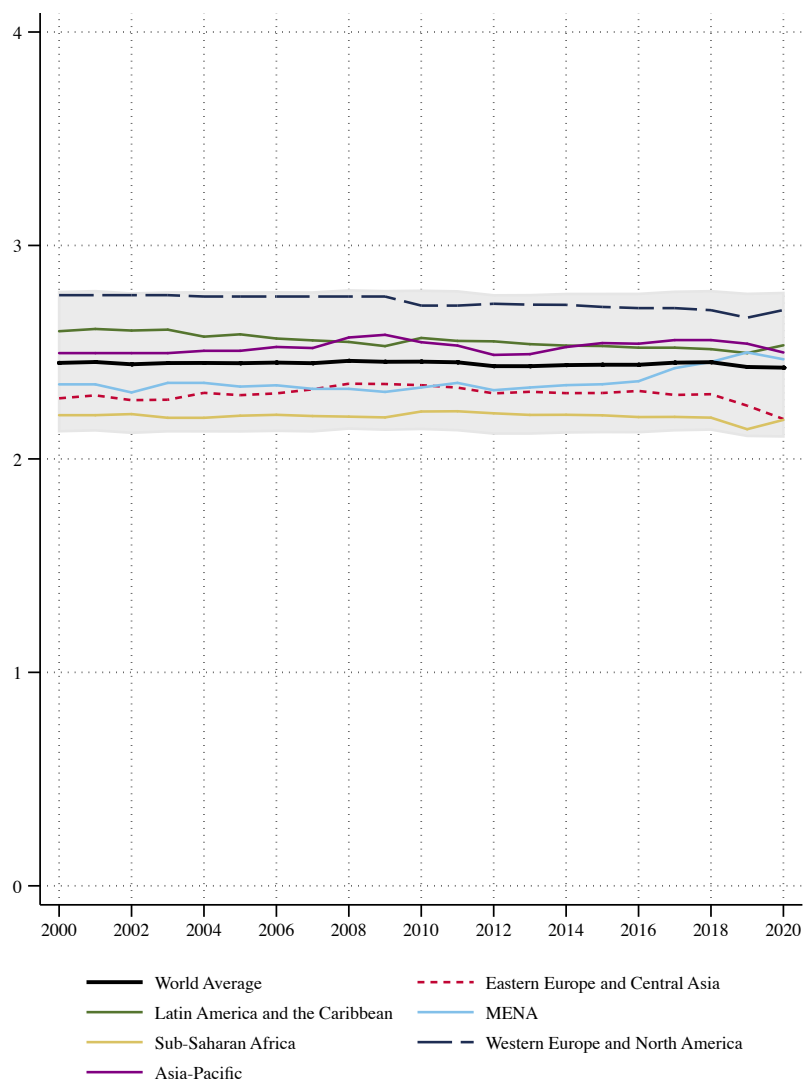
Figure 10: Fractionalization of Online Media Perspectives, 2000-2020

Conclusion

We hope this report provides compelling insights useful for a wide audience of academics, policy-makers and citizens, interested in the development of digital media and its relationship to democracy.

digital
society
project

References

Adler, Julie (2011). The Public's Burden in the Digital Age: Pressures on Intermediaries and the Privatization of Internet Censorship. Journal of Law and Policy 20(1):231–266.

Bak, Daehee, Surachanee Sriyai & Stephen Meserve (2018). The internet and state repression: A cross-national analysis of the limits of digital constraint. *Journal of Human Rights* 17(5):642–659.

Bunce, Valerie J and Sharon L Wolchik (2010). "Defeating Dictators : Electoral Change and Stability in Competitive Authoritarian Regimes". In: *World Politics* 62.1, pp. 43-86.

Castells, Manuel (2009). The Information Age: Economy, Society, and Culture. Wiley Blackwell.

Chowdhury, Mridul (2008). "The Role of the Internet in Burma's Saffron Revolution". In: *SSRN Electronic Journal.*

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, Joshua Krusell, Kyle L. Marquardt, Daniel Pemstein, Josefine Pernes, Natalia Stepanova, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, and Steven Wilson. (2018). "V-Dem Methodology v8" Varieties of Democracy (V-Dem) Project.

Deibert, R.J., J.G. Palfrey, R. Rohozinski & J. Zittrain (2008). *Access Denied: The Practice and Policy of Global Internet Filtering.* Cambridge, MA: MIT Press.

Diamond, Larry (2010). "Liberation technology". In: *Journal of Democracy* 21.3, pp. 69-83.

Dyczok, Marta (2005). "Breaking Through the Information Blockade : Election and Revolution in Ukraine 2004". In: *Canadian Slavonic Papers* 47.4, p. 241.

Garcia, Cristobal et al. (2013). "What can Social Media teach us about protests? Analyzing the Chilean 2011-12 Student Movement's Network evolution through Twitter data". In: *arXiv preprint arXiv:1308.2451.*

Gentzkow, Matthew and Jesse M. Shapiro (Apr 2010). "Ideological Segregation Online and Offline". In: *National Bureau of Economic Research.*

Tarleton Gillespie (2018). Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. New Haven: Yale University Press.

Gohdes, Anita R. (2015). Pulling the Plug: Network Disruptions and Violence in Civil Conflict. *Journal of Peace Research* 52(3):352–367.

Gohdes, Anita R. (Forthcoming). Repression technology: Internet accessibility and state violence. *American Journal of Political Science.*

Goldstein, Joshua (2007). "The role of digital networked technologies in the Ukrainian Orange Revolution". In: *Berkman Center Research Publication* 2007-14.

Hofheinz, Albrecht (2005). "The Internet in the Arab World: Playground for Political Liberalization". In: *International Politics and Society*, pp. 78-96.

Kemp, Simon. (2019). Digital in 2019. Essential insights into Internet, social media, mobile, and ecommerce use around the world. https://wearesocial.com/uk/digital-2019 Accessed April 1st, 2019.

King, Gary, Jennifer Pan & Margaret E. Roberts (2013). How Censorship in China Allows Government Criticism but Silences Collective Expression. *American Political Science Review* 107(02):326–343.

Kirkegaard, Emil O. W. (2015). The general socioeconomic factor among Colombian departments. Tech. rep.

Kyj, Myroslaw J. (Jan. 2006). "Internet use in Ukraine's Orange Revolution". In: *Business Horizons* 49.1, pp. 71-80.

Mackell, Austin (2011). "The IMF versus the Arab spring".pp. 2-4.

Meserve, Stephen (2018). Why Do Governments Censor? Expanding from State Survival to Content Regulation Theories in Political Science. *APSA Comparative Politics Newsletter* 18(2):55–59.

Meserve, Stephen A. & Daniel Pemstein (2018). Google Politics: The Political Determinants of Internet Censorship in Democracies. P*olitical Science Research & Methods* 6:245– 263.

Morozov, Evgeny (Feb. 2012). The Net Delusion: The Dark Side of Internet Freedom. Public Affairs.

Murphy, Emma C (Sept. 2006). "Agency and Space: the political impact of information technologies in the Gulf Arab states". In: *Third World Quarterly* 27.6, pp. 1059-1083.

Oghia, Michael J and Helen Indelicato (2010). "Ruling the Arab Internet : An Analysis of Internet Ownership Trends of Six Arab Countries".

Page, Scott E. (2008). The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton University Press.

Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y. T., Krusell, J., & Miri, F. (2018). The V-Dem measurement model: latent variable analysis for cross-national and cross-temporal expert-coded data.

Ramakrishnan, Naren et al. (2014). Beating the News with EMBERS: Forecasting Civil Unrest using Open Source Indicators. Tech. rep.

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

Scherman, Andres, Arturo Arriagada, and Sebastian Valenzuela (2015). "Student and Environmental Protests in Chile: The Role of Social Media". In: *Politics* 35.2, pp. 151-171. issn: 1467-9256. doi: 10.1111/1467-9256.12072.

Sereghy, Zsolt, Sarah Bunk, and Bert Preiss, eds. (2012). The Arab Revolutions: Reflections on the Role of Civil Society, Human Rights and New Media in the Transformation Processes. Austrian Study Center for Peace and Conflict Resolution.

Shirky, Clay (Feb. 2009). Here Comes Everybody: The Power of Organizing Without Organizations. Penguin.

Stepanova, Ekaterina (2011). "The role of information communication technologies in the Arab Spring". In: *Ponars Eurasia* 15, pp. 1-6.

Sunstein, Cass R (2009). *Republic.com 2.0*. Princeton University Press.

Tufekci, Zeynep (2017). Twitter and Tear Gas: The Power and Fragility of Networked Protest. New Haven: Yale University Press.

Tufecki, Zeynep and Christopher Wilson (2012). Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. *Journal of Communication* 62(2), 363-379.

Valenzuela, Sebastian (2013). "Unpacking the use of social media for protest behavior the roles of information, opinion expression, and activism". In: *American Behavioral Scientist* 57.7, pp. 920-942.

Warren, T Camber (2015). Explosive Connections? Mass Media, Social Media, and the Geography of Collective Violence in African States. *Journal of Peace Research* 52(3):297– 311.

Wilson, Steven Lloyd (2016). "Information and Revolution". PhD thesis. University of Wisconsin-Madison.

Wilson, Steven Lloyd (2017). Detecting Mass Protest through Social Media. *Journal of Social Media in Society*, 6(2), 5–25.

Wojcieszak, Magdalena E and Diana C Mutz (2009). "Online groups and political discourse: Do online discussion spaces facilitate exposure to political disagreement?" In: *Journal of Communication* 59.1, pp. 40-56.

Zhuo, Xiaolin, Barry Wellman, and Justine Yu (2011). "Egypt: The First Internet Revolt?" In: *Peace Magazine* Jul/Sep, pp. 6-10.

Zittrain, J. (2003). Internet Points of Control. *Boston College Law Review* 44:653–688.