Digital Society Survey

Codebook

Valeriya Mechkova (University of Gothenburg)
Daniel Pemstein (North Dakota State University)
Brigitte Seim (University of North Carolina, Chapel Hill)
Steven Wilson (University of Nevada, Reno)

April 2019

Contents

1 Introduction

The Digital Society Survey, designed by the Digital Society Project, contains questions pertaining to the political environment of the internet and social media. These data, which we collected using expert-coded surveys, provide information on topics related to coordinated information operations, digital media freedom, online media polarization, social cleavages, and state internet regulation capacity and approach.

For more information, please visit `https://www.digitalsocietyproject.org`

1.1 Funders

The Varieties of Democracy Institute (V-Dem) provided the use of its infrastructure for this data collection project. The V-Dem data team processed this survey using the standard V-Dem measurement modeling and quality control processes, using the V-Dem Institute infrastructure for collection and aggregating expert-based data on democracy, reaching out to a network of over 3,200 scholars from more than 180 countries. To learn more about V-Dem, and its funders, please visit: `https://www.v-dem.net/`.

In addition, DSP received support from Facebook to cover the costs of initial data collection. The National Science Foundation provided support (Grant No. SES-1423944) for the development of the methodological tools upon which we rely.

For data enquires: contact@digitalsocietyproject.org

1.2 Cautionary Notes

Both DSP and V-Dem are firmly committed to full transparency and data sharing. We ask users to take the following cautions into consideration when using the dataset.

- The *V-Dem Methodology* assumes five or more coders for the "contemporary" period. DSP spans a subset of this period, covering 2000–2018. Given that this was the first round of data collection for the DSP variables, we were not always able to recruit five country-experts to rate each observation. We urge users to exercise caution when working with observations that were rated by fewer than five experts, and to pay careful attention to estimates of uncertainty around the point estimates that we provide for each observation. We strongly advise against using point estimates for country-variable-years with three or fewer (≤ 3) ratings. We suggest filtering these observations out before conducting any type of analysis. For this purpose, a special count-variable for each Country-Expert coded variable, which is suffixed with "_nr", is included in the dataset.

1.3   Variable Types

The DSP data is gathered following the V-Dem rules and procedures. While V-Dem contains a number of types of data, all DSP variables are what V-Dem calls 'C-type' variables:

- **Type C: Variables coded by Country Experts**
  A Country Expert is typically a scholar or professional with deep knowledge of a country and of a particular political institution. Furthermore, the expert is usually a citizen or resident of the country. Multiple experts (usually 5 or more) code each variable. More information about the Country Experts can be found in the *V-Dem Methodology* document.

1.4   Variable Versions and Suffixes

Just like the V-Dem Dataset, the DSP data set contains several versions of the variables coded by country experts (type C variables).

- **Model Estimates**
  "Model Estimates" — Measurement Model Output:
  This version has no special suffix (*e.g.* v2smgovshut). This version of the variables provides country-year (country-date in the alternative dataset) point estimates from the V-Dem measurement model (see Pemstein et al. 2019). The measurement model aggregates the ratings provided by multiple country experts and, taking disagreement and measurement error into account, produces a probability distribution over country-year scores on a standardized interval scale (see the *V-Dem Methodology* document). The point estimates are the median values of these distributions for each country-year. The scale of a measurement model variable is similar to a normal ("Z") score (*e.g.* typically between -5 and 5, with 0 approximately representing the mean for all country-years in the sample) though it does not necessarily follow a normal distribution. For most purposes, these are the preferred versions of the variables for time series regression and other estimation strategies.

  "Model Estimates Measure of Uncertainty" — Measurement Model Highest Posterior Density (HPD) Intervals:
  This version has the suffixes: "codelow" and "codehigh" (*e.g.* v2smgovshut_codelow and v2smgovshut_-codehigh). These two kinds of variables ["code low" and "code high"] demarcate the interval in which the measurement model places 68 percent of the probability mass for each country-year score, which is approximately equivalent to one standard deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (*e.g.*, v2smgovshut_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Original Scale (*_osp)**
  "Original Scale" — Linearized Original Scale Posterior Prediction:
  This version has the suffix _osp, (*e.g.* v2smgovshut_osp). In this version of the variables, we have linearly translated the measurement model point estimates back to the original ordinal scale of each variable (*e.g.* 0–4 for v2smgovshut_osp) as an interval measure. The decimals in the _osp version roughly indicate the distance between the point estimate from the linearized measurement model posterior prediction and the threshold for reaching the next level on the original ordinal scale. Thus, a _osp value of 1.25 indicates that the median measurement model posterior predicted value was closer to the ordinal value of 1 than 2 on the original scale. Technically, it calculates the sum of the posterior probabilities that the estimate is in a particular category: If a particular country-year-variable has a probability of 90% to be in category 4, a 10% probability of being in category 3, and 0% probability of being in categories 2, 1, and 0, the result is a value of 3.9 (4*0.9 + 3*0.1 = 3.6+0.3). Since there is no conventional theoretical justification for linearly mapping ordinal posterior predictions onto an interval scale, these scores should primarily be used for heuristic purposes. Using the Ordinal Scale estimates— or incorporating the properties of ordinal probit models into the estimation procedure—is thus preferable to using the _osp estimates in statistical analyses. However, since the _osp version maps onto the coding criteria found in the V-Dem Codebook, and is strongly correlated with the

Measurement Model output (typically at .98 or higher), some users may find the _osp version useful in estimating quantities such as marginal effects with a clear substantive interpretation. If a user uses _osp data in statistical analyses it is imperative that she confirm that the results are compatible with estimations using Measurement Model output.

"Original Scale Measure of Uncertainty" — Linearized Original Scale HPD Intervals:
This version has the suffixes "codelow" and "codehigh" (*e.g.* v2smgovshut_osp_codelow and v2smgovshut_osp_codehigh). We estimate these quantities in a similar manner as the Measurement Model Highest Posterior Density Intervals. These two variables ["code low" and "code high"] demarcate the interval in which the measurement model places 70 percent of the probability mass for each country-year score, which is approximately equivalent to one standard deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (*e.g.*, v2smgovshut_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Ordinal Scale (*_ord)**
  "Ordinal Scale" — Measurement Model Estimates of Original Scale Value:
  This version has the suffix "_ord" (*e.g.* v2smgovshut_ord). This method translates the measurement model estimates back to the original ordinal scale of a variable (as represented in the Codebook) after taking coder disagreement and measurement error into account. More precisely, it represents the most likely ordinal value on the original codebook scale into which a country-year would fall, given the average coders usage of that scale. More specifically, we assign each country-year a value that corresponds to its integerized median ordinal highest posterior probability category over Measurement Model output.

  "Ordinal Scale Measure of Uncertainty" — Original Scale Value HPD Intervals:
  This version has the suffixes - "codelow" and "codehigh" (*e.g.* v2smgovshut_ord_codelow and v2smgovshut_ord_codehigh). We estimate these values in a similar manner as the Measurement Model Highest Posterior Density Intervals. These two variables ["code low" and "code high"] demarcate the interval in which the measurement model places 70 percent of the probability mass for each country-year score, which is approximately equivalent to one standard deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (*e.g.* v2smgovshut_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Number of Coders per Country, Variable and Year/Date (*_nr)**
  The number of V-Dem Country Experts (regular coders, bridge- and lateral coders) who provided data on country, variable and year. V-Dems methodology is based on the assumption that we have a minimum of five Country Experts for every single country-variable-year. Sometimes, however, we end up with fewer than five Country Experts. From v7 of the Country-Year, and the Country-Date type datasets, we provide all data we have for full transparency. By providing the number of Country Experts for each country-variable-year/date, we suggest that users primarily base analyses on observations based on five or more coders. We strongly advise against using observations based on three or fewer coders. This concerns all C type variables.

1.5 Aggregation

C-variables, ratio/percentage variables, and High-Level/Mid-Level Democracy indices are aggregated from the country-date level to the country-year level by the day-weighted mean.

1.6 Variable Information

The following information is available per variable (if applicable):

Additional versions : Indicates if the variable is also available in the following versions; *_-osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean and/or *_nr. Detailed information about the different versions can be found in section 1.5 (*Variable Versions and Suffixes*).

Available versions: Lists the available variable types (Only applicable for ordinalized versions of indices).

Question: The question that the variable attempts to measure.

Clarification: Definition of key terms, clarification of scope-conditions, contexts, and any other features needed to understand the question (if any). All key terms appear in the Glossary (*Appendix B*), unless they are specific to a single section (in which case they only appear in the introduction to that section or in the clarifications for particular questions). Key terms are sometimes cross-referenced with hypertext.

Responses: Numeric, Percentage, Text, Date, Countries, or specific response categories (listed below under "Answer-types" and "Scales").

Answer-Types:
Multiple-choice: Where a coder can select only one answer. This is the usual protocol and is therefore not noted.
Multiple-selection: Where a coder can select more than one answer. For most multiple-selection variables, the dataset contains both the original variable as well as a set of dummies for each of the responses.

Ordering (only applicable to a selection of C variables): This relates to the ordering of questions when the coding of one indicator depends upon the coding of other indicators (*i.e.*, whenever there is some alteration of the serial ordering of questions as listed in this document).

Aggregation (only applicable to indices): Explanation of how an index is constructed.

Scale: Dichotomous, Nominal, Ordinal, or Interval/Ratio (Extra response options such as N/A or Other, are not counted as part of this classification).

Cross-Coder Aggregation (only applicable to C variables): IRT, Bayesian ordinal item response theory measurement model (see the *V-Dem Methodology* document). Available in mode and mean.

Cleaning: Specifies if observations are set to missing based on values on other variables.

Citation: Suggested citation when using the specific variable.

Years: Available coverage for the respective variable. For more information on country-specific year coverage, see *the country table.*

Note: Additional information about the variable.

1.7 Suggested Citation

Nota bene: If a variable drawn from the DSP dataset plays an important role in your project (published or unpublished), please cite the following:

- **DSP Dataset:**
  Mechkova, Valeriya, Daniel Pemstein, Brigitte Seim, and Steven Wilson. 2019. "DSP [Country-Year] Dataset v9" Digital Society Project (DSP).

- **DSP Introduction:**
  Mechkova, Valeriya, Daniel Pemstein, Brigitte Seim, and Steven Wilson. 2019. "Measuring Internet Politics: Introducing the Digital Society Project" Digital Society Project (DSP).

- **V-Dem Methodology:**
  Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Kyle L. Marquardt, Juraj Medzihorsky, Daniel Pemstein, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Eitan Tzelgov, Yi-ting Wang, and Steven Wilson. 2019. "V-Dem Methodology v9" Varieties of Democracy (V-Dem) Project.

- **V-Dem Measurement Model**:
  Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, ăJuraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer. 2019.

Given the DSP's intellectual debt to the large V-Dem project, we strongly encourage that users of the data include some or all of the following citations:

- **V-Dem Dataset:**
  Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Steven Wilson, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Nina Ilchenko, Joshua Krusell, Valeriya Mechkova, Juraj Medzihorsky, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2019. "V-Dem [Country-Year/Country-Date] Dataset v9" Varieties of Democracy (V-Dem) Project.

- **V-Dem Codebook:** Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Valeriya Mechkov, Johannes von Römer, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2019. "V-Dem Codebook v9" Varieties of Democracy (V-Dem) Project.

- **V-Dem Country Coding Units:**
  Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Vlad Ciobanu, and Lisa Gastaldi. 2019. "V-Dem Country Coding Units v9" Varieties of Democracy (V-Dem) Project.

- **V-Dem Organization and Management:**
  Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Lisa Gastaldi, Nina Ilchenko, Josefine Pernes, Johannes von Römer, and Natalia Stepanova. 2019. "V-Dem Organization and Management v9" Varieties of Democracy (V-Dem) Project.

1.8 Countries

The following table contains all country units (and their year coverage) that are included in the V-Dem Dataset. Some countries are coded prior to independence, and some have gaps in their coding periods. For a more detailed description of the country units and their year coverage please consult the V-Dem *Country Coding Units* document.

Name	ID	Coverage	Name	ID	Coverage
Afghanistan	36	2000–2018	Ecuador	75	2000–2018
Albania	12	2000–2018	Egypt	13	2000–2018
Algeria	103	2000–2018	El Salvador	22	2000–2018
Angola	104	2000–2018	Equatorial Guinea	160	2000–2018
Argentina	37	2000–2018	Eritrea	115	2000–2018
Armenia	105	2000–2018	Estonia	161	2000–2018
Australia	67	2000–2018	Ethiopia	38	2000–2018
Austria	144	2000–2018	Fiji	162	2000–2018
Azerbaijan	106	2000–2018	Finland	163	2000–2018
Bahrain	146	2000–2018	France	76	2000–2018
Bangladesh	24	2000–2018	Gabon	116	2000–2018
Barbados	147	2000–2018	Georgia	118	2000–2018
Belarus	107	2000–2018	Germany	77	2000–2018
Belgium	148	2000–2018	Ghana	7	2000–2018
Benin	52	2000–2018	Greece	164	2000–2018
Bhutan	53	2000–2018	Guatemala	78	2000–2018
Bolivia	25	2000–2018	Guinea	63	2000–2018
Bosnia and Herzegovina	150	2000–2018	Guinea-Bissau	119	2000–2018
Botswana	68	2000–2018	Guyana	166	2000–2018
Brazil	19	2000–2018	Haiti	26	2000–2018
Bulgaria	152	2000–2018	Honduras	27	2000–2018
Burkina Faso	54	2000–2018	Hong Kong	167	2000–2018
Burma/Myanmar	10	2000–2018	Hungary	210	2000–2018
Burundi	69	2000–2018	Iceland	168	2000–2018
Cambodia	55	2000–2018	India	39	2000–2018
Cameroon	108	2000–2018	Indonesia	56	2000–2018
Canada	66	2000–2018	Iran	79	2000–2018
Cape Verde	70	2000–2018	Iraq	80	2000–2018
Central African Republic	71	2000–2018	Ireland	81	2000–2018
Chad	109	2000–2018	Israel	169	2000–2018
Chile	72	2000–2018	Italy	82	2000–2018
China	110	2000–2018	Ivory Coast	64	2000–2018
Colombia	15	2000–2018	Jamaica	120	2000–2018
Comoros	153	2000–2018	Japan	9	2000–2018
Costa Rica	73	2000–2018	Jordan	83	2000–2018
Croatia	154	2000–2018	Kazakhstan	121	2000–2018
Cuba	155	2000–2018	Kenya	40	2000–2018
Cyprus	156	2000–2018	Kosovo	43	2000–2018
Czech Republic	157	2000–2018	Kuwait	171	2000–2018
Democratic Republic of the Congo	111	2000–2018	Kyrgyzstan	122	2000–2018
Denmark	158	2000–2018	Laos	123	2000–2018
Djibouti	113	2000–2018	Latvia	84	2000–2018
Dominican Republic	114	2000–2018	Lebanon	44	2000–2018

Name	ID	Coverage	Name	ID	Coverage
Lesotho	85	2000–2018	Saudi Arabia	197	2000–2018
Liberia	86	2000–2018	Senegal	31	2000–2018
Libya	124	2000–2018	Serbia	198	2000–2018
Lithuania	173	2000–2018	Seychelles	199	2000–2018
Luxembourg	174	2000–2018	Sierra Leone	95	2000–2018
Macedonia	176	2000–2018	Singapore	200	2000–2018
Madagascar	125	2000–2018	Slovakia	201	2000–2018
Malawi	87	2000–2018	Slovenia	202	2000–2018
Malaysia	177	2000–2018	Solomon Islands	203	2000–2018
Maldives	88	2000–2018	Somalia	130	2000–2018
Mali	28	2000–2018	Somaliland	139	2000–2018
Malta	178	2000–2018	South Africa	8	2000–2018
Mauritania	65	2000–2018	South Korea	42	2000–2018
Mauritius	180	2000–2018	South Sudan	32	2011–2018
Mexico	3	2000–2018	South Yemen	23	2000–2000
Moldova	126	2000–2018	Spain	96	2000–2018
Mongolia	89	2000–2018	Sri Lanka	131	2000–2018
Montenegro	183	2000–2018	Sudan	33	2000–2018
Morocco	90	2000–2018	Suriname	4	2000–2018
Mozambique	57	2000–2018	Swaziland	132	2000–2018
Namibia	127	2000–2018	Sweden	5	2000–2018
Nassau	366	2000–1866	Switzerland	6	2000–2018
Nepal	58	2000–2018	Syria	97	2000–2018
Netherlands	91	2000–2018	Taiwan	48	2000–2018
New Zealand	185	2000–2018	Tajikistan	133	2000–2018
Nicaragua	59	2000–2018	Tanzania	47	2000–2018
Niger	60	2000–2018	Thailand	49	2000–2018
Nigeria	45	2000–2018	The Gambia	117	2000–2018
North Korea	41	2000–2018	Timor-Leste	74	2000–2018
Norway	186	2000–2018	Togo	134	2000–2018
Oman	187	2000–2018	Trinidad and Tobago	135	2000–2018
Pakistan	29	2000–2018	Tunisia	98	2000–2018
Palestine/British Mandate	209	2000–1948	Turkey	99	2000–2018
Palestine/Gaza	138	2000–2018	Turkmenistan	136	2000–2018
Palestine/West Bank	128	2000–2018	Uganda	50	2000–2018
Panama	92	1903–2018	Ukraine	100	2000–2018
Papua New Guinea	93	2000–2018	United Arab Emirates	207	2000–2018
Paraguay	189	2000–2018	United Kingdom	101	2000–2018
Peru	30	2000–2018	United States of America	20	2000–2018
Philippines	46	2000–2018	Uruguay	102	2000–2018
Poland	17	2000–2018	Uzbekistan	140	2000–2018
Portugal	21	2000–2018	Vanuatu	206	2000–2018
Qatar	94	2000–2018	Venezuela	51	2000–2018
Republic of the Congo	112	2000–2018	Vietnam	34	2000–2018
Romania	190	2000–2018	Yemen	14	2000–2018
Russia	11	2000–2018	Zambia	61	2000–2018
Rwanda	129	2000–2018	Zanzibar	236	2000–2018
Sao Tome and Principe	196	2000–2018	Zimbabwe	62	2000–2018
			Total number of countries	**202**	

1.9 Identifier Variables in the V-Dem and DSP Datasets

1.9.1 Country Name (country_name)

Name of coded country. A V-Dem country is a political unit enjoying at least some degree of functional and/or formal sovereignty. For more details on country units consult the V-Dem *Country Coding Units* document.

Response: Text.

1.9.2 V-Dem Country ID (country_id)

Unique country ID designated for each country. A list of countries and their corresponding IDs used in the V-Dem dataset can be found in the country table in the codebook, as well as in the V-Dem *Country Coding Units* document.

Response: Numeric.

1.9.3 Country Name Abbreviation (country_text_id)

Abbreviated country names.

Response: Text.

1.9.4 Year (year)

V-Dem year coded annually from 2000–2018. This variable is included in the V-Dem Country Year as well as Country Date datasets.

Response: Date.

1.9.5 Start of Coding Period (codingstart)

The DSP country coding starts in 2000, or from when a country first enjoyed at least some degree of functional and/or formal sovereignty. For detailed information, please see the V-Dem *Country Coding Units* document.

Response: Date.

1.9.6 Gap in Coding Period Starts (gapstart)

Time periods when a country does not fulfill V-Dems coding period criteria are not coded. The date that indicates the gap start is the last date coded before the gap. For more details about V-Dem country coding periods, please see the V-Dem *Country Coding Units* document.

Response: Date.

1.9.7 Gap in Coding Period Ends (gapend)

The periods of when a country does not fulfill V-Dems coding period macriteria are not coded. The date that indicates the gap end is the first date coded after the gap. For more details about V-Dem country coding periods, please see the V-Dem *Country Coding Units* document.

Response: Date.

1.9.8 End of Coding Period (codingend)

The DSP country coding ends in 2018, or from when a country formally stopped enjoying at least some degree of functional and/or formal sovereignty. For detailed information, please see the V-Dem *Country Coding Units* document.

Response: Date.

2 Digital Society Survey

The Digital Society Survey, designed by the Digital Society Project, contains questions pertaining to the political environment of the internet and social media. The data collected through expert-coded surveys provides information on topics related to coordinated information operations, digital media freedom, online media polarization, social cleavages, as well as state internet regulation capacity and approach.

Principal investigators for the Digital Society Project are Valeriya Mechkova, Daniel Pemstein, Brigitte Seim, Steven Wilson.

For more information, please visit `https://www.digitalsocietyproject.org`.

Instructions to the coders (as shown in the surveys)

Digital society: The following survey contains questions pertaining to the political environment of the Internet and social media. Please bear in mind the following definitions as you respond to questions on this survey:

The government and its agents include official government organs, such as bureaucracies, courts, intelligence services, and the military, but also unofficial agents, such as officially unaffiliated cyber-warfare operatives who perform services, even off-book work, on behalf of the government.

Major political parties include the group of political parties that hold a significant number of seats in national legislative body(-ies), or earn a significant number of votes in elections for the executive. When we ask you to consider major political parties, you do not need to consider parties that run in elections but receive only a small minority of seats or votes, or those that receive no seats at all.

We define the Internet as all information that people access over public and private digital networks, worldwide. The Internet includes both publicly accessible digital spaces and private or gated information transmission platforms. The Internet does not include traditional media transmission mechanisms such as paper, television, traditional voice telephone, and radio.

Social media are a subset of Internet platforms that enable normal individuals to create and share content with networks of other people. Social media platforms are available to the public, although content on such networks may be shared privately within subgroups of users. Social media includes both publicly visible, or semi-public platforms, like Facebook, Flickr, Friendster, Google+, Instagram, Myspace, LinkedIn, Twitter, VKontakte, and Weibo and private social networking and messaging platforms like Signal, Slack, Snapchat, or WhatsApp.

Domestic online media is any media source originating in the country in question. For example, the New York Times website is domestic online media in the United States, but not in India, even though it operates bureaus in India. Media includes any source reporting on current events or political issues, ranging from well-established brands to newsletters and websites run by an individual.

Cyber security threats include penetration of private digital networks, using means ranging from exploiting software vulnerabilities, password cracking, or social engineering (e.g., tricking individuals into revealing passwords or other information necessary to break into a digital system) to obtain information or disrupt an organization or individuals use of digital networks and tools. They also include unauthorized alterations of an individual or organizations digital presence, such as defacing websites and commandeering social media accounts. These threats range from unsophisticated (e.g., exploitation of failure to password protect private networks or use of common passwords by authorized users, and spear phishing) to moderate (e.g., embedding malicious code in emails or exploiting well-known software flaws that organizations have failed to patch), to sophisticated (e.g., exploiting unknown exploits in commonly used software or even embedding exploits into commercial systems unbeknownst to their creators).

Clarification: When we discuss shutting down online content, please consider instances where a web-

site (or websites) have been taken entirely offline as well as instances where a website (or websites) have been slowed down or had access similarly intentionally inhibited, such that use of this website is challenging. In other words, both outright shutting down and more subtle measures that inhibit access should be considered when answering these questions.

Clarification: When we discuss censorship or censoring content online, we are not concerned with censorship of topics such as child pornography, highly classified information such as military or intelligence secrets, or defamatory speech, unless this sort of censorship is used as a pretext for censoring political information or opinions.

2.1 Coordinated Information Operations

2.1.1 Government dissemination of false information domestic (C) (v2smgovdom)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How often do the government and its agents use social media to disseminate misleading viewpoints or false information to influence its own population?

Responses:
  0: Extremely often. The government disseminates false information on all key political issues.
  1: Often. The government disseminates false information on many key political issues.
  2: About half the time. The government disseminates false information on some key political issues, but not others.
  3: Rarely. The government disseminates false information on only a few key political issues.
  4: Never, or almost never. The government never disseminates false information on key political issues.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.1.2 Government dissemination of false information abroad (C) (v2smgovab)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Responses:
  0: Extremely often. The government disseminates false information on all key political issues.
  1: Often. The government disseminates false information on many key political issues.
  2: About half the time. The government disseminates false information on some key political issues, but not others.
  3: Rarely. The government disseminates false information on only a few key political issues.
  4: Never, or almost never. The government never disseminates false information on key political issues.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.1.3 Party dissemination of false information domestic (C) (v2smpardom)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Responses:

0: Extremely often. Major political parties and candidates disseminate false information on all key political issues.

1: Often. Major political parties and candidates disseminate false information on many key political issues.

2: About half the time. Major political parties and candidates disseminate false information on some key political issues, but not others.

3: Rarely. Major political parties and candidates disseminate false information on only a few key political issues.

4: Never, or almost never. Major political parties and candidates never disseminate false information on key political issues.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.1.4 Party dissemination of false information abroad (C) (v2smparab)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Responses:

0: Extremely often. Major political parties and candidates disseminate false information on all key political issues.

1: Often. Major political parties and candidates disseminate false information on many key political issues.

2: About half the time. Major political parties and candidates disseminate false information on some key political issues, but not others.

3: Rarely. Major political parties and candidates disseminate false information on only a few key political issues.

4: Never, or almost never. Major political parties and candidates never disseminate false information on key political issues.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.1.5 Foreign governments dissemination of false information (C) (v2smfordom)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Responses:

0: Extremely often. Foreign governments disseminate false information on all key political issues.

1: Often. Foreign governments disseminate false information on many key political issues.

2: About half the time. Foreign governments disseminate false information on some political issues, but not others.

3: Rarely. Foreign governments disseminate false information on only a few key political issues.

4: Never, or almost never. Foreign governments never disseminate false information on key political issues.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.1.6 Foreign governments ads (C) (v2smforads)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Responses:

0: Extremely often. Foreign governments disseminate false information on all key political issues.

1: Often. Foreign governments disseminate false information on many key political issues.

2: About half the time. Foreign governments disseminate false information on some political issues, but not others.

3: Rarely. Foreign governments disseminate false information on only a few key political issues.

4: Never, or almost never. Foreign governments never disseminate false information on key political issues.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2 Digital Media Freedom

2.2.1 Government Internet filtering capacity (C) (v2smgovfilcap)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Responses:

0: The government lacks any capacity to block access to any sites on the Internet.

1: The government has limited capacity to block access to a few sites on the Internet.

2: The government has adequate capacity to block access to most, but not all, specific sites on the Internet if it wanted to.

3: The government has the capacity to block access to any sites on the Internet if it wanted to.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.2 Government Internet filtering in practice (C) (v2smgovfilprc)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How frequently does the government censor political information (text, audio, images, or video) on the Internet by filtering (blocking access to certain websites)?

Responses:
0: Extremely often. It is a regular practice for the government to remove political content, except to sites that are pro-government.
1: Often. The government commonly removes online political content, except sites that are pro-government.
2: Sometimes. The government successfully removes about half of the critical online political content.
3: Rarely. There have been only a few occasions on which the government removed political content.
4: Never, or almost never. The government allows Internet access that is unrestricted, with the exceptions mentioned in the clarifications section.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.3 Government Internet shut down capacity (C) (v2smgovshutcap)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Clarification: A domestic Internet connection is any connection originating physically within the country, whether over wired, wireless, or satellite networks. This question asks what proportion of potential Internet connections of domestic origin the government has the capacity to render inoperable.

Responses:
0: The government lacks the capacity to shut down any domestic Internet connections.
1: The government has the capacity to shut down roughly a quarter of domestic access to the Internet.
2: The government has the capacity to shut down roughly half of domestic access to the Internet.
3: The government has the capacity to shut down roughly three quarters of domestic access to the Internet.
4: The government has the capacity to shut down all, or almost all, domestic access to the Internet.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.4   Government Internet shut down in practice (C) (v2smgovshut)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How often does the government shut down domestic access to the Internet?

Responses:

0: Extremely often. It is a regular practice for the government to shut down domestic access to the Internet.

1: Often. The government shut down domestic access to the Internet numerous times this year.

2: Sometimes. The government shut down domestic access to the Internet several times this year.

3: Rarely but there have been a few occasions throughout the year when the government shut down domestic access to Internet.

4: Never, or almost never. The government does not typically interfere with the domestic access to the Internet.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.5   Government social media shut down in practice (C) (v2smgovsm)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How often does the government shut down access to social media platforms?

Responses:

0: Extremely often. It is a regular practice for the government to shut down access to social media.

1: Often. The government shuts down access to social media numerous times this year.

2: Sometimes. The government shuts down access to social media several times this year.

3: Rarely. There have been a few occasions throughout the year when the government shuts down access to social media.

4: Never, or almost never. The government does not interfere with the access to social media, except in the cases mentioned in the clarifications section.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.6   Government social media alternatives (C) (v2smgovsmalt)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How prevalent is the usage of social media platforms that are wholly controlled by either the government or its agents in this country?

Responses:

0: Essentially all social media usage takes place on platforms controlled by the state.
1: Most usage of social media is on state-controlled platforms, although some groups use non-state-controlled alternatives.
2: There is significant usage of both state-controlled and non-state-controlled social media platforms.
3: While some state-controlled social media platforms exist, their usage only represents a small share of social media usage in the country.
4: Practically no one uses state-controlled social media platforms.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.7 Government social media monitoring (C) (v2smgovsmmon)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How comprehensive is the surveillance of political content in social media by the government or its agents?

Responses:
0: Extremely comprehensive. The government surveils virtually all content on social media.
1: Mostly comprehensive. The government surveils most content on social media, with comprehensive monitoring of most key political issues.
2: Somewhat comprehensive. The government does not universally surveil social media but can be expected to surveil key political issues about half the time.
3: Limited. The government only surveils political content on social media on a limited basis.
4: Not at all, or almost not at all. The government does not surveil political content on social media, with the exceptions mentioned in the clarifications section.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.8 Government social media censorship in practice (C) (v2smgovsmcenprc)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: To what degree does the government censor political content (i.e., deleting or filtering specific posts for political reasons) on social media in practice?

Responses:
0: The government simply blocks all social media platforms.
1: The government successfully censors all social media with political content.
2: The government successfully censors a significant portion of political content on social media, though not all of it.
3: The government only censors social media with political content that deals with especially sensitive issues.
4: The government does not censor political social media content, with the exceptions mentioned in the clarifications section.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.9 Government cyber security capacity (C) (v2smgovcapsec)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Does the government have sufficiently technologically skilled staff and resources to mitigate harm from cyber-security threats?

Responses:
0: No. The government does not have the capacity to counter even unsophisticated cyber security threats.
1: Not really. The government has the resources to combat only unsophisticated cyber attacks.
2: Somewhat. The government has the resources to combat moderately sophisticated cyber attacks.
3: Mostly. The government has the resources to combat most sophisticated cyber attacks.
4: Yes. The government has the resources to combat sophisticated cyber attacks, even those launched by highly skilled actors.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.2.10 Political parties cyber security capacity (C) (v2smpolcap)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Do the major political parties have sufficiently technologically skilled staff and resources to mitigate harm from cyber security threats?

Responses:
0: No. The government does not have the capacity to counter even unsophisticated cyber security threats.
1: Not really. The government has the resources to combat only unsophisticated cyber attacks.
2: Somewhat. The government has the resources to combat moderately sophisticated cyber attacks.
3: Mostly. The government has the resources to combat most sophisticated cyber attacks.
4: Yes. The government has the resources to combat sophisticated cyber attacks, even those launched by highly skilled actors.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.3 State Internet Regulation Capacity and Approach

2.3.1 Internet legal regulation content (C) (v2smregcon)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: What type of content is covered in the legal framework to regulate Internet?

Responses:

　　0: The state can remove any content at will.

　　1: The state can remove most content, and the law protects speech in only specific, and politically uncontroversial contexts.

　　2: The legal framework is ambiguous. The state can remove some politically sensitive content, while other is protected by law.

　　3: The law protects most political speech, but the state can remove especially politically controversial content.

　　4: The law protects political speech, and the state can only remove content if it violates well-established legal criteria.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.3.2   Privacy protection by law exists (C) (v2smprivex)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Does a legal framework to protect Internet users privacy and their data exist?

Responses:

　　0: No. (Skip to v2smregcap)

　　1: Yes

Ordering: if 0 no, Skip to v2smregcap

Scale: yes/no

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (see V-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.3.3   Privacy protection by law content (C) (v2smprivcon)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: What does the legal framework to protect Internet users privacy and their data stipulate?

Responses:

　　0: The legal framework explicitly allows the government to access any type of personal data on the Internet.

　　1: The legal framework explicitly allows the government to access most types of personal data on the Internet.

　　2: The legal framework explicitly allows the government to access many types of personal data on the Internet.

　　3: The legal framework explicitly allows the government to access only a few types of personal information on the Internet.

　　4: The legal framework explicitly allows the government to access personal information on the Internet only in extraordinary circumstances.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.3.4 Government capacity to regulate online content (C) (v2smregcap)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Does the government have sufficient staff and resources to regulate Internet content in accordance with existing law?

Responses:
0: No, almost all online activity happens outside of reach of the state, where it lacks the capacity to remove illegal content.
1: Not really. The state has extremely limited resources to regulate online content.
2: Somewhat. The state has the capacity to regulate only some online content or some portions of the law.
3: Mostly. The state has robust capacity to regulate online content, though not enough to regulate all content and all portions of the law.
4: Yes, the government has sufficient capacity to regulate all online content.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeǎV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.3.5 Government online content regulation approach (C) (v2smregapp)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Responses:
0: All online content monitoring and regulation is done by the state.
1: Most online content monitoring and regulation is done by the state, though the state involves private actors in a limited way.
2: Some online content monitoring and regulation is done by the state, but the state also involves private actors in monitoring and regulation in various ways.
3: The state does little online content monitoring and regulation, and entrusts most of the monitoring and regulation to private actors.
4: The state off-loads all online content monitoring and regulation to private actors.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeǎV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.3.6 Defamation protection (C) (v2smlawpr)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Does the legal framework provide protection against defamatory online content, or hate speech?

Responses:

    0: No. The law provides no protection against Internet defamation and hate speech.

    1: Not really. The law provides a weak protection and to very limited range of circumstances.

    2: Somewhat. The law provides some protection against Internet defamation and hate speech but in limited circumstances, or only to particular groups of people.

    3: Mostly. The law provides protection against Internet defamation and hate speech under many circumstances, and to most groups of people.

    4: Yes. The law provides comprehensive protection against Internet defamation and hate speech.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.3.7   Abuse of defamation and copyright law by elites (C) (v2smdefabu)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: To what extent do elites abuse the legal system (e.g., defamation and copyright law) to censor political speech online?

Responses:

    0: Regularly. Elites abuse the legal system to remove political speech from the Internet as regular practice.

    1: Often. Elites commonly abuse the legal system to remove political speech from the Internet.

    2: Sometimes. Elites abuse the legal system to remove political speech from the Internet about half the time.

    3: Rarely. Elites occasionally abuse the legal system to remove political speech from the Internet.

    4: Never, or almost never. Elites do not abuse the legal system to remove political speech from the Internet.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.4   Online Media Polarization

2.4.1   Online media existence (C) (v2smonex)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Do people consume domestic online media?

Responses:

    0: Not at all. No one consumes domestic online media. Skip next question if this answer is selected.

    1: Limited. Domestic online media consumption is limited.

    2: Relatively extensive. Domestic online media consumption is common.

    3: Extensive. Almost everyone consumes domestic online media.

Ordering: if 0, skip v2smonper

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.4.2 Online media perspectives (C) (v2smonper)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Do the major domestic online media outlets represent a wide range of political perspectives?

Responses:
0: The major domestic online media outlets represent only the governments perspective.
1: The major domestic online media outlets represent only the perspectives of the government and a government approved, semi-official opposition party.
2: The major domestic online media outlets represent a variety of political perspectives but they systematically ignore at least one political perspective that is important in this society.
3: All perspectives that are important in this society are represented in at least one of the major domestic online media outlets.
4: All perspectives that are important in this society are represented in many major domestic online media outlets.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.4.3 Online media fractionalization (C) (v2smmefra)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Do the major domestic online media outlets give a similar presentation of major (political) news?

Responses:
0: No. The major domestic online media outlets give opposing presentation of major events.
1: Not really. The major domestic online media outlets differ greatly in the presentation of major events.
2: Sometimes. The major domestic online media outlets give a similar presentation of major events about half the time.
3: Mostly. The major domestic online media outlets mostly give a similar presentation of major events.
4: Yes. Although there are small differences in representation, the major domestic online media outlets give a similar presentation of major events.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5 Social Cleavages

2.5.1 Online harassment groups (C) (v2smhargr)

Additional versions: *_nr

Question: Which groups are targets of hate speech or harassment in online media?

Clarification: Multiple selection. Choose all that apply.

Responses:
    0: Women [v2smhargr_0]
    1: LGBTQ groups and individuals [v2smhargr_1]
    2: Specific religious groups [v2smhargr_2]
    3: Specific ethnic groups [v2smhargr_3]
    4: Specific caste [v2smhargr_4]
    5: Specific language groups [v2smhargr_5]
    6: Specific race [v2smhargr_6]
    7: People with physical or cognitive disabilities [v2smhargr_7]
    8: People from specific regions [v2smhargr_8]
    9: Other (specify in the next question) [v2smhargr_9]
    10: No group is a specific target [v2smhargr_10]

Scale: Series of dichotomous scales.

Data release: 9.

Cross-coder aggregation: Mean.

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5.2 Other online harassment groups (C) (v2smhargrtxt)

Question: Which other groups are targets of hate speech or harassment in online media?

Clarification: Skip if the question does not apply to this country.

Scale: Text.

Data release: 9. Available upon request, subject to review and approval.

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87)

Years: 2000-2018

2.5.3 Use of social media to organize offline violence (C) (v2smorgviol)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How often do people use social media to organize offline violence?

Responses:
    0: Frequently. There are numerous cases in which people have used social media to organize offline violence.
    1: Sometimes. There are a few cases in which people have used social media to organize offline violence.
    2: Never. People have never used social media to organize offline violence.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5.4 Average peoples use of social media to organize offline action (C) (v2smorgavgact)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How often do average people use social media to organize offline political action of any kind?

Responses:

0: Never or almost never. Average people have almost never used social media to organize offline political action.

1: Rarely. Average people do not typically use social media to organize offline political action.

2: Sometimes. There are a few cases in which average people have used social media to organize offline political action.

3: Often. There have been several cases in which average people have used social media to organize offline political action.

4: Regularly. There are numerous cases in which average people have used social media to organize offline political action.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5.5 Elites use of social media to organize offline action (C) (v2smorgelitact)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How often do domestic elites use social media to organize offline political action of any kind?

Responses:

0: Never or almost never. Elites have almost never used social media to organize offline political action.

1: Rarely. Elites do not typically use social media to organize offline political action.

2: Sometimes. There are a few cases in which elites have used social media to organize offline political action.

3: Often. There have been several cases in which elites have used social media to organize offline political action.

4: Regularly. There are numerous cases in which elites have used social media to organize offline political action.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5.6 Types of organization through social media (C) (v2smorgtypes)

Additional versions: *_nr

Question: What types of offline political action are most commonly mobilized on social media?

Clarification: Multiple selection. Choose all that apply.

Responses:

0: Petition signing [v2smorgtypes_0]

1: Voter turnout [v2smorgtypes_1]

2: Street protests [v2smorgtypes_2]

3: Strikes/labor actions [v2smorgtypes_3]
4: Riots [v2smorgtypes_4]
5: Organized rebellion [v2smorgtypes_5]
6: Vigilante Justice (e.g., mob lynching, stalking harassment) [v2smorgtypes_6]
7: Terrorism [v2smorgtypes_7]
8: Ethnic cleansing/genocide [v2smorgtypes_8]
9: Other (specify in the next question) [v2smorgtypes_9]

Scale: Series of dichotomous scales.

Data release: 9.

Cross-coder aggregation: Mean.

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5.7 Other types of organization through social media (C) (v2smorgtypestxt)

Question: What other types of offline political action are most commonly mobilized on social media?

Clarification: Skip if the question does not apply to this country.

Scale: Text.

Data release: 9. Available upon request, subject to review and approval.

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87)

Years: 2000-2018

2.5.8 Party/candidate use of social media in campaigns (C) (v2smcamp)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: To what extent do major political parties and candidates use social media during electoral campaigns to communicate with constituents?

Responses:
0: None. Major political parties and candidates do not use social media during electoral campaigns to communicate with constituents.
1: A little. Major political parties and candidates rarely use social media during electoral campaigns to communicate with constituents.
2: Somewhat. Major political parties and candidates sometimes use social media during electoral campaigns to communicate with constituents.
3: Substantial. Major political parties and candidates frequently use social media during electoral campaigns to communicate with constituents.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5.9 Arrests for political content (C) (v2smarrest)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: 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?

Responses:
0: Extremely likely.
1: Likely.

2: Unlikely.

3: Extremely unlikely.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5.10   Polarization of society (C) (v2smpolsoc)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How would you characterize the differences of opinions on major political issues in this society?

Clarification: While plurality of views exists in all societies, we are interested in knowing the extent to which these differences in opinions result in major clashes of views and polarization or, alternatively, whether there is general agreement on the general direction this society should develop.

Responses:

0: Serious polarization. There are serious differences in opinions in society on almost all key political issues, which result in major clashes of views.

1: Moderate polarization. There are differences in opinions in society on many key political issues, which result in moderate clashes of views.

2: Medium polarization. Differences in opinions are noticeable on about half of the key political issues, resulting in some clashes of views.

3: Limited polarization. There are differences in opinions on only a few key political issues, resulting in few clashes of views.

4: No polarization. There are differences in opinions but there is a general agreement on the direction for key political issues.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Methodology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

2.5.11   Political parties hate speech (C) (v2smpolhate)

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: How often do major political parties use hate speech as part of their rhetoric?

Clarification: Hate speech is any speech that is intended to insult, offend, or intimidate members of specific groups, defined by race, religion, sexual orientation, national origin, disability, or similar trait.

Responses:

0: Extremely often.

1: Often.

2: Sometimes.

3: Rarely.

4: Never, or almost never.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 9.

Cross-coder aggregation: Bayesian item response theory measurement model (seeăV-Dem Method-

ology).

Citation: Mechkova *et al.* (2019, *V-Dem Working Paper Series* 2019:87); Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 2000-2018

3 Appendix B: Glossary

Attributes: This section includes the most specific conceptual building blocks we use to discuss democracy and related concepts. Many of our survey questions attempt to ask about a single attribute, for example, "What percentage of the lower (or unicameral) chamber of the legislature is directly elected in popular elections?" Although any of these questions could also be seen as a compendium of multiple attributes (What does it mean to be a legislature? What is a "popular" election?), in a project covering all countries for more than a century, there are degrees of specificity that it is not practical to approach, so attributes are the most specific concepts that we consider feasible to measure.

Country: A sovereign state or semi-sovereign territory. All political units of concern to V-Dem are referred to as countries, even though their status in international law varies — some being colonies others being nation-states or empires.

Cyber Security Threats: Cyber security threats include penetration of private digital networks, using means ranging from exploiting software vulnerabilities, password cracking, or social engineering (*e.g.*, tricking individuals into revealing passwords or other information necessary to break into a digital system) to obtain information or disrupt an organization or individuals use of digital networks and tools. They also include unauthorized alterations of an individual or organizations digital presence, such as defacing websites and commandeering social media accounts. These threats range from unsophisticated (*e.g.*, exploitation of failure to password protect private networks or use of common passwords by authorized users, and spear phishing) to moderate (*e.g.*, embedding malicious code in emails or exploiting well-known software flaws that organizations have failed to patch), to sophisticated (*e.g.*, exploiting unknown exploits in commonly used software or even embedding exploits into commercial systems unbeknownst to their creators).

Domestic Online Media: Domestic online media is any media source originating in the country in question. For example, the New York Times website is domestic online media in the United States, but not in India, even though it operates bureaus in India. Media includes any source reporting on current events or political issues, ranging from well-established brands to newsletters and websites run by an individual.

Geographic Group: Geographic group refers to those living in rural or urban areas. Urban areas are defined as an area that meets the following conditions: population density exceeds a threshold of 150 persons per square kilometer, there is access to a sizeable settlement of 50,000 people or more within some reasonable travel time, for example 60 minutes by road. (World Development Report, 2009: 54).

Government: The executive branch of the government, including its head of state (HOS) and/or head of government (HOG) — whichever is most prominent, or both if they are both powerful — along with the cabinet, ministries, and top civil servants. We are only concerned here with the government that actually resides within the country or semi-sovereign territory. Thus, in a typical British colony the government would include the governor-general and his local administration but not the King/Queen of England or the government of England.

Government and its Agents: The government and its agents include official government organs, such as bureaucracies, courts, intelligence services, and the military, but also unofficial agents, such as officially unaffiliated cyber-warfare operatives who perform services, even "off-book" work, on behalf of the government.

Internet: We define the Internet as all information that people access over public and private digital networks, worldwide. The Internet includes both publicly accessible digital spaces and private or gated information transmission platforms. The Internet does not include traditional media transmission mechanisms such as paper, television, traditional voice telephone, and radio.

Major Political Parties: Major political parties include the group of political parties that hold a significant number of seats in national legislative body(-ies), or earn a significant number of votes in elections for the executive. When we ask you to consider "major political parties", you do not need to consider parties that run in elections but receive only a small minority of seats or votes, or those that receive no seats at all.

National Government: The highest level of aggregation recognized by the V-Dem project. Refers to the national government of a sovereign state or the territorial level of government for a semi-sovereign colony or territory. Thus, the "national" government of India prior to independence — the British Raj — was situated in New Delhi, *not* in London — even though decisions affecting the Indian colony were often made in London.

Political Groups: Political groups are defined as those who are affiliated with a particular political party or candidate, or a group of parties/candidates. A common form of partisan exclusion is when state services or regulations are implemented in a way that seeks to reward incumbent political supporters and punish non-supporters.

Political Party: An organization that nominates candidates for public office. The term includes a longstanding coalition such as the CDU/CSU in Germany if that coalition functions in most respects like a single party. Sometimes, the identity of a party is obscured by name changes. However, if the party changes names but retains key personnel and is still run by and for the same constituencies then it should be considered the same organization.

Public Authorities: Includes the government as well as subnational governments, agencies, parastatals, and the like. Compare State.

Semisovereign Territory: This refers to a country that is not fully sovereign but nonetheless exercises some — at least minimal — level of self-determination. Many of the countries of concern to this project began as colonies of an empire. If a country moved from semi-sovereign status to sovereign status over the course of the twentieth century — maintaining comparable borders — then we want to code both entities. Likewise, we want to include countries like Taiwan that are not universally recognized as sovereign but nonetheless enjoy self-determination (in part or in full).

Most questions pertaining to semi-sovereign territories ask you to reflect on the practices and institutions located within that territory — rather than the empire or nation-state that may claim ultimate sovereignty over the territory. Thus, a question about the government or judicial bodies seated within a British colony would refer to the governor-general and his local administration rather than the King/Queen or government of England.

Social Group: A social group is differentiated within a country by caste, ethnicity, language, race, region, religion, migration status, or some combination thereof. (It does *not* include identities grounded in sexual orientation, gender, or socioeconomic status.) Social group identity is contextually defined and is likely to vary across countries and through time. Social group identities are also likely to cross-cut, so that a given person could be defined in multiple ways, *i.e.*, as part of multiple groups. Nonetheless, at any given point in time there are social groups within a society that are understood — by those residing within that society — to be different, in ways that may be politically relevant. Contrast Identity group.

Social Media: Social media are a subset of Internet platforms that enable normal individuals to create and share content with networks of other people. Social media platforms are available to the public, although content on such networks may be shared privately within subgroups of users. Social media includes both publicly visible, or semi-public platforms, like Facebook, Flickr, Friendster, Google+, Instagram, Myspace, LinkedIn, Twitter, VKontakte, and Weibo and private social networking and messaging platforms like Signal, Slack, Snapchat, or WhatsApp.

State: A political organization that organizes compulsory domination over a fixed territory on a continual basis.

Variable: A measure of a small number of attributes. Synonymous with "indicator."