Measuring Media Freedom: An Item Response Theory Analysis of Existing Indicators

Supplemental Appendix

Jonathan A. Solis
College of William & Mary

Philip D. Waggoner
University of Chicago

January 30, 2020

Contents

List of Tables

List of Figures

A General Media Freedom Indicators

In this section we provide a more detailed description of the indicators we use in the IRT model.

A.1 Freedom House

Freedom House's (2017) *Freedom of the Press* index remains one of the most widely used datasets for measuring press freedom in political science and economics (Brunetti and Weder, 2003; Egorov, Guriev and Sonin, 2009; Schoonvelde, 2014; Kellam and Stein, 2016). The non-governmental organization began issuing reports in 1980, covering the previous year 1979, and continues to the present. However, the index's managers have changed the dataset's methodology numerous times since its inception. From 1980 until 1988, Freedom House provided separate rankings for a country's print and broadcast sectors, ranking each either *free*, *partially free*, or *not free*. Freedom House provides no aggregate ranking of the country's media system during this period. From 1989 until the present, Freedom House assigns countries a sector aggregated, countrywide *free*, *partially free*, or *not free* ranking. Freedom House does not make the index's survey methodology or criteria readily available from 1980 to 1993.

From 1994 until the present, Freedom House introduced a continuous score ranging from 0 to 100 to accompany each country's categorical ranking, with 0 being the most free and 100 being the least. Initially under the 100-point scale, the index evaluated media freedom based on four criteria: 1) law, 2) political pressure, 3) economic influence, and 4) repressive actions. The index managers considered both print and broadcast sectors separately and then assigned an overall freedom score. From 1997 until 2001, they used the same basic structure but modified the point distribution. In 2002 Freedom House introduced a new coding scheme that it still employs today. The newest methodology evaluates three different areas of each country's media system: the legal, political, and economic environments.

The legal environment category evaluates a country's laws and regulations regarding media and the government's inclination to use them to restrict the press falling on a scale from 0 to 30 with 0

being the most free and 30 being the least. The political environment category measures the degree of political control over the media by state and non-state actors including intimidation, detention, imprisonment, and violent assault on a range from 0 to 40 with 0 being the most free and 40 being the least.

Finally, the economic environment category assesses the structure, concentration, and transparency of media ownership including the government's distribution of advertisement, bribery, and the cost of establishing media among other criteria. Like the legal environment category, Freedom House utilizes a 0 to 30 range. These three combined scores determine a country's press freedom level. Countries that score from 0 to 30 are considered free, those from 31 to 60 rank as partially free, and those above 60 receive the label not free.[1]

For Freedom House's data before 1988, we take the average of the print and broadcast sectors to generate a rank. We score a country-year partly free (2) for any average greater than one or less than three. To illustrate this coding scheme, we provide a number of different examples. Australia in 1979 has two 3 scores for each sector, which makes their overall score 3: ((3 + 3)/2 = 3). Brazil's overall score in 1980 would be 2: ((2 + 2)/2 = 2). Equatorial Guinea 's overall score in 1981 would be 1: ((1 + 1)/2 = 1). All other print and broadcast sector combinations results in a 2 (partially free) coding. For instance, Freedom House gives Argentina a 2 rank for the print sector and a 1 rank for the broadcast sector. This average equals 1.5, so we assign it a 2 (partially free) in accordance to our criterion. Jamaica in 1987 has a score of 3 and 2 in the print and broadcast sectors respectively. This average equals 2.5, so we assign it a 2 (partially free). In rare cases where one sector is 3 and the other is 1 (like Malta in 1986), we again assign a 2 (partially free) ranking ((3 + 1)/2 = 2).

A.2 Global Media Freedom

The Global Media Freedom Dataset (Whitten-Woodring and Van Belle, 2017) orients its data collection methodology by first defining media freedom as "an environment in which journalists are able to safely criticize political and economic elites at both the national and local levels."

[1] See Freedom House's website for more information: https://freedomhouse.org/report-types/freedom-press.

(Whitten-Woodring and Van Belle, 2017, 180).[2] The authors use simple, ordinal categories of free, imperfectly free, and not free, and orient their data collection methodology by first defining media freedom as "an environment in which journalists are able to safely criticize political and economic elites at both the national and local levels." (Whitten-Woodring and Van Belle, 2017, 180). From this starting point, they employ a simple coding system based on three categories obtained from historical documents about each country's media:

- **Free**—Countries where criticism of government and government officials is a common and normal part of the political dialogue in the mediated public sphere.
- **Imperfectly Free**—Countries where social, legal, or economic costs related to the criticism of government or government officials limits public criticism, but investigative journalism and criticism of major policy failings can and does occur.
- **Not Free**—Countries where it is not possible to safely criticize government or government officials.

The Global Media Freedom (GMF) dataset provides media freedom measures for 205 countries from 1948 to 2014.

A.3 Press Freedom Index (Reporters Without Borders)

The French-based watchdog group Reporters Without Borders (RSF) has released its Press Freedom Index since 2002 (Reporters Without Borders, 2018).[3] The index provides press freedom scores and country rankings based on surveys from journalists, scholars, and human rights activists.[4] RSF focuses heavily on harassment against media, attacks against journalists, and self-censorship, though they consider other criteria such as economic and legal conditions. Generally, scores range from 0 to 100, with 0 representing perfect press freedom and 100 indicating the least perfect, though some years inexplicably possess negative scores or scores above 100. Unfortunately, RSF does not publish the survey questionnaires or methodology it employs for each year (Schneider, 2014).

[2] See http://faculty.uml.edu/Jenifer_whittenwoodring/MediaFreedomData_000.aspx for more information.
[3] See https://rsf.org/en/world-press-freedom-index for more information.
[4] Most often used as a robustness check. See Freille, Haque and Kneller (2007), Egorov, Guriev and Sonin (2009), and Stier (2015).

Next, we describe how we generate an ordinal variable for RSF's data. The RSF score ranges from -14 to 144, and we convert it to four categories, dividing the data by 25th percentiles. We then code the categories so that higher numbered categories indicate greater media freedom.

A.4 Varieties of Democracy Indicators

The Varieties of Democracy (V-Dem) dataset provides original data for a number of indicators often associated with democracy (Coppedge et al., 2018a).[5] In their *Media* section, they offer a number of variables that evaluate different aspects of media quality and media freedom. V-Dem generates scores by asking country experts to rank each country according to a specified variable concept. After receiving the responses, the V-Dem researchers run the results through an item response theory (IRT) model to compile a cross-coder aggregated score (Pemstein et al., 2018). Below we report the question and scores/rankings for each V-Dem variable we use in this analysis including government censorship of traditional media (the press, television, and radio), government internet censorship, presence of a critical media, presence of various perspectives in media, harassment of journalists, the prevalence of self-censorship, media bias, media corruption, and access to media critical of the government.

A.4.1 Government Censorship

Government Censorship measures government censorship efforts against traditional media outlets like the press, television, and radio. The authors clarify that this includes indirect means of censorship such as politically motivated financial and official (awarding broadcast frequencies) support of supportive media outlets and other restrictions including a high barrier to receive a broadcasting license or taxes. The survey asked experts the following (Coppedge et al., 2018c, 181):

> *Does the government directly or indirectly attempt to censor the print or broadcast media?*

[5] See https://www.v-dem.net/en/data/data-version-8/.

From here, V-Dem offers five possible responses (Coppedge et al., 2018c, 181):

0. Attempts to censor are direct and routine.
1. Attempts to censor are indirect but nevertheless routine.
2. Attempts to censor are direct but limited to especially sensitive issues.
3. Attempts to censor are indirect and limited to especially sensitive issues.
4. The government rarely attempts to censor major media in any way, and when such exceptional attempts are discovered, the responsible officials are usually punished.

A.4.2 Critical Media

Critical Print and Broadcast Media measures the degree to which major media outlets criticize the government. The survey asked experts the following (Coppedge et al., 2018c, 182):

> *Of the major print and broadcast outlets, how many routinely criticize the government?*

From here, V-Dem offers four possible responses (Coppedge et al., 2018c, 182):

0. None.
1. Only a few marginal outlets.
2. Some important outlets routinely criticize the government but there are other important outlets that never do.
3. All major media outlets criticize the government at least occasionally.

A.4.3 Print and Broadcast Media Perspectives

Print and Broadcast Media Perspectives measures the degree to which major media outlets report different perspectives. The lowest score represents a situation where media only report the government's perspective, while the highest score scores situations where all society's important perspectives are present in the media. The survey asked experts the following (Coppedge et al., 2018c, 182):

> *Do the major print and broadcast media represent a wide range of political perspectives?*

From here, V-Dem offers four possible responses (Coppedge et al., 2018c, 182-83):

0. The major media represent only the government's perspective.
1. The major media represent only the perspectives of the government and a government-approved, semi-official opposition party.
2. The major media 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 media.

A.4.4 Harassment of Journalists

Harassment of Journalist measures the degree to which journalists face harassment from governments and powerful non-governmental organizations. The survey asks experts the following basic question (Coppedge et al., 2018c, 183):

> *Are individual journalists harassed - i.e., threatened with libel, arrested, imprisoned, beaten, or killed – by governmental or powerful nongovernmental actors while engaged in legitimate journalistic activities?*

From here, V-Dem offers five possible responses (Coppedge et al., 2018c, 183):

0. No journalists dare to engage in journalistic activities that would offend powerful actors because harassment or worse would be certain to occur.
1. Some journalists occasionally offend powerful actors but they are almost always harassed or worse and eventually are forced to stop.
2. Some journalists who offend powerful actors are forced to stop but others manage to continue practicing journalism freely for long periods of time.
3. It is rare for any journalist to be harassed for offending powerful actors, and if this were to happen, those responsible for the harassment would be identified and punished.
4. Journalists are never harassed by governmental or powerful nongovernmental actors while engaged in legitimate journalistic activities.

A.4.5 Media Self-censorship

Media Self-censorship measures self-censorship's prevalence among journalists in a country. The survey asked experts the following (Coppedge et al., 2018c, 183):

> *Is there self-censorship among journalists when reporting on issues that the government considers politically sensitive?*

From here, V-Dem offers four possible responses (Coppedge et al., 2018c, 184):

0. Self-censorship is complete and thorough.
1. Self-censorship is common but incomplete.
2. There is self-censorship on a few highly sensitive political issues but not on moderately sensitive issues.
3. There is little or no self-censorship among journalists.

A.4.6 Media Bias

This variable measures the degree to which media as a whole holds overt biases against opposition parties or candidates. The survey asked experts the following (Coppedge et al., 2018c, 184):

> *Is there media bias against opposition parties or candidates?*

The survey then adds the following clarification (Coppedge et al., 2018c, 184):

> We ask you to take particular care in rating the year-to-year variation on this question if media bias tends to increase or decrease in election years. Coverage can be considered "more or less impartial" when the media as a whole present a mix of positive and negative coverage of each party or candidate.

From here, V-Dem offers four possible responses (Coppedge et al., 2018c, 184):

0. The print and broadcast media cover only the official party or candidates, or have no political coverage, or there are no opposition parties or candidates to cover.
1. The print and broadcast media cover more than just the official party or candidates but all the opposition parties or candidates receive only negative coverage.
2. The print and broadcast media cover some opposition parties or candidates more or less impartially, but they give only negative or no coverage to at least one newsworthy party or candidate.
3. The print and broadcast media cover opposition parties or candidates more or less impartially, but they give an exaggerated amount of coverage to the governing party or candidates.
4. The print and broadcast media cover all newsworthy parties and candidates more or less impartially and in proportion to their newsworthiness.

A.4.7 Media Corruption

This variable measures the level of corruption among journalists and media personnel in a country's media system. The survey asked experts the following (Coppedge et al., 2018c, 185):

Do journalists, publishers, or broadcasters accept payments in exchange for altering news coverage?

From here, V-Dem offers five possible responses (Coppedge et al., 2018c, 185):

0. The media are so closely directed by the government that any such payments would be either unnecessary to ensure pro-government coverage or ineffective in producing anti-government coverage.
1. Journalists, publishers, and broadcasters routinely alter news coverage in exchange for payments.
2. It is common, but not routine, for journalists, publishers, and broadcasters to alter news coverage in exchange for payments.
3. It is not normal for journalists, publishers, and broadcasters to alter news coverage in exchange for payments, but it happens occasionally, without anyone being punished.
4. Journalists, publishers, and broadcasters rarely alter news coverage in exchange for payments, and if it becomes known, someone is punished for it.

B Pairwise Correlation Matrix w/ Significance Levels

Table 1: Pairwise Correlation Matrix; 10 Media Freedom Variables

	GMF	Freedom House	RSF	Trad M.Cen	M.Critical	M.Perspectives	M.Harassment	M.Self-Cen	M.Bias	M.Corrupt
GMF	1									
	(10,080)									
Freedom House	0.8418*	1								
	(6,377)	(6,781)								
RSF	0.8307*	0.8364*	1							
	(2,181)	(2,545)	(2,730)							
V-Dem: Trad M.Cen	0.7481*	0.7477*	0.7403*	1						
	(9,341)	(6,153)	(2,627)	(9,938)						
V-Dem: M.Critical	0.6791*	0.6952*	0.6206*	0.8192*	1					
	(9,341)	(6,153)	(2,627)	(9,938)	(9,938)					
V-Dem: M.Perspectives	0.6841*	0.6819*	0.6273*	0.8342*	0.8644*	1				
	(9,341)	(6,153)	(2,627)	(9,938)	(9,938)	(9,938)				
V-Dem: M.Harassment	0.7341*	0.7150*	0.7458*	0.8245*	0.7955*	0.7910*	1			
	(9,341)	(6,153)	(2,627)	(9,938)	(9,938)	(9,938)	(9,938)			
V-Dem: M.Self-Cen	0.6832*	0.7042*	0.6847*	0.8383*	0.8417*	0.8203*	0.8114*	1		
	(9,341)	(6,153)	(2,627)	(9,938)	(9,938)	(9,938)	(9,938)	(9,938)		
V-Dem: M.Bias	0.7150*	0.7004*	0.6571*	0.8472*	0.8574*	0.8798*	0.8022*	0.8417*	1	
	(9,341)	(6,153)	(2,627)	(9,938)	(9,938)	(9,938)	(9,938)	(9,938)	(9,938)	
V-Dem: M.Corrupt	0.7130*	0.7093*	0.6737*	0.8073*	0.8119*	0.8134*	0.7978*	0.8053*	0.8458*	1
	(9,341)	(6,153)	(2,627)	(9,938)	(9,938)	(9,938)	(9,938)	(9,938)	(9,938)	(9,938)

$*p < 0.01$; Observations in parentheses

C Posterior Prediction Distribution

Figure 1 in the manuscript shows the predicted distribution to the actual distribution of y_{rky} to assess how well the measurement model predicts each indicator's manifest distribution. To provide further detail for this figure, we present the raw numbers that make-up that figure in table 2 below. We highlight the two categories that fall outside the predicted distribution: 1) the first category of V-Dem's Range of Perspective variable, and 2) the fourth category of the V-Dem media corruption variable.

Table 2: Observed Ratings and The Posterior Predictive Distribution, Raw Numbers

GMF	1	2	3		
97.50%	5207	2424	2726		
actual	5172	2248	2656		
2.50%	5034	2216	2558		
misses: 0					
FH: FoP	1	2	3		
97.50%	2419	2262	2319		
actual	2397	2107	2270		
2.50%	2277	2087	2175		
misses: 0					
RSF	1	2	3	4	
97.50%	694	735	791	712	
actual	678	684	680	681	
2.50%	601	623	674	611	
misses: 0					
VDem-Trad M.Cen	1	2	3	4	5
97.50%	3091	1297	1468	2717	1774
actual	3081	1191	1314	2616	1729
2.50%	2933	1135	1283	2513	1639
misses: 0					
VDem-Critical.M	1	2	3	4	
97.50%	1705	2233	3404	2919	
actual	1682	2063	3321	2865	
2.50%	1570	2051	3213	2771	
misses: 0					
VDem-M.Perspective	1	2	3	4	
97.50%	2504	1639	2275	3840	
actual	2504	1458	2141	3828	
2.50%	2365	1471	2086	3686	
misses: 1					
VDem-M.Harassment	1	2	3	4	5
97.50%	1550	2806	3104	2243	652
actual	1491	2697	3007	2109	627
2.50%	1397	2610	2893	2068	540
misses: 0					
VDem-M.Selfcen	1	2	3	4	
97.50%	2147	2242	3994	1886	
actual	2112	2064	3923	1832	
2.50%	2000	2049	3799	1747	
misses: 0					
VDem-M.Bias	1	2	3	4	5
97.50%	2411	1309	1223	2784	2607
actual	2401	1176	1082	2700	2572
2.50%	2265	1148	1068	2587	2458
misses: 0					
VDem-M.Corruption	1	2	3	4	5
97.50%	2455	1335	1756	2703	2124
actual	2386	1125	1728	2631	2061
2.50%	2299	1162	1579	2501	1973
misses: 1					

D Test IRT's Single-Trait Assumption

Item response theory (IRT) models assume a single latent trait among the set of indicators it analyzes. To provide evidence for this assumption, we run factor analytic (FA) and principal component analytic (PCA) models on the 10 media freedom variables that we include in the measurement model. For FA, we expect to find evidence that the data contain a single factor, while we expect to find a single component for PCA. The results presented below confirm these expectations. However, due to pair-wise deletion inherent in FA and PCA processes, we opt to also fit Bayesian factor analytic (BFA) models. The Bayesian structure does not drop observation through pair-wise deletion. Like the results from the FA and PCA models, we find evidence for a single dimension using BFA.

D.1 Factor Analytic Models

The factor analytic (FA) model is a latent trait analysis that "identifies the relationships among measured variables for the purposes of reducing data, such as collapsing several items on a test into subscales, and/or evaluating theoretical structures" (Brown et al., 2011, 141). Here, we wish to identify the data's structure. FA identifies inferred latent variables known as factors (Brown et al., 2011, 141). In accordance with the IRT model's assumption, we expect the data to contain a single factor. To this end, we run a FA model and examine each factor's eigenvalues. Eigenvalues over 1 indicate a single factor (Kaiser, 1958). We therefore expect to find only one factor with an eigenvalue over 1.

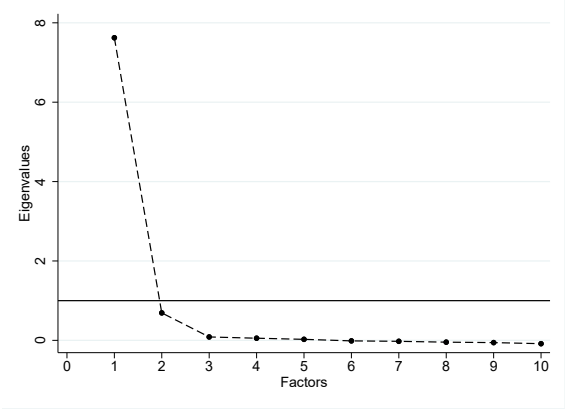
Figure 1a shows scree plots of the FA results with all 10 variables. The results indicate the first factor with a eigenvalue of about 7.62 with all other factors being less than 1. As expected, the results indicate a single factor. However, FA requires the presence of all variable rows and discards observations that do not meet this criterion. As a result, pairwise deletion greatly diminishes the sample size. We therefore repeat the analysis and remove variables with less coverage to expand the sample size. First, we remove *RSF*, which runs from 2002 to 2017, and find the same result

(figure 1b). We then remove *RSF* and *Freedom House*, which runs from 1979 to 2016, and we again find the expected result (figure 1c). In sum, using factor analysis, we find evidence that these variables contain a single factor, which supports the IRT model's single trait assumption.

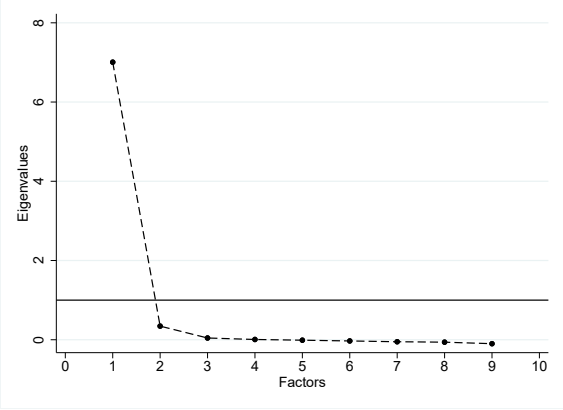
D.1.1 Scree Plots

Figure 1: Factor Analysis Scree Plots, Different Specifications

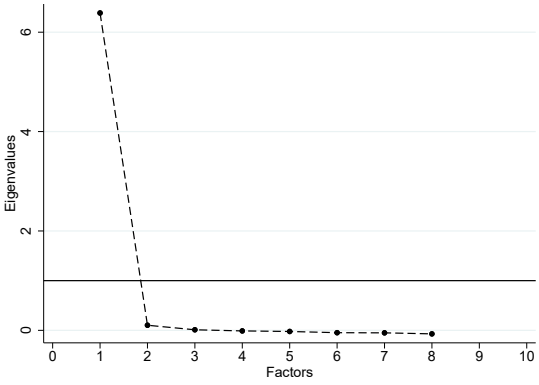
(a) All 10 Indicators (n=2,119)



(b) Without RSF (n=5,801)



(c) Without RSF and Freedom House (n=9,341)



D.2 Principal Component Analytic Models

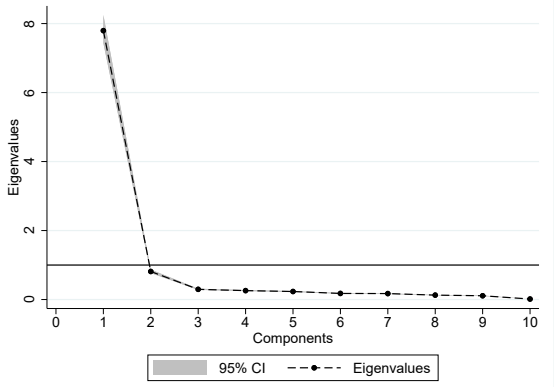
The principal component analytic (PCA) model is a data reduction technique that linearly transforms intercorrelated variables into smaller sets of uncorrelated variables that contain most of the original dataset's information (Dunteman, 1989, 7). Researchers can use PCA to reduce multicollinearity among highly correlated variable or examine data structure. We aim to do the latter, and we expect the data to contain a single component, in accordance with the IRT model's single-trait assumption. To this end, we run a PCA and examine each component's eigenvalues. Eigenvalues over 1 indicate a single component (Kaiser, 1958), and we therefore expect to find only one component with an eigenvalue over 1.

Figure 2a shows scree plots with 95 percent confidence intervals of the PCA results with all 10 variables. The results show the first component with a eigenvalue of about 7.79 with all other components falling below the 1 criterion. As expected, these results indicate a single component. However, PCA requires the presence of all variable rows and discards observations that do not meet this criterion. As a result, pairwise deletion greatly diminishes the sample size. We therefore repeat the analysis and remove variables with less coverage to expand the sample size. First, we remove *RSF*, which runs from 2002 to 2017, and find the same result (figure 2b). We then remove *RSF* and *Freedom House*, which runs from 1979 to 2016, and we again find the expected result (figure 2c). In sum, using PCA, we find evidence that these variables contain a single component, which supports the IRT model's single trait assumption.

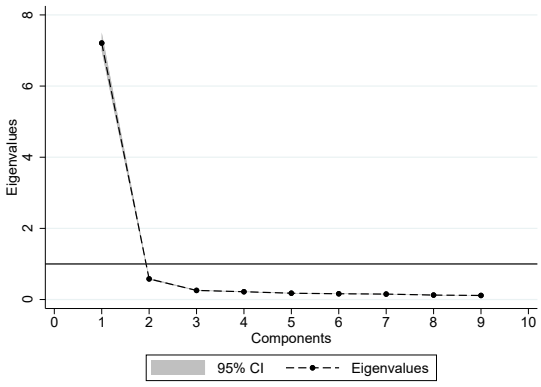
D.2.1 Scree Plots

Figure 2: PCA Scree Plots, Different Specifications

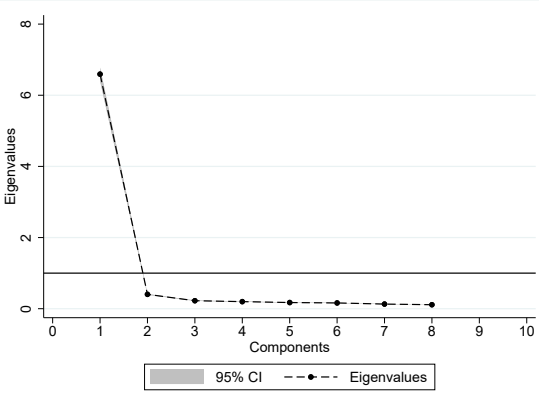
(a) All 10 Indicators (n=2,119)



(b) Without RSF (n=5,801)



(c) Without RSF and Freedom House (n=9,341)



D.3 Bayesian Factor Analytic Models

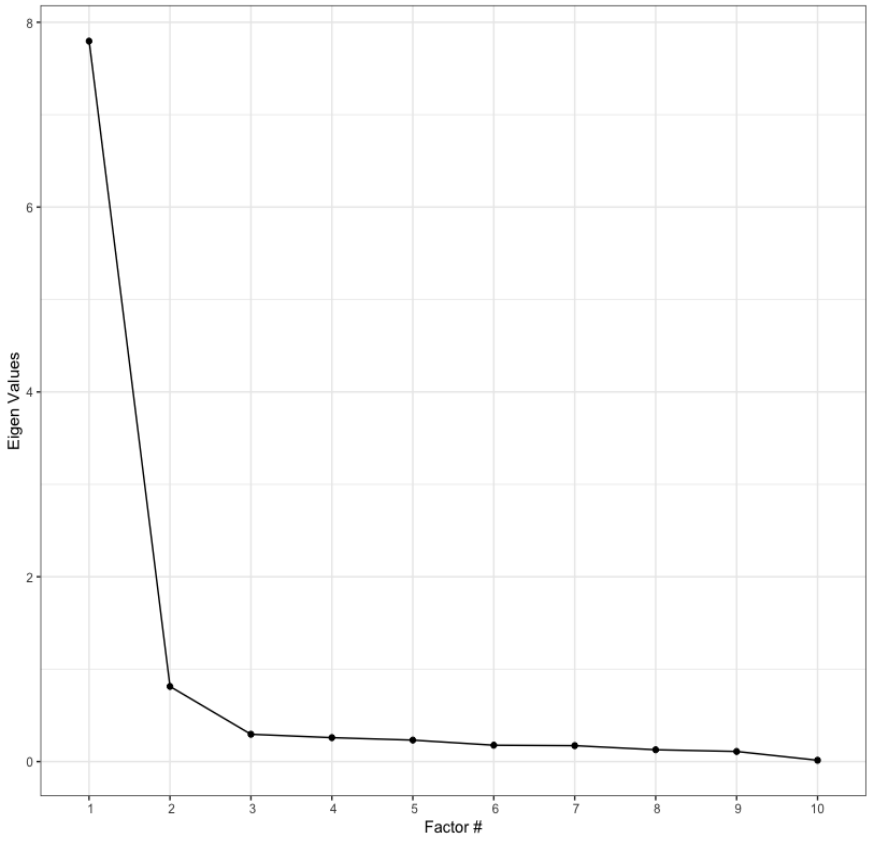
In line with our previous findings, the Bayesian factor analytic (BFA) results demonstrate that a single dimension accurately characterizes our feature space. Table 3 below shows highly stable and consistent loadings along with no problematic cross loadings. These findings suggest a strong likelihood of a single dimension being the most parsimonious and descriptive characterization of the space (Model 1: D1 in the unidimensional specification, and Model 2: D2 in the multidimensional specification). Furthermore, we display a scree plot, similar to the scree plots in figures 1 and 2, in figure 3. The pattern strongly supports the single dimension based on the eigenvalue greater than 1 for only the first dimension, and less than 1 for all subsequent dimensions in line with Kaiser criterion to drop all factors with values less than 1—as we also employ in the preceding FA and PCA analysis. In sum, these result suggest a unidimensional concept among our data in line with our expectations and discussion.

Table 3: Bayesian Factor Analysis Loadings: 1 and 2 Dimensions

	Model 1: D1	**Model 2: D1**	**Model 2: D2**
GMF	2.112193713	0.920740174	3.317962357
Freedom House	2.085784206	1.077890832	3.738413726
RSF	2.048439237	0.701223964	2.433030605
Vdem: Trad M.Censor	2.784529112	-0.239387663	2.716688836
Vdem: M.Critical	3.143672582	-0.727514925	3.266019117
Vdem: M.Perspectives	3.143053118	-0.644468848	3.121336621
Vdem: M.Harassment	2.424771164	-0.174529874	2.34798901
Vdem: M.Self-censor	2.929314959	-0.493303479	2.831894694
Vdem: M.Bias	3.199232661	-0.520211259	3.072580855
Vdem: M.Corrupt	2.318612175	-0.131576045	2.264934156

D.3.1 Scree Plot

Figure 3: BFA Scree Plot



E Assess MSF Scores: Content and Convergent Validity

While we discuss core validation techniques for our new measure in the manuscript, space constraints do not allow us to fully address them there. In this section, we further explore the validity of our new measure. Based on criteria suggested by McMann et al. (2016), we investigate to what degree our measure holds *content validity* and *convergent validity*. We address each validity category in the subsections below, structuring them based on the authors' guiding questions to assess data validity (table 1, pg. 6). This structure helps us organize our thoughts around data validation.

E.1 Content Validity

To what extent does the measure capture the higher-level theoretical construct it is intended to capture and exclude irrelevant elements? (McMann et al., 2016, pg. 6)

We include a lengthy section in the manuscript discussing media freedom theoretically as a latent measure. We root the measured in both academic conceptualizations of media freedom. Using this as guidance to select variables for the measurement models ensures our measurement model reflects the underlying media freedom concept appropriate for academic and policy researchers (that is, it has *resonance* and *domain*). By drawing theoretical boundaries around the concept, we become better able to *differentiate* against variables that exist beyond our conceptualization. We also take the various, complex ideas that relate to media freedom and create a parsimonious concept that is specific enough to describe the concept (that is, *fecundity*), but also allows for *consistency* across countries and time. *Coherence* comes from our IRT model that provides a single measure of the latent concept that remains consistent both across countries and over time.

To what extent is the measure useful to research? (McMann et al., 2016, pg. 6)

The measure produces time-series, cross-sectional data on media freedom that is analyzable from 1948 to 2017 in 197 countries. The dataset includes more years and countries than any other media freedom dataset currently available. To demonstrate its usefulness in empirical research, we utilize it in four replications studies. We present one in the manuscript and three additional ones

here.

E.1.1 Replications (cont.)

Besides the Egorov, Guriev and Sonin (2009) that we present in the manuscript, we replicated three other studies that empirically analyze media freedom: Whitten-Woodring (2009), Schoonvelde (2014), and Stier (2015). We present them in this section below. Like the manuscript's replication, we focus only on models that reflect each study's core findings. Future work may wish to broaden the scope of these replications to examine some of their robustness checks, extend the samples by both countries and years, as well as include more empirical studies of media freedom beyond the four in the manuscript and appendix. The last subsection presents the summary statistics for all data used in replications both below and in the manuscript.

Whitten-Woodring (2009)

Whitten-Woodring (2009) argues that media freedom's influence on government respect for human rights differs depending on the level of democratic institutions.[6] Her nuanced theory produces a number of testable, empirical implications. The paper's core findings indicate that in autocracies, human rights violations will increase *if media are free*. She argues that journalists in autocracies will readily report human rights violations when they remain free to do so, exposing those violations to citizens. However, because citizens do not have a democratic, institutional outlet to address these violations (such as removing leaders through elections), they instead protest against the regime. This action leads to more human rights violations once the government suppresses the protesters. She also argues that a free press effectively reduces human rights violations in the most consolidated democracies. Further, media freedom should not impact respect for human rights in regimes between these two extremes.

Though she runs a number of regressions, Whitten-Woodring spends much of the analysis interpreting and discussing model 6 in table 3 (pg. 612). However, when we replicate the marginal effects graph of this model (a figure crucial to evaluating the author's conditional hypotheses),

[6]Replication material available on Harvard's Dataverse. See https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/AGZHQW. As of 12/14/2018, Google Scholar indicates the paper has 114 citations.

it does not match the author's figure 3 (pg. 613). Because only a direct comparison makes this replication useful, we focus on model 8 in the same table for this replication; we were able to successfully replicate the marginal effects figure (Technical Report, figure A4 (pg. 622)) with the author's data.[7]

We briefly describe Whitten-Woodring's model and estimation approach below. The author uses the CIRI variable that evaluates government respect for the human right of physical integrity as her dependent variable (Cingranelli and Richards, 2006). Higher values indicate greater government respect for human rights. Next, she uses Van Belle's (1997) media freedom measure. We note that this variable is the forerunner to the Global Media Freedom Dataset. The author collapses the index to create a binary variable with 1 indicating a free media and a 0 indicating a media that is not free. She also uses an ordinal, five-category score of democracy based on the polity dataset (Marshall and Jaggers, 2017), with higher scores indicating that a country possesses more democratic institutions. She also uses an interaction term of the media freedom dummy and the five-point polity score to test the conditional effect about which she hypothesized. To complete the model, she uses controls including the logs of *GDP per capita* and *population*, pulling both from the Penn World Table (Heston, Summers and Aten, 2002). Finally, Whitten-Woodring includes variables that measure *interstate* and *intrastate* conflict (Gleditsch et al., 2002; Strand et al., 2005). She estimates the model using a panel-corrected standard error regression with a first-order autocorrelation on a sample of 93 countries from 1981 to 1995. We present the replications and reanalyses below in table 4.

Model 1 in table 4 below replicates Whitten-Woodring's original model 8 in her table 3 (pg. 612). We highlight the key findings here. Media freedom is negative and statistically significant, indicating that countries with a free media see less government respect for human rights. The five-point polity score is positive and statistically significant, suggesting that higher levels of democracy associate with greater respect for human rights. The interaction term is also positive and statistically significant.

[7]The author's use of a lagged dependent variable differentiates the models. We note that we use code from Brambor, Clark and Golder (2005) for both figure replications, and present this code in the replication materials.

However, to evaluate the conditional hypotheses, she graphs the interaction term to better interpret democracy's conditional effect on the relationship between media freedom and government respect for human rights. We replicate her figure A4 in our figure 4a below to show the marginal effect of a free media on government respect for human right conditional on democracy level. As expected, it shows a reductive effect for a free media on government respect for human rights at 0—the lowest category in her five-point polity coding scheme. Also in accordance to her theory, a free media increases government respect for human rights at the highest democracy levels. Finally, the figure indicates a null effect for the second and third categories in the five-point polity variable. The results largely confirm her arguments—a free media increases government respect for human rights at the highest levels of democracy, while it decreases government respect for human rights in the most autocratic regimes.

Model 2 in table 4 below reanalyzes model 1 but replaces Van Belle's (1997) binary variable with our continuous, Media System Freedom (MSF) point estimates. In model 2, media free-

Table 4: Whitten-Woodring (2009) Replication, Model 2 (Table 3)

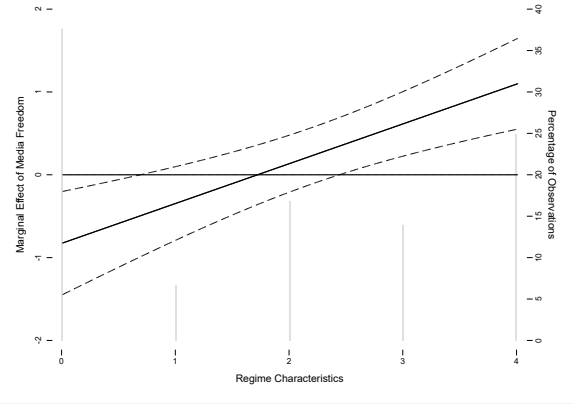
	Model 1 Replication	Model 2 w/ MSF	Model 3 MCS
	β/(SE)	β/(SE)	β/(SE)
Media Freedom	-.827***	-.696	-.748
	(.318)	(.662)	(.601)
Polity (Ord.)	.186***	-.612***	-.516***
	(.066)	(.135)	(.134)
Media Freedom*Polity (Ord.)	.48***	1.427***	1.311***
	(.121)	(.182)	(.187)
GDP p/c (log)	.355***	.139	.164*
	(.091)	(.094)	(.088)
Population (log)	-.388***	-.407***	-.405***
	(.04)	(.034)	(.035)
Interstate Conflict	.021	.022	.031
	(.218)	(.189)	(.201)
Internal Conflict	-1.868***	-1.795***	-1.829***
	(.195)	(.189)	(.191)
N	1,395	1,395	1,395
Countries	93	93	93
R^2	.396	.4234	-

$***p < .01, **p < .05, *p < .1$; β = Coefficient Estimates; (SE)= Standard Errors; Constants not reported; Sample covers years 1981-1995; Monte Carlo simulation (MCS) w/ 750 iterations: β = mean of 750 coefficient estimates, (SE)= mean of 750 standard error estimates.

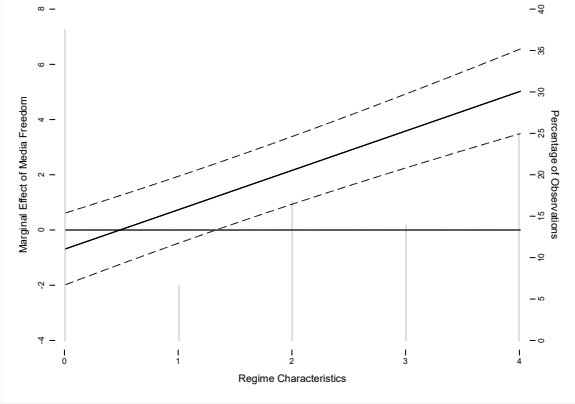
dom's effect on government's respect for human rights changes as a result. *Media freedom* loses significance, though it retains its negative direction. The five-point democracy score retains its statistical significance but its sign switches to positive. Furthermore, the interaction between the two remains positive and statistically significant. We also note that the controls' statistical significance and sign remains the same, though *log GDP p/c* does lose statistical significance. In figure 4b below, we graph the interaction terms from our model 2, and the results appear different from Whitten-Woodring's original figure. The most autocratic category no longer displays the statistically significant reductive effect. Similar to Whitten-Woodring's original figure, the next highest polity category is insignificant, however the remaining polity levels are statistically significant. This suggests a positive, significant effect of media freedom on government respect for human rights in countries at or above 3 in this five-point democracy score. Finally, model 3 utilizes the posterior standard deviations to run a Monte Carlo simulation (as we described in the manuscript's *Replication* section). The results remain largely similar to model 2, though *log GDP p/c* becomes statistical significant at the 90% confidence level.

Figure 4: Marginal Effect of Media Freedom on Gov't Respect for Human Rights, 1981-1995

(a) Whitten-Woodring (2009) Fig. A4



(b) w/ MSF Score



Note: Marginal effect is across Whitten-Woodring's five-point democracy score.

In sum, not all of Whitten-Woodring's results hold when we utilize the MSF data. Specifically, the results suggest that the author's initial finding that media freedom's reductive effect on government respect for human rights in the most autocratic regimes does not hold. Furthermore and

to confirm the author's initial finding, the results provide strong evidence for the claim that media freedom increases government respect for human rights in more democratic countries.

Stier (2015)

Stier (2015) examines media freedom variation among regime types.[8] Like Whitten-Woodring, he makes a number of nuanced arguments and tests multiple hypotheses. First, he argues liberal democracy is positively associated with higher levels of media freedom. However, the author further argues that media freedom varies across the autocratic spectrum. Specifically, he suggests that electoral autocracies, monarchies and military regimes have the freest media, while regimes with a communist ideology have the least.

Here, we replicate Stier's models 1 and 3 in table 1 (pg. 1284). Model 1 examines the effect of liberal democracy on media freedom, while model 3 analyzes the effects among differing autocratic regime-type on media freedom. Both models employ Freedom House's *Freedom of the Press* media freedom variable as their dependent variables. To measure the various regime types, the author uses Kailitz's (2013) binary variables on *liberal democracy* and autocratic regimes. The various autocratic regime-type variables include: *monarchy*, *military*, *one-party rule*, *personalist*, and *communist*. *Electoral autocracy* is the reference category. The author's controls include *log GDP p/c* (Bolt and van Zanden, 2013), *internal war* (Gleditsch et al., 2002; Harbom and Wallensteen, 2012), *natural resource income* (Ross, 2011), as well as *internet users* and *log population* (World Bank, 2017). Stier estimates the models using a panel-corrected standard error specification with a panel specific, first-order autocorrelation on a sample of 149 countries from 1993 to 2010. The model also includes region and decade dummies. We present the replications and reanalyses below in table 5.

Model 4 in table 5 replicates Stier's model 1 in table 1 (pg. 1284). The data show *liberal democracy* is positive and statistically significant, indicating that media freedom is higher in liberal democracies, on average. These results hold when we include our MSF point estimates in model 5 and run the Monte Carlo simulation in model 6. We also note that *log population* and *internet*

[8]We thank the author for directly providing replication materials. As of 12/14/2018, Google Scholar indicates the paper has 19 citations.

Table 5: Stier (2015) Replication, Models 1 & 3 (Table 1)

	Model 4 Replication	Model 5 w/ MSF	Model 6 MCS	Model 7 Replication	Model 8 w/ MSF	Model 9 MCS
	β/(SE)	β/(SE)	β/(SE)	β/(SE)	β/(SE)	β/(SE)
Monarchy				.301	.043***	.014
				(2.023)	(.01)	(.038)
Military regime				-4.175	-.001	-.026
				(2.365)	(.009)	(.028)
One-party regime				-5.509**	-.038***	-.083***
				(1.743)	(.01)	(.026)
Personalist regime				-6.988**	-.042***	-.138***
				(2.233)	(.011)	(.033)
Communist ideocracy				-21.658***	-.373***	-.365***
				(3.105)	(.02)	(.029)
Liberal Democracy	9.186***	.067***	.132***			
	(1.37)	(.01)	(.018)			
Log GDP per capita	9.507***	.060***	.067***	.297	-.020**	-.026**
	(.983)	(.007)	(.012)	(1.135)	(.007)	(.011)
Log population	-1.303***	-.004	.004	-.523	.042***	.032***
	(.369)	(.004)	(.005)	(.724)	(.005)	(.008)
Oil and gas income	-.00019**	-.0000016**	-.000003**	.001	-.001	-.001
	(.00007)	(.0000007)	(.000001)	(.001)	(.001)	(.001)
Internet users	-.055**	.001	.001	-.091*	.0004*	.001
	(.02)	(.001)	(.001)	(.042)	(.0002)	(.001)
Internal war	-2.863***	-.016***	-.017*	-2.815***	-.002	-.007
	(.563)	(.004)	(.008)	(.645)	(.003)	(.008)
N	2,477	2,477	2,477	1,092	1,092	1,092
Countries	149	149	149	80	80	80
R^2	.8667	.9512	-	.7453	.9295	-

$***p < .001, **p < .01, *p < .05$; β = Coefficient Estimates; (SE)= Standard Errors; Regions and decade dummies not reported; Constants not reported; Sample covers years 1993-2010; Monte Carlo simulation (MCS) w/ 750 iterations: β = mean of 750 coefficient estimates, (SE)= mean of 750 standard error estimates.

users lose statistical significance when we include the MSF scores in model 5 and the Monte Carlo simulation model.

Model 7 replicates Stier's model 3 in table 1 (pg. 1284). The replication indicates that *one-party*, *personalist*, and *communist* regimes are negative and statistically significant. As expected, the results also indicate that *one-party*'s magnitude is the least of the statistically significant, autocratic regime-type variables, while *personalist*'s coefficient is larger, and *communist* appears the largest. Model 8 replaces the Freedom House variables with our MSF point estimates. The results for *one-party*, *personalist*, and *communist* regimes remain largely the same, though we find that *monarchy* is now positive and statistically significant. This suggests monarchies have higher media freedom levels than the reference regime—*electoral autocracy*. Model 8 also shows different

results for all controls. Both *log GDP per capita* and *log population* have switched directions to positive and become statistically significant. *Internet users* has switched signs from negative to positive though it remains statistically significant. *Internal war* remains negative, but loses statistical significance. Finally, the Monte Carlo simulation in model 9 returns largely similar results to model 8, though *monarchy* and *internet users* are no longer statistically significant.

In sum, the results provide robust support for Stier's core arguments. Liberal democracies possess higher media freedom levels than non-liberal democracies. The data also suggest strong reductive associations for various authoritarian regime types. Communist regimes carry the lowest media freedom, followed by personalist, then one-party regimes, and then electoral autocracies. These results remain robust after a Monte Carlo simulation that incorporates the MSF posterior standard deviations—the measure of uncertainty among extant media freedom measures.

Schoonvelde (2014)

Schoonvelde (2014) investigates the conditions that encourage greater political knowledge among individuals, arguing that the political knowledge that citizens possess depends in part on the relationship between government and media.[9] His core findings suggest that a freer media positively increases political learning as well as conditions the individual-level effect of education on political knowledge. The author puts forth two hypotheses: the *Media Freedom Hypothesis* and the *Education Hypothesis*. We focus here on the former, which essentially states that individuals in freer media systems are more knowledgable about politics than those in less free media systems.

We replicate Schoonvelde's model 2 and 4 in his table 4 (pg. 172), which evaluates the *Media Freedom Hypothesis*. Model 2 features *political knowledge* as its dependent variable, and the author uses a variable from the first wave of the Comparative Study of Electoral Systems (CSES) dataset (Sapiro and Shively, 2003). The dataset features multi-level, cross-national survey responses on elections in numerous countries. According to Schoonvelde, CSES measures *political knowledge* by taking "the absolute distance between an individual's perception of the largest party's ideological location and the mean perception of the electorate at large" (pg. 167). He also

[9]Replication material available on Harvard's Dataverse. See https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/24122. As of 12/14/2018, Google Scholar indicates the paper has 9 citations.

inverts the variable so that larger scores correspond to higher knowledge levels. He also notes that "the variable is logged and recoded so that it fits a normal distribution and varies between 0 and 1" (pg. 167). Next, Freedom House's *Freedom of the Press* is the key independent variable, though for ease of interpretation the author converts the measure to a 0 to 1 scale and inverts it so that higher scores indicate increased media freedom. We note that this is the same scale as the MSF scores.

To complete the model, the author adds a number of individual and country-level variables. The individual-level variables include *income*, *education*, *union membership*, and *age* taken from the CSES survey. As for country-level ones, the author includes the *Herfindahl Index* that measures media market competitiveness in both the newspaper and television sectors (Hirschman, 1964). These variables range from 0 to 1, with higher scores indicating greater monopoly in that media sector. The author also includes Baek's (2009) *public broadcasting percentage* variable that measures "the audience share of public broadcasting channels as the percentage of total audience size of the five largest television stations" (Schoonvelde, 2014, 168). Next, the model includes *newspaper demand*, or daily newspaper subscribers per capita in the population[10] and *GDP per capita*.[11] Other political variables at the country level include dummies for a parliamentary system (Golder, 2005), where compulsory voting is mandatory (Sapiro and Shively, 2003), where party lists exist (Sapiro and Shively, 2003), and majoritarian systems. To complete the model, he includes the *effective number of electoral parties*.[12] To estimate the model, Schoonvelde employs a random intercept (multi-level) model. We also replicate Schoonvelde's model 4. It is similar to model 2 but removes some of the controls, which increases the sample from 23 to 31 countries and the number of observations (respondents) from 28,450 to 37,077.[13] we present the replications and reanalyses below in table 6.

Model 10 in table 6 below displays our replication of Schoonvelde's model 2 in table 3 (pg.

[10]UNESCO Institute for Statistics: http://www.uis.unesco.org/Pages/default.aspx.

[11]World Bank: http://data.worldbank.org/.

[12]The author does not reveal from where he drew the majoritarian systems dummy or *effective number of electoral parties*.

[13]We note that models 1 and 3 in Schoonvelde's table 4 (pg. 172) are similar to models 2 and 4 except that they lack *GDP per capita*.

Table 6: Schoonvelde (2014) Replication, Models 2 & 4 (Table 4)

	Model 10 Replication	Model 11 w/ MSF	Model 12 MCS	Model 13 Replication	Model 14 w/ MSF	Model 15 MCS
	β/(SE)	β/(SE)	β/(SE)	β/(SE)	β/(SE)	β/(SE)
Individual-level variables						
Income	.023***	.023***	.023***	.024***	.024***	.023***
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Education	.044***	.044***	.044***	.041***	.041***	.042***
	(.003)	(.003)	(.003)	(.002)	(.002)	(.003)
Union Membership	-.006***	-.006***	-.006***			
	(.001)	(.001)	(.001)			
Age	-.008	-.007	-.008			
	(.0046)	(.004)	(.0043)			
Country-level variables						
Media Freedom	.138***	.18*	.024	.13***	.12***	.054
	(.035)	(.09)	(.045)	(.025)	(.003)	(.028)
Parliamentary System	.041***	.028	.039**	.022**	.019	.025**
	(.01)	(.015)	(.014)	(.008)	(.01)	(.01)
Compulsory Voting	-.012	-.028	-.021	-.03***	-.0298**	-.031**
	(.013)	(.019)	(.018)	(.008)	(.0099)	(.009)
Party List	.01	-.004	.023	.012	.007	.013
	(.011)	(.021)	(.016)	(.009)	(.011)	(.011)
Majoritarian System	-.009	-.018	.008	-.009	-.008	-.008
	(.014)	(.024)	(.02)	(.012)	(.01)	(.014)
Effective Number of Parties	-.0013	.001	-.001			
	(.002)	(.003)	(.002)			
Newspaper Competitiveness	-.002	-.097	-.05			
	(.062)	(.092)	(.087)			
Television Competitiveness	.084	.097	.031			
	(.053)	(.081)	(.077)			
Newspaper Circulation	-.004	-.03	-.014			
	(.066)	(.094)	(.088)			
Public Broadcasting Pct.	.045*	.002	.026			
	(.022)	(.034)	(.032)			
GDP Per Capita	-.001	-.001	-.001	-.000001*	-.001	-.001
	(.001)	(.001)	(.001)	(.0000005)	(.001)	(.001)
Observations	28,450	28,450	28,450	37,077	37,077	37,077
Elections	23	23	23	31	31	31
R^2 (individual)	.0146	.0146	-	.0127	.0127	-
R^2 (election)	.8449	.7715	-	.783	.717	-
AIC	-31,790.51	-31,787.49	-	-40,264.52	-40,258.51	-

$***p < .001, **p < .01, *p < .05$; β = Coefficient Estimates; (SE)= Standard Errors; Constants not reported; Sample covers surveys conducted between 1996 and 2002; Monte Carlo simulation (MCS) w/ 750 iterations: β = mean of 750 coefficient estimates, (SE)= mean of 750 standard error estimates.

172). To save space, we will focus our interrelations on *media freedom* but note any changes in controls. *Media freedom* is positive and statistically significant, indicating that individuals in countries with higher media freedom tend to have higher political knowledge, on average. This finding supports the *Media Freedom Hypothesis*. When we include our MSF point estimates in

model 11, this result holds; *Media freedom* remains positive and statistically significant. However, the result does not hold when we run the Monte Carlo simulation in model 11; *Media freedom* remains positive put loses statistical significance. Significance levels for the controls remain largely the same, though *parliamentary system* loses significance in model 11 and *public broadcasting percentage* loses statistical significance in models 11 and 12. We note that *age* was not statistically significant in model 10, the replication model, contrary to the author's published results ($p = .077$).

We next turn to replicating Schoonvelde's model 4 in table 3 (pg. 172), which only differs from model 2 by removing some of the controls. This increases the sample size both in terms of respondents and countries covered. The replication in our model 13 indicates that *media freedom* is positive and statistically significant. This result holds when we replace the Freedom House scores with the MSF point estimates in model 14, though the coefficient becomes insignificant when we run the Monte Carlo simulation in model 15.[14] The controls remain largely the same, though *parliamentary system* loses significance in model 14 and *compulsory voting* becomes negative while remaining statistically significant in models 14 and 15. In addition, *GDP per capita* loses significance in models 14 and 15.

In sum, the results from this replication and reanalysis are mixed. Schoonvelde's finding that media freedom is positively associated with media freedom generally holds when we use the MSF point estimates. However, when we employ the posterior standard deviations in the two Monte Carlo simulations—a more rigorous test—, this result does not hold.

[14]Though the simulation returns a t-statistic of 1.951, just below the required 1.96 score to achieve statistical significance at the 95% confidence level.

Summary Statistics

We present the summary statistics for all variables that we use for our replication analyses in tables 7, 8, and 9 below.

Table 7: Variables Used in Replications, I

Variable	Obs	Mean	Std. Dev.	Min	Max
Egorov et al. (2009)					
Model 2 (Table 1)					
Freedom House (lead)	2,056	53.18823	22.81544	1	95
MSF (lead)	2,056	.620897	.2631726	.0045	.9988
MSF, st. dev. (lead)	2,056	.0297868	.0117465	.0009	.0589
Log Oil Reserves Value	2,056	3.954202	1.210115	2.814676	9.785866
Polity	2,056	3.562257	6.463825	-10	10
Log GDP p/c, PPP	2,056	8.427315	1.307141	5.492956	11.16642
Log Population	2,056	16.12022	1.506039	12.77992	20.99467
Log Govt. Expend./GDP	2,056	2.662588	.4189498	1.051084	4.241943
Model 5 (Table 1)					
Freedom House (lead)	958	36.29749	14.3977	1	74
MSF (lead)	958	.4346698	.1984967	.0045	.9113
MSF, st. dev. (lead)	958	.0315472	.009021	.0032	.0547
Log Oil Reserves Value	958	4.146661	1.391827	2.814676	9.785866
Log GDP p/c, PPP	958	7.847244	1.201062	5.492956	11.16642
Log Population	958	15.96781	1.509885	12.77992	20.99467
Log Govt. Expend./GDP	958	2.566447	.4395115	1.051084	3.734825
Whitten-Woodring (2009)					
Physical Integrity (CIRI)	1,395	5.117563	2.326739	0	8
Media Freedom (Freedom House)	1,395	.4258065	.494642	0	1
Media Freedom (MSF)	1,395	.5510014	.303437	.0223	.9987
MSF, st. dev.	1,395	.0298411	.0114226	.001	.0569
Polity (Ord.)	1,395	1.817921	1.635719	0	4
Log GDP p/c	1,395	8.330455	1.065341	6.078282	10.25446
Log Population	1,395	9.270705	1.490713	6.426488	14.0006
Interstate Conflict	1,395	.0408602	.1980373	0	1
Internal Conflict	1,395	.1870968	.3901291	0	1

Table 8: Variables Used in Replications, II

Variable	Obs	Mean	Std. Dev.	Min	Max
Schoonvelde (2014)					
Model 2 (Table 4)					
Voter Knowledge	28,450	.2724482	.1425685	0	1
Media Freedom (Freedom House)	28,450	.8057052	.2076309	.2857143	1
Media Freedom (MSF)	28,450	.9110795	.1481192	.5027	.9987
MSF, st. dev.	28,450	.0176303	.0133329	.001	.0436
Income	28,450	.5176714	.3493366	0	1
Education	28,450	.5225577	.2880803	0	1
Union Membership	28,450	.2713884	.4446838	0	1
Age	28,450	.3594053	.189955	.0232558	1
Parliamentary System	28,450	.6040422	.489064	0	1
Compulsory Voting	28,450	.3054482	.4606051	0	1
Party List	28,450	.5210545	.4995653	0	1
Majoritarian System	28,450	.2578559	.4374619	0	1
Effective Number of Parties	28,450	4.493525	2.163641	2.12	9.725
Newspaper Competitiveness	28,450	.0804998	.0582153	.003	.23
Television Competitiveness	28,450	.2122676	.094734	.047	.377
Newspaper Circulation	28,450	.2298846	.1344587	.038	.591
Public Broadcasting Pct.	28,450	.4942148	.2338514	0	.96
GDP Per Capita	28,450	18878.49	11194	835	38291
Model 4 (Table 4)					
Voter Knowledge	37,077	.2696175	.1454472	0	1
Media Freedom (Freedom House)	37,077	.7763426	.2264835	0	1
Media Freedom (MSF)	37,077	.8827002	.1752878	.1331	.9987
MSF, st. dev.	37,077	.0201443	.0145609	.001	.0505
Income	37,077	.5104243	.3416119	0	1
Education	37,077	.533406	.2888178	0	1
Parliamentary System	37,077	.5168433	.499723	0	1
Compulsory Voting	37,077	.3140222	.4641315	0	1
Party List	37,077	.5504221	.4974578	0	1
Majoritarian System	37,077	.2194892	.4139061	0	1
GDP Per Capita	37,077	17854.39	11486.53	835	38291

Table 9: Variables Used in Replications, III

Variable	Obs	Mean	Std. Dev.	Min	Max
Stier (2015)					
Model 1 (Table 1)					
Media Freedom (Freedom House)	2,477	51.36294	24.00041	0	95
Media Freedom (MSF)	2,477	.6010674	.2902484	.0029	.9988
MSF, st. dev.	2,477	.0290213	.0120632	.0009	.0589
Lib. Democracy, lag	2,477	.5478401	.4978066	0	1
GDP (log), lag	2,477	8.250111	1.130093	5.429179	10.36265
Population (log), lag	2,477	16.21922	1.47926	12.81821	21.00939
Oil and gas income, lag	2,477	889.2129	3425.546	0	50468.05
Internet users, lag	2,477	11.32736	19.69421	0	92.08
Internal war, lag	2,477	.1473557	.3545317	0	1
Region 2	2,477	.0343157	.1820756	0	1
Region 3	2,477	.02826	.1657482	0	1
Region 4	2,477	.0637868	.2444222	0	1
Region 6	2,477	.1199031	.3249138	0	1
Region 7	2,477	.132822	.3394507	0	1
Region 8	2,477	.0193783	.1378783	0	1
Region 9	2,477	.1812677	.3853176	0	1
Region 10	2,477	.152604	.3596779	0	1
Decade	2,477	.614453	.4868226	0	1
Model 3 (Table 1)					
Media Freedom (Freedom House)	1,092	30.66484	14.5902	0	78
Media Freedom (MSF)	1,092	.3482845	.2106878	.0029	.8914
MSF, st. dev.	1,092	.029151	.0102474	.0019	.0547
Monarchy (Kailitz), lag	1,092	.1593407	.3661612	0	1
Military (Kailitz), lag	1,092	.0430403	.2030408	0	1
One-party (Kailitz), lag	1,092	.0320513	.176217	0	1
Personalist (Kailitz), lag	1,092	.0769231	.2665914	0	1
Communist (Kailitz), lag	1,092	.0796703	.2709062	0	1
GDP (log), lag	1,092	7.746229	.9561215	5.429179	10.19553
Population (log), lag	1,092	16.14676	1.471619	12.89828	21.00939
Oil and gas income, lag	1,092	1543.902	4682.514	0	50468.05
Internet users, lag	1,092	4.401282	10.04681	0	69.9
Internal war, lag	1,092	.2051282	.4039803	0	1
Region 2	1,092	.03663	.1879378	0	1
Region 3	1,092	.03663	.1879378	0	1
Region 4	1,092	.1034799	.3047241	0	1
Region 6	1,092	.239011	.4266749	0	1
Region 7	1,092	.0302198	.17127	0	1
Region 8	1,092	.0119048	.1085072	0	1
Region 9	1,092	.1630037	.3695383	0	1
Decade	1,092	.5879121	.4924363	0	1

E.2 Convergent Validity

Does the measure accurately capture actual cases? (McMann et al., 2016, pg. 6)

To investigate the degree to which our measure captures actual cases, we present case studies that detail the case histories of media freedom levels in the eight countries in our manuscript's figure 4. These include six studies conducted by the authors (Spain, North Korea, the United States, Nigeria, and Brazil), as well as two conducted by a data-blinded research assistant (Jamaica and Armenia). We present these results below.

E.2.1 Case Studies by Authors

Poland

After World War II and a Soviet military occupation, communists with close links to the Soviet Union took power in Poland and imposed harsh censorship laws (Bajomi-Lázár, 2014, 104). Poland's media system remained an instrument of the communist regime until a number of protests in the late 1970s and early 1980s (notably including the Solidarity movement) slightly improved Poland's civil liberties for a short time (Jakubowicz, 1992). However, the communist party would impose martial law from 1981 to 1983 in response to the protests. By 1988, the Solidarity movement engaged in dialogue with the now unpopular Polish communist leaders. Change in media freedom followed these events with the communist leadership giving the Solidarity movement certain broadcasting and publishing rights in 1989. For example, after years as an underground publication, officials allowed editors of the daily newspaper *Gazeta Wyborcza* to openly publish on May, 8, 1989 (Bajomi-Lázár, 2014, 107). Finally, in June and July of 1989 Poland held semi-competitive parliamentary and presidential elections that allowed non-communist leaders to take power (Bajomi-Lázár, 2014, 105-06). As the political transition continued, Soviet-era censorship officially ended in Poland on April 11, 1990 (Kitschelt, Mansfeldova, Markowski and Tóka, 1999; Ost, 2001), and by 1992 Polish leaders amended the constitution to ban government censorship and guarantee freedom of speech (Curry, 2003).

The MSF scores reflect the qualitative evidence. Beginning in 1948 during the communist

regime, Poland displays a score at nearly 0 until about 1980 when it slightly increased during societal protests that briefly improved civil liberties. The score then decreases around 1981 when the government introduced martial law. Then, around 1990 when the political transition to democracy resulted in officials lifting the most restrictive media repression laws, the score rises steeply to nearly 1. Poland held this high MSF score until 2016 when it dropped to about .708 from a 2015 point estimate of about .935. According to Chapman (2017), the conservative Law and Justice Party (PiS) began attempting to impose controls on public and private media once it took office October 2015. The measures included firing and replacing heads of broadcasting outlets with individuals friendly to the party, such as appointing former PiS member of the lower house (*Sejm*) Jacek Kurski as head of the *Telewizja Polska* television station (pg. 16). Chapman (2017) also reports the government redirects advertising funds to more conservative-leaning outlets and talks of "repolonizing" the media, which the author argues is an effort to replace foreign media owners with domestic ones. The temporal data captures this 2016 downturn. In sum, given Poland's case history, the MSF scores reflect manifestations of nearly absent media freedom as well as nearly perfect media freedom. In addition, it registers instances when nearly imperfect media freedom increased slightly and when nearly full media freedom decreased.

Spain

At the beginning of Spain's MSF scores in 1948, General Francisco Franco had been Spain's authoritarian ruler since 1939 at the Spanish Civil War's conclusion. During the civil war, Franco's forces took measures to control the printing press in 1936 and broadcast media in 1938 (Deacon, 1999, 311). With the civil war's conclusion, media served as the government's mouthpiece. The state tightly controlled Spain's media, though some officials attempted meager reforms later in Franco's life. In 1966, a law abolished prior censorship, meaning the press could publish stories without prior government approval. However, this new law did not amount to any substantial change in media freedom (Deacon, 1999, 313). Franco would die in 1975, but officials were slow to lift restrictions (Deacon, 1999, 313). In 1978, the Spanish people voted to ratify a new constitution that included a declaration of freedom of expression and freedom to receive and disseminate

information.[15] It also barred government censorship.[16] Finally, the 1980 Statute of Radio and Television guaranteed plurality of broadcast media, including opening media to areas with regional identities that Franco's regime suppressed, such as the Basque Country and Catalonia.

The MSF scores reflect these changes. The score remains near 0 from 1948 until 1975 when the authoritarian Franco dies, but gradually increases as the case study suggests — .3479 in 1976, .5803 in 1977, .8225 in 1978, .9304 in 1979, and then .9648 in 1980. Spain's estimate remains near 1 around .96 until the sample ends in 2017.

North Korea

The Democratic People's Republic of Korea (North Korea) had been a sovereign, communist regime since 1948 when the Soviet Union ended its occupation. At the beginning of the MSF scores, North Korea's authoritarian regime had already heavily restricted media (Byman and Lind, 2010, 54). This policy has endured to the present.[17] In 2015, the Committee to Protect Journalists (CPJ) named North Korea the second worst media censor in the world,[18] while in 2017 the watch-dog group described North Korea's dictator Kim Jong Un as having an "absolute grip on the flow of public information and [a] deadly approach to dissent."[19]

The MSF scores reflect this description. North Korea begins the time series with a score near zero, indicating a virtual absence of media system freedom, and retains it throughout the sample.

United States

The United States' constitution has explicitly guaranteed press freedom since the country's founders ratified it in 1788. In addition, constitutions in nine of the North American country's first 13 states included a clause for press freedom (Powe, 1991, 26). Though the government did impose restrictions on the press at various times in the country's development,[20] the media in the United

[15] Spanish Constitution of 1978, Art 20.

[16] Spanish Constitution of 1978, Art 2.

[17] *BBC* "North Korea's tightly controlled media" Dec. 19, 2011. Available at: http://www.bbc.com/news/world-asia-pacific-16255126.

[18] *Committee to Protect Journalists* "10 Most Censored Countries 2015" Available at: https://cpj.org/2015/04/10-most-censored-countries.php

[19] *Committee to Protect Journalists* "Supervised Access" April 25, 2017. Available at: https://cpj.org/2017/04/supervised-access.php

[20] Most notably the 1798 Sedition Act (Lewis, 2007, 11) and the 1918 amendments to the 1917 Espionage Act (Lewis, 2007, 28) (Lewis 2007, 28).

States have generally operated in a mostly free environment. However, subnational government restrictions on media occurred throughout the early to mid 20th century, notably in the US states of Minnesota (Lewis, 2007, 43-44) and Louisiana (Powe, 1991, 222). Lewis (2007) and Powe (1991) both point to three Supreme Court rulings that resulted in media's increased protection from libel suits in the United States: *New York Times v. Sullivan* in 1964, *New York Times v. United States* in 1971, and *Miami Herald Publishing Co. v. Tornillo* in 1974. Kalven (1988) describes the First Amendment "working itself pure" during this time period (pg. xvii).

The United States' MSF scores reflect these changes. At the beginning of the time series in 1948, when media were generally free but still subject to prior restraint and libel law suits in certain circumstances, the country scores a .7185 and steadily increases. In 1967, its score increases to about .8. By 1975, the US scores at .9379 and remains at about this score—near 1—until the dataset's end in 2017. The MSF scores detect these events and change accordingly.

Nigeria

Nigeria achieved independence from Great Britain in 1960 and has experienced the most instability of the six countries we display in the manuscript's figure 5. After declaring independence, Nigeria ratified a new constitution in 1963 that lasted until 1966 when a coup occurred. The coup would eventually trigger a civil war that lasted from 1967 to 1970. At the civil war's conclusion, stability returned to the west African country until another coup occurred in 1975. By 1977, a new constitution ushered in the Second Republic, but in 1983 a military coup replaced it with an authoritarian regime. Yet another coup occurred two years later in 1985. The Nigeria government survived an unsuccessful coup in 1990, and then annulled the results of a 1993 presidential election. Turmoil from the disputed election gave Defense Minister and army general Sani Abacha an excuse to seize power later that year. Abacha would die of a heart attack in 1998, and power passed to General Abdulsalami Abubakar who allowed a transition back to civilian rule in 1999. The May 1999 presidential election established the Fourth Republic, the form of government under which Nigeria remains today. It was in this shuffle of civilian and military regimes that Nigeria's media system would operate.

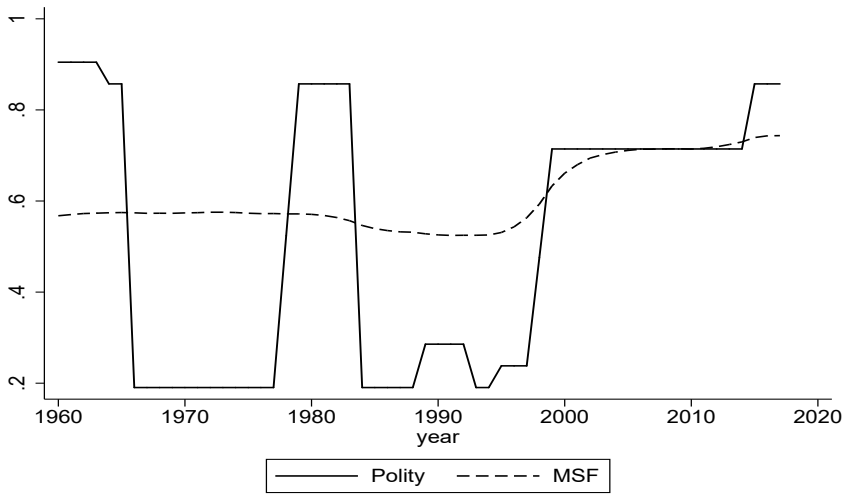
Before Nigeria's independence, British colonists established a press in Nigeria in 1859 (Oso, Odunlami and Adaja, 2011, 1). During colonization, the British restricted the press but outlets published content critical of the colonial powers (Mohammed and James, 2017). By the time Nigeria became independent, various Nigerian leaders, both democratic and non-democratic, used colonial-era media laws to restrict media (Eribo and Jong-Ebot, 1997, 63). Looking back at Nigerian media since independence, Mohammed and James (2017) note that "Nigerian media have been playing a very significant role in setting the agenda for public discourse and molding the direction of public opinions on vital issues in the country." Tejumaiye and Abelabu (2011) argue that Nigerian media did not develop as watchdogs of government, and so did not always perform this role (pg. 73). Eribo and Jong-Ebot (1997) also argue the worst restrictions occurred during military rule. But even once Nigerian leaders established the Fourth Republic and included press freedom protections in the 1999 constitution,[21] the government continued to restrict media. For example, in 2004 the government closed *Weekly Insider* magazine citing national security concerns, and imprisoned *Midwest Herald* publisher Orobosa Omo-Ojo for a story about the Nigerian president's wife in 2005. In reviewing Nigerian state-press relations since independence, Ogbondah (2011) concludes that "the state still utilises arbitrary actions and extra-legal measures that were adopted by erstwhile military regimes in attempts to cow the press and suppress the dissemination of diverse views and information in the media" (pg 46). Given that Nigeria has maintained an active, independent media that democratic and authoritarian regimes have restricted, we expect Nigeria's media freedom to remain relatively constant with the exception of military rule when authors agree government attacks against media intensified.

Nigeria's MSF scores generally reflect this case history. The case history does not suggest that Nigeria should begin the series either fully free or fully not free, so we cannot place Nigeria's MSF score precisely on the latent scale in 1960. However, the case history suggests changes at certain times that the MSF scores reflect. The data begins in 1960 at .5676 and increases nominally to around .5732 by 1975. It remains constant until it dips slightly around 1983 when military

[21] Sections 16, 22, 39(1), and 39(2) of the 1999 Nigerian Constitution.

rule replaced the Second Republic. The score then rises to about .7019 around 2003 when the country transitioned to civilian rule but remains at that level. We highlight that while Nigeria's democracy level changed many times during the sample, media freedom remained largely stable. In figure 5 below, Nigeria's *polity* score changes quite often, while its MSF score remains more stable. Nigeria's case demonstrates that 1) MSF scores do not necessarily depend on regime type by default, and 2) a media system can have a degree of freedom even under authoritarian rule.

Figure 5: Nigeria's Polity and MSF Scores, 1960-2017



Note: Polity standardized to 0 - 1 scale.

Brazil

In contrast to Nigeria's case, Brazil experienced moments of political instability during the sample's time period that *did* affect its media system's freedom. The South American country begins the sample as a democratic regime (Cheibub, Gandhi and Vreeland, 2010; Marshall and Jaggers, 2017), but the military staged a coup in 1964 that forced the country into a dictatorship. In 1985, the country transitioned back to a democracy and has remained one since.

During a democratic period in Brazil that lasted from 1946 to 1964, the press often served as an instrument for political parties (Albuquerque, 2012, 80). For example, backers of the former dictator Getúlio Vargas created the *Última Hora* paper to support their candidate's 1951 election campaign. Also during this time, the press was not market-driven and depended on advertising from the government and state-owned operations as well as bribes (Albuquerque, 2012, 80-81). However, once the generals installed a military regime in 1964, the authoritarian government dissolved political parties and media outlets became either subservient to the regime or authorities censored them (Smith, 1997). For example, during the initial transition *Globo* became a dominant media outlet and developed an "authoritarian model of journalism" by allying with the military dictatorship (Porto, 2012, 61).[22]

Strict censorship lasted until 1974 when General Ernesto Geisel introduced the *abertura* policy that relaxed censorship, in addition to other reforms (Williamson, 2009, 431). This period saw the authorities search for democratic legitimacy for their authoritarian regime. *Abertura* continued under Brazil's next military ruler who abolished the censorship of books and newspapers in 1979, though radio and television censorship continued (Williamson, 2009, 432). The military dictatorship endured until 1985, and a democratic constitution took effect in 1988. The document declares freedom of the press and bars government censorship.[23] Media changed to a more market-driven model as the country's democracy consolidated (Albuquerque, 2012, 81). The current media system remains relatively free, though media ownership remains heavily concentrated (Hervieu, 2013)

[22] *Globo* actually offered an apology for this alliance in 2013 (http://articles.latimes.com/2013/sep/04/world/la-fg-wn-brazil-globo-network-military-20130904).

[23] Censorship and press freedom addressed in Title II (Chapter 1) and Title XIII (Chapter V) respectively.

and the judiciary often censors outlets.[24, 25]

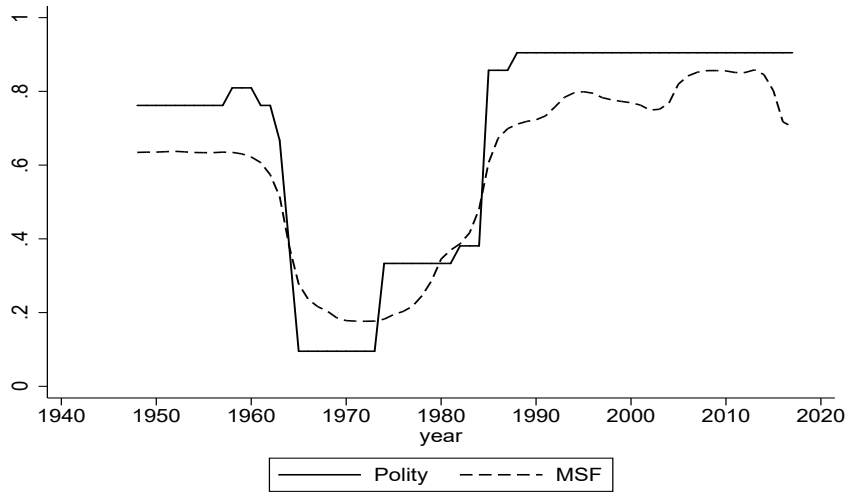
Brazil's MSF scores reflect this case history. The case history does not suggest that Brazil should begin the series either fully free or fully not free, so we cannot place Brazil's MSF score precisely on the latent scale in 1948 based on this history. However, the case history suggest changes at certain times. In 1948, Brazil's score is around .6345 but then decreases in 1964 when the authoritarian regime took power and exerted influence over the country's media. The score steadily decreases for nearly a decade until about 1974 when General Geisel enacted the *abertura* policy. Throughout the *abertura* period until Brazilians ratified the 1988 constitution, the score gradually rises to about .7109. It rises again during this most recent democratic period, though never reaches above about .85. Given media ownership concentration and judicial censorship, we do not expect Brazil's score to represent a nearly fully free media system. The last two years indicate a dip that the country has not seen since the late-1980s: .7174 in 2016 and .7034 in 2917.

In sum and as expected, Brazil's MSF scores reflect its media system case history. And contrary to Nigeria, the country's media freedom score moves with the authoritarian regime—as the military regime took power MSF scores decrease and as the regime opened the score steadily rise. To show this relationship, we graph Brazil's polity and MSF scores in figure 15 below.

[24] *Committee to Protect Journalists* Carlos Lauría "Attacks launch: Judicial censorship strikes a chord in Brazil," Feb. 17, 2011. Available at: https://cpj.org/blog/2011/02/at-attacks-launch-judicial-censorship-strikes-a-ch.php

[25] *Committee to Protect Journalists* Carlos Lauría "Violence and Judicial Censorship Mar Brazil's Horizon," 2013. Available at: https://cpj.org/2014/02/attacks-on-the-press-brazil-analysis.php.

Figure 6: Brazil's Polity and MSF Scores, 1948-2017



Note: Polity standardized to 0 - 1 scale.

E.2.2 Data-blinded Case Studies by RA

To provide further validation for our new measure, we conduct a data-blinded case study for two countries in our sample. While the case studies above provide strong convergent validation for the data that our measurement models produced, we recognize that knowing the trends before conducting the case studies may have induced some bias into the results. We therefore enlisted a student Research Assistant (RA) to conduct data-blind case studies on randomly selected countries from our sample. We present the instructions for the procedure and the results.

Procedure

[We provided the RA the following instructions]

Please complete a case history on media freedom for two (2) countries by following the steps below.

Step 1: Case Selection

Criteria for countries so that the Research Assistant (RA) has enough material to make an adequate evaluation:

- At least 30 years old;
- Population should be larger than 1 million during most of the time series; and
- Academic/historical material that covers all time periods in the Media System Freedom (MSF) data readily available.

Researchers will randomly select five countries for potential analysis.[26] The RA should start at the beginning of the following list and determine if that country meets the criteria.

1. St. Vincent and the Grenadines
2. Jamaica
3. Armenia
4. Libya
5. Zimbabwe

Step 2: Evaluate country's media freedom

a) Spend some time reading about the country's general history during the sample.

b) Evaluate each country based on this definition of media freedom:

[26] We randomly grab four rows from the dataset, as each row is a unique country using the following R code: *ccode[sample(nrow(ccode), 4),]*.

The ability for media (including journalists, photo journalists, editors, opinion writers, and other media personnel) to produce content without undue influence.

We note that undue influence can stem from government and non-government actors including opposition parties, criminal organizations, guerilla/terrorist groups, and advertisers.

c) RA should determine media freedom level in 1948 or oldest year for the selected country (for instance, Azerbaijan would begin in 1991) based on a scale from 1 to 3 for the level of media freedom from RA's perspective.

- 1 = Low Media Freedom
- 2 = Medium-level Freedom
- 3 = High Media Freedom

RA should try to the best of their ability to map this score over the years, noting when they believe the score changes over time. They should also record noticeable increases or decreases that do not necessarily result in a rank change.

For example, if RA determines *Country A* to be a 3 in 1948, then the RA should investigate if this rating generally holds through time. When the RA identifies a change, they should mark when it changes and to which other level. If there are slight improvements or regressions that do not necessarily result in a rank change in the RA's estimation, then please note those as well. NOTE: RA may begin their research but determine that the country does not meet condition 3 from Step 3. If so, then please move to the next country.

d) In their investigation, the RA should not include sources from watchdog groups that also provide data for our media freedom measurement model: *Freedom House* and *Reporters without Borders*. Also, avoid the book *Historical Guide to World Media Freedom: A Country-by-Country Analysis* by Jenifer Whitten-Woodring and Douglas A. Van Belle, as it is also a source of data for the measurement model. The RA should begin their search in the following places:

- Google Scholar
- Academic books from library

If the RA is unsure about a source or about anything during this step, please contact a Researcher for assistance.

Step 3: Debrief interview

[After the RA finished their research, we met them for a debriefing interview. We asked them the following questions:]

1. Please walk us through your findings beginning with the first country on the list that meets the criteria in Step 1.

2. Do these results [MSF scores] mirror your own expectations? Are you surprised by these results?

3. If the results were divergent from your own in any way, why do you think this was the case?
4. After researching each country's case history, do you think our approach to measuring media freedom has a degree of face validity, in other words is it intuitive from your perspective?
5. Did you find enough data to make a reasonable assessment?
6. *Asked after both countries reviewed* Do you have any questions for us regarding our measure or your participation in this exercise?

Results

Below we detail the debriefing interview that we had with the RA. Her findings generally corroborate the data and show a high level of convergent validity.

St. Vincent and Grenadines

Does not meet the criteria outlined in *Step 1*.

Jamaica

Step 3, Question 1: Jamaica begins the sample in 1962 after it gained independence from Great Britain. The RA found Jamaica to have higher levels of media freedom, though certain periods reflected a slight decrease. Specifically, the RA noted that decreases occurred in the late-1960s and early 1970s due to a series of state emergencies. She also noted instances of government institutions protecting media freedom, such as a 2001 law that ensured freedom of information as well as legislation that repealed a law that criminalized defamation in 2017. According to the RA, "On the whole, Jamaica has high levels of media freedom, despite the newspapers' frequent engagement with politics."

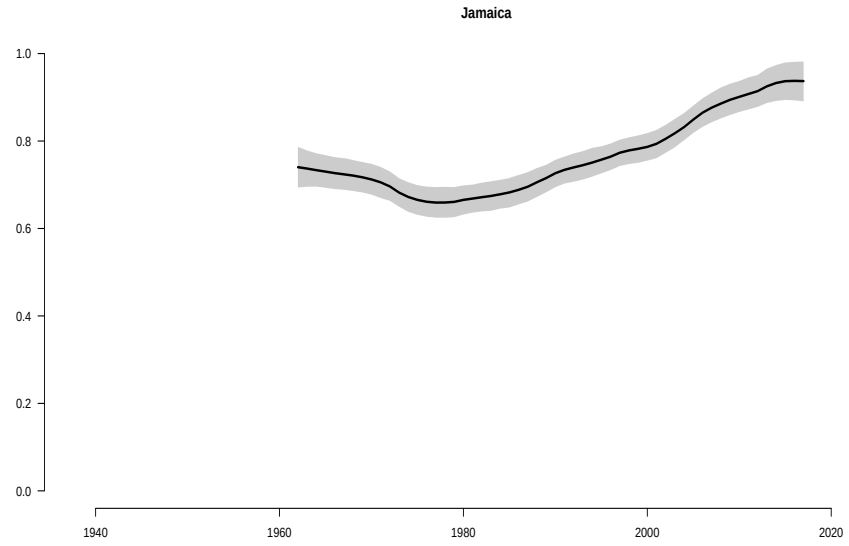
Step 3, Question 2: When we showed her the data (see figure 7 below), she was not surprised. While the media freedom levels seemingly decreased in the late-1960s and early 1970's, they only did so incrementally (in accordance to her findings).

Step 3, Question 3: n/a

Step 3, Question 4: The RA generally agreed with the face validity of the data, noting that the dip in the late-1960s and early-1970s fits with her findings.

Step 3, Question 5: The RA stated that she had enough material to make a reasonable assessment of Jamaica's media freedom during the sample.

Figure 7: Jamaica's MSF Scores, 1962-2017



Armenia

Step 3, Question 1: Armenia became an independent country when the Soviet Union dissolved in 1991. The RA generally found it had a medium level of media freedom for much of the sample, with instances of dips to a low-medium categorization. The RA states, "Among the former Soviet republics, Armenia has higher than average media freedom; however, it still faces challenges stemming from that time period." She generally described the country as attempting to guarantee media freedom through legislation, however the laws were vague and ineffective at guaranteeing media freedom. For instance, the Armenian legislature passed an amendment in 1995 guaranteeing media freedom, but the verbiage was generally vague and uninformed. In 2008, the legislative body would pass another amendment to improve the 1995 one on media freedom. In addition, a freedom of information act passed in 2003, though it did not come into effect until 2015. She also notes that defamation was decriminalized in 2010.

However, she did note events in the early 2000s that affected media freedom in the country, including the shutdown of opposition news channel *A1+*, as well as several high profile killing of journalists including Tigran Nagdalyan, chairman of Armenia's Public Television and Radio

in 2002. She concludes, "While Armenia has made—and continues to make—a lot of progress compared to other former Soviet republics, it still has medium-low levels of media freedom, due mainly to the self-censorship of media outlets."

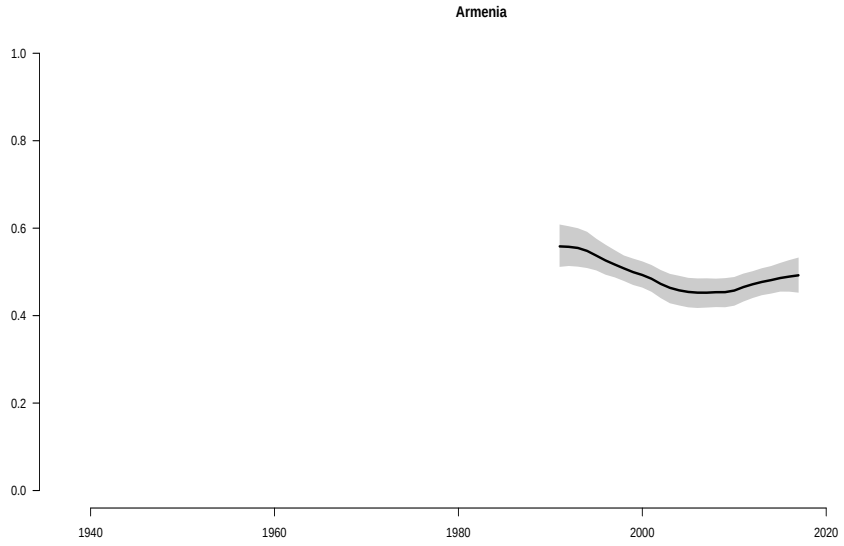
Step 3, Question 2: When we showed her Armenia's data (see figure 8 below), she thought it generally reflected her findings.

Step 3, Question 3: She did think that around the early 2000s the data was closer to low-medium, but that potentially the confidence intervals covered this.

Step 3, Question 4: Even given this slight discrepancy, she did feel the data were generally valid.

Step 3, Question 5: The RA stated that she had enough material to make a reasonable assessment of Armenia's media freedom during the sample.

Figure 8: Armenia's MSF Scores, 1991-2017



Summary

In sum, the data-blinded case studies reveal that our data follow a general pattern of media freedom that reflect the case history. While it remains difficult to pinpoint exact continuous numbers, the trends from the data-blinded case history generally reflect our new data; Indicating the data carry a high level of convergent and face validity. We note that when we asked the RA *Step 3, Question 6*, she did feel the definition of media freedom we provided was a bit vague but that this

did not impact her work. In future iterations, we would have let the RA read the entire paper that better outlines the latent concept of media freedom and/or have a pre-meeting where we verbally describe and discuss the concept of media freedom as we define it in the paper.

E.2.3 Empirical Evaluation of Convergence

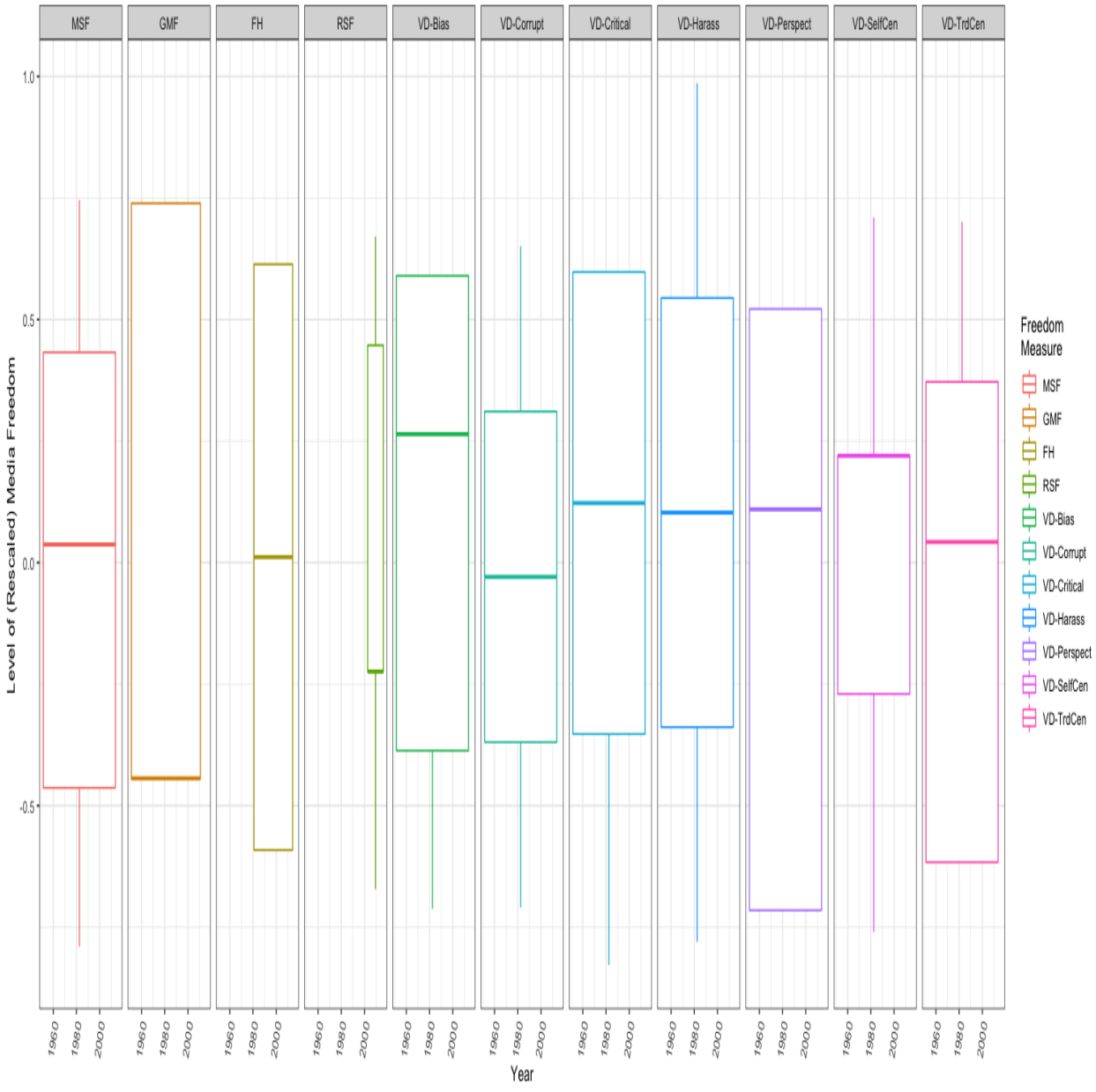
Where multiple coders exist, to what extent do they converge? (McMann et al., 2016, pg. 6)

This section provides further evidence of convergent validity using empirical analysis. First, we present box plots of the manifest variables compared to our new MSF score in figure 9 below. We rescale the measures by centering on zero and dividing by two standard deviations (Gelman, 2008). This method is efficient for both showing and comparing various distributions, which includes more information than individual country plots. As all of these measures are distributions in their own rites, we suggest this is an efficient method for direct comparison across measures in all countries for all years for which there are data. Note that the x-axis includes the full range of data beginning in 1948. However, limited space forces us to set the tick marks every 20 years, from 1960, 1980, and 2000. This figure indicates that our measure both reflects and also disagrees with the other manifest variables, both of which are valuable points in assessing and demonstrating the validity of our measure.

Regarding agreement in comparison to the other measures, the distributions are mostly centered around zero, with similar inter-quartile ranges. Interestingly, the median values are useful proxies for capturing the disagreement across the measures. For example, while we might expect the V-dem measures to possess similar median scores hovering around zero, our measure, the *RSF* measure, and *Freedom House* measure each trend slightly more negative. This also suggests a world with generally less free media than the world the V-Dem measures portray. This suggests that the similarities and disagreements across the measures supports our IRT approach, where our measurement strategy acts as a "smoother" to synthesize the differences and similarities (hence loading individual items onto a single latent dimension). The result is a *capturing* and accounting for these similarities and differences, rather than treating each measure—and thus their differences

and similarities—as isolated quantities of interest. As such, the consistency we find across the measures points to high *convergent* validity, as outlined in the techniques for assessing convergent validity laid out by McMann et al. (2016). Still, the differences that exist across the measures point to the value of our IRT approach, but also acts as a guard against overfitting. If our measurement patterns perfectly match the trends of other existing indicators, it would suggest we are essentially replicating other measures, instead of generating a generalizable measure that captures valuable nuance in media freedom broadly defined, including both similarities *and* differences. In sum, by placing all of these measures (ours and the manifest variables) on the same scale and then showing each distribution side by side as boxplots, our measure appears to retain a high degree of convergent and even face validity.

Figure 9: Box Plot to Compare MSF and Manifest Variable Distributions



In addition, we follow McMann et al.'s (2016) approach to check convergent validity and provide scatter plots for our new measure compared to the 10 manifest variables. We present these scatter plots in figure 10 below.[27] Like figure 9 above, we find a high level of convergence. All 10 manifest variables correlate to our MSF measure at .82 or higher, with all V-Dem variables yielding a correlation coefficient of .9 or higher. Both the box plots and scatter plots reveal that our MSF converges well with the other media freedom variables, revealing our new measure closely resembles the concept that these variables wish to capture.

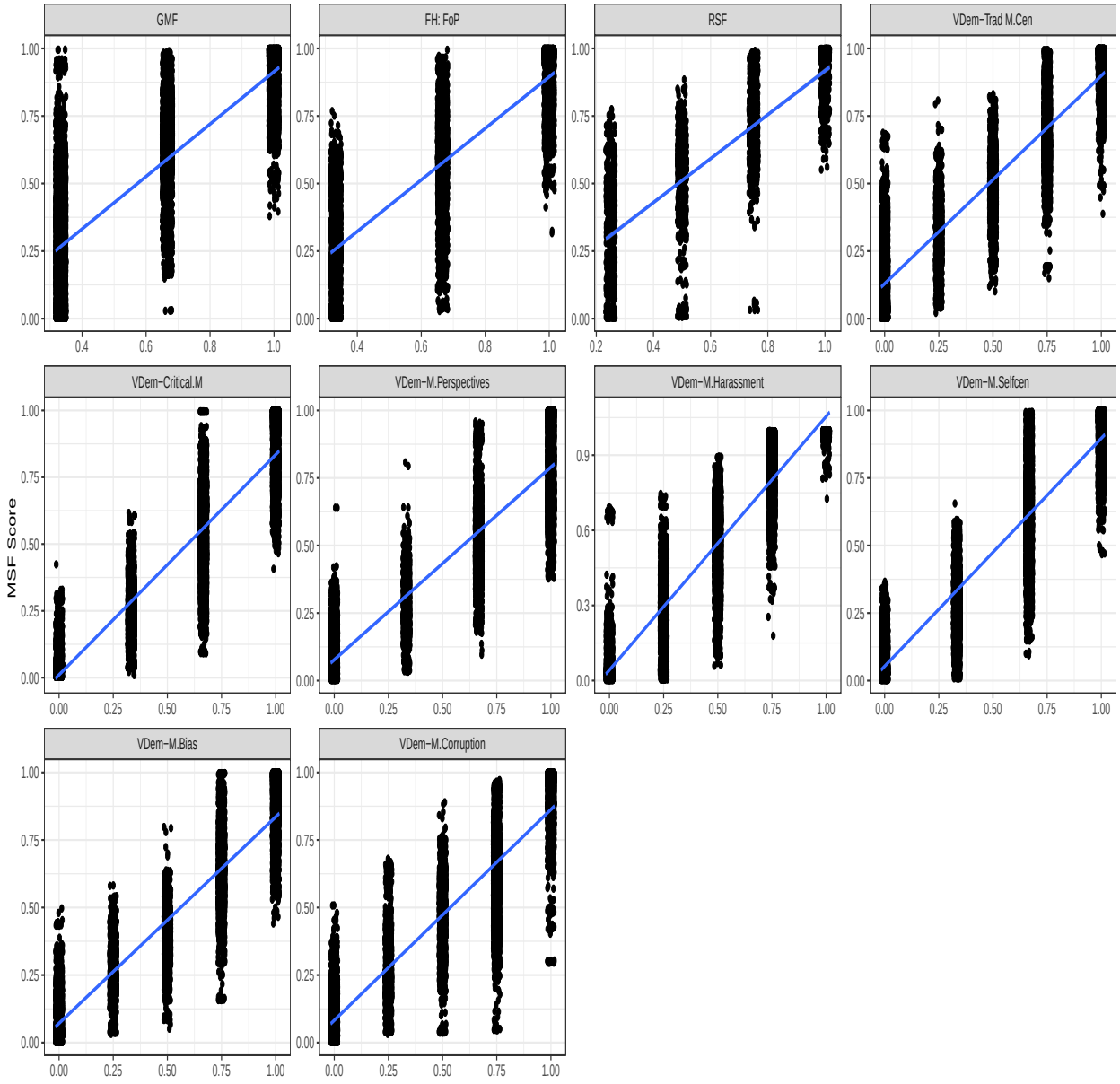
To what extent do the data produced by the measure correlate with data produced by other measures of the construct, and are areas of low correlation thoroughly understood? (McMann et al., 2016, pg. 6)

McMann et al. (2016) warn researchers that high correlations between a new measure and extant ones alone do not necessarily establish validity.[28] Indeed, figure 10 shows that some observations stray a bit from the fitted line and have higher residuals than others. If these higher divergences are systematic due to rater bias, then we have grounds to question the measure's validity. We therefore follow McMann et al.'s (2016) suggestion and investigate if any systematic biases exist among the raters. Because McMann et al. (2016) evaluate the bias of survey respondents—which varies significantly from our data generating process, we must take a different route to explore systematic bias in the data. Though we do not possess individual level data for our dataset, we treat each manifest variable as a respondent and investigate systematic bias among these raters for cases that diverge from our new measure.

Our approach first identifies country-years with sufficiently high levels of disagreement. To develop a measure for rater disagreement with our MSF score, we fit a simple regression model between the MSF score and each of the 10 manifest variables (see figure 10 above). From here, we extract the residuals and mark country-years where disagreement between our measure and a manifest variable was significantly high. We consider disagreement significant if the residual is two standard deviations above or below the residual's mean for that regression. From here,

[27] All manifest variables converted to a 0-1 scale to ease interpretation.
[28] Their warning is actually a paraphrase of Hawken and Munck (2009, pg. 4).

Figure 10: Scatter-plots w/ Fitted Line: MSF and 10 Manifest Variables



Note: Manifest variables standardized to 0 - 1 scale.

we construct a dataset that identifies country-years with significant rater disagreement with our measure. We also include a variable that counts the number of times significant disagreement occurred. We note that of the 89,090 total country-years in all 10 regressions, only 3,929 (or about 4.4%) country-years logged a significant level of disagreement. Most countries in most regressions had at least one significant disagreement throughout its sample, with the most being

four. The following country-years had four significant disagreements: Colombia 1948, Seychelles 1976, Nepal 2000, Pakistan 2004, Gambia 2017, and Oman 2017. Singapore had by far the most total country-year disagreements with 122 (or about 25% of possible country-years). However, this case is hardly typical. Gambia—the next highest—also contained 76 significant disagreements (or about 15% of possible country-years). In total, the mean number of disagreements is about 19 while the median is about 15 across 10 variables for each country's total country-years, indicating these agreements are generally uncommon.

From these data we create two separate variables: 1) a binary that indicates a country-year contains a significant disagreement, and 2) the count of significant disagreement in that country-year. We then we construct other variables based on traits of the dataset that may indicate a rater's (in our case, a manifest variable) bias. We first create a variable for the number of manifest variables a country-year carries. For example, if a country-year is missing all seven V-Dem variables, this data point would be *3* ($10-7=3$). We also create a dummy for country-years where V-Dem variables are missing. Next, we create a dummy for country-years where non-academic "watchdog" groups like *Freedom House* and *Reporters Without Borders* contribute data. Because these groups' seek to advocate for media freedom (rather than dispassionately study it), they may approach rating media freedom in a less rigorous and more visceral manner, perhaps inducing a bias. Finally, we include region dummies as well.[29] The measure should indicate whether certain regions contain more systematic disagreement than others. Finally, we include the posterior standard deviation that the measurement model produces. This variable captures the level of uncertainty for each country year among the raters.

We run logit and ordered logit models on our two variables. For each model, we include year effects, as well as cluster the standard errors by country. We use Western Europe and North America as the base category for the region variables. This will indicate if non-Western European countries carry more disagreement than Western European and North American countries. We

[29]We draw these data from Teorell et al. (2017, pg.357). While they provide 10 categories, we collapse them to 6 regions: 1) Post-communist countries, 2) Latin America and the Caribbean, 3) Middle East and North Africa, 4) Sub-Saharan Africa, 5) Western European and North America (including Australia and New Zealand), 6) East Asia and the Pacific, and 7) South and South East Asia.

present the results in table 10 below. For both models we find that all variables are insignificant, with the exception of *Posterior St.D.*—the measure of uncertainty among raters. Higher levels of uncertainty increase the probability that a country will see disagreement between our new MSF score and the manifest variables. The results also indicate that non-Western European countries do not see more or less disagreement than Western European countries. In sum, while cases like Singapore and Gambia may carry higher levels of significant disagreement than other countries, the analyses indicate that 1) there likely is not systemic traits from raters (that is, the manifest variables) that bias the data, and 2) that if researchers utilize the posterior standard deviations provided by the IRT model in empirical analysis, then they can assure the validity of their results—though the evidence indicates the data are generally valid without them.

Overall, these data indicate a high level of convergent validity of the MSF measure. The level of coder disagreement is relatively small, but when disagreement does occur we find no systemic rater traits to account for this variance—uncommon as it is.

Table 10: Explaining Divergence, Logit and Ordinal Regression Models

	(1) Logit	(2) Ordinal Logit
Posterior St.D.	29.424***	27.783***
	(7.028)	(7.386)
V-Dem Missing	-1.340	-.848
	(2.590)	(2.353)
Manifest Variables	.179	.241
	(.354)	(.322)
Watchdog Rater	-.216	-.197
	(.598)	(.539)
region dummies		
Latin America/Carib.	-.231	-.178
	(.277)	(.275)
MENA	.299	.310
	(.260)	(.251)
Sub-Saharan Africa	.246	.292
	(.233)	(.225)
East Asia-Pacific	-.079	-.067
	(.586)	(.555)
South/South-East Asia	.120	.176
	(.379)	(.401)
Countries	197	197
N	10,739	10,739

Standard errors in parentheses (clustered by country); Year effects and Constant not reported;
Western Europe and North America base category for region; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

F Robustness Checks: Alternative IRT Estimations

To check the robustness of our new measure, we estimate the IRT model using different specifications. First, we perform different estimations by down-weighting the seven V-Dem variables. We do this by replace them with 1) another V-Dem variable, the Freedom of Expression and Alternative Source of Information (FEASI) index, and 2) scores from a Bayesian Factor Analytic (BFA) model of the seven V-Dem variables. Second, we run several IRT models incorporating uncertainty from the seven V-Dem variables from our main model. Next, we estimate an unconstrained, IRT model that removes the 0-1 scale constraint allowing it to vary freely. Finally, we run seven subset IRT models including the three *non*-V-Dem variables with a single V-Dem variable to compare the discrimination parameters to the ones in our full model. Overall, we find that these checks confirm that robustness of our initial data and results.

F.1 Downweighted IRT Estimations

While all the variables we use fall within our media freedom theoretical framework, seven of the 10 manifest variables originate from the same data source—Varieties of Democracy. Here, we investigate if their inclusion over-weights them in the resulting index. First, we replace our seven V-Dem measures with V-Dem's Freedom of Expression and Alternative Sources of Information (FEASI) index. According to V-Dem, the guiding question of this variable asks:

To what extent does government respect press and media freedom, the freedom of ordinary people to discuss political matters at home and in the public sphere, as well as the freedom of academic and cultural expression?

While this measure is continuous from 0 to 1, we create categorical variables based on the variable's quartiles in order to include it into the graded measurement model. Second, we run a BFA model on the seven V-Dem variables in our measurement model, then extract the resulting scores. Like the V-Dem model, we create ordinal variables based on the variable's quartiles. Once we create these two variables, we run two separate measurement models that replaces them with the original seven V-Dem variables. After running these separate models, we extract the index and

compare them to the manuscript MSF score.

Table 11 below shows the Pearson's correlation coefficients between the original MSF scores and the scores from these two new specifications. The results indicate the down-weighted variables correlate highly at .94 or higher. As a point of comparison, we also provide figures for the eight case studies in figure 4 in the manuscript. We find the trends remain largely the same, though they seem to not detect subtle changes in the trends that the case studies would suggest. For instance, the case studies suggest that the United States did not have very high levels of media freedom until a series of landmark Supreme Court decisions in the late 1960s-early 1970s. However, when we include the FEASI variables, we see the United States has high levels and entirely smooth throughout—outside of expectations of the case study. We also see that this measure shows a bump for the 1980s Solidarity movement, but does not decrease back to near 0 levels like the case study suggests.

Table 11: Correlation Coefficients: Compare MSF Scores to Down-weighted Specifications

Dataset	Correlation Coefficient
IRT w/ BFA Scores	0.9458
IRT w/ V-Dem FEASI Index	0.9638

While the two alternative scores highly correlate with our MSF scores, we explain here why we prefer our original model. For the BFA scores specification, the trends in figure 12 smooth and seemingly lose variation in some cases. We suspect that reducing the data through the BFA process *then reducing it again* suppresses much of the rich variation that the seven separate V-Dem variables provide. We find evidence to support this when comparing the measurement model's cutpoints in figure 11. The figure indicates that BFA index does about equally well at determining higher and lower levels of media freedom as the *Reporters Without Border* data. This is in stark contrast to the V-Dem's variables much better ability to discriminate high and low levels of media freedom in the cutpoint figure in the manuscript (fig. 5). We therefore prefer our 10-variable model over this specification.

Next, we also find similar trends from the FEASI model—also created from a BFA estimation

of various indicators by V-Dem—in figure 14. Though again, some trends diverge from what the case study suggests. We also have theoretical reason for not including this variable. While freedom of expression and media freedom are similar concepts that overlap to some extent, they are not necessarily the same thing. The variable does include some of the variables we include in our measurement model, it does not include all (for example, *media corruption*). In addition, it include some components that exist beyond the bounds of our theory, such as indicators that measure freedom of cultural and academic expression and freedom of discussion. Freedom of expression as a whole represents a broader concept that includes individuals and academic expression. Media freedom expressibly deals with freedom of expression but for the press and other professionals in media. Therefore, this variable does not theoretically map to our conceptualization of media freedom. We suspect its inclusion would decrease content validity by failing to differentiate irrelevant concepts from relevant ones.

Finally, we emphasize that including the seven V-Dem variables—though from the same data generators—is a strength that allows us to capture more variation in media system freedom. Each variable examines related but *still different* aspects of journalists' ability to produce content without undue influence from state and non-state actors. So asking a rater to examine, for example, government censorship attempts will trigger a different mindset in evaluating the media system than if asked to evaluate harassment of media personnel of journalists' self-censorship. Again, we reiterate that including these seven variables *strengthens* our research design.

Overall, we prefer our measurement model over these down-weighted specifications. This decision comes from both empirical but also theoretical reasoning.

Figure 11: Threshold Estimates for *Non*-V-Dem and BFA Scores Specification

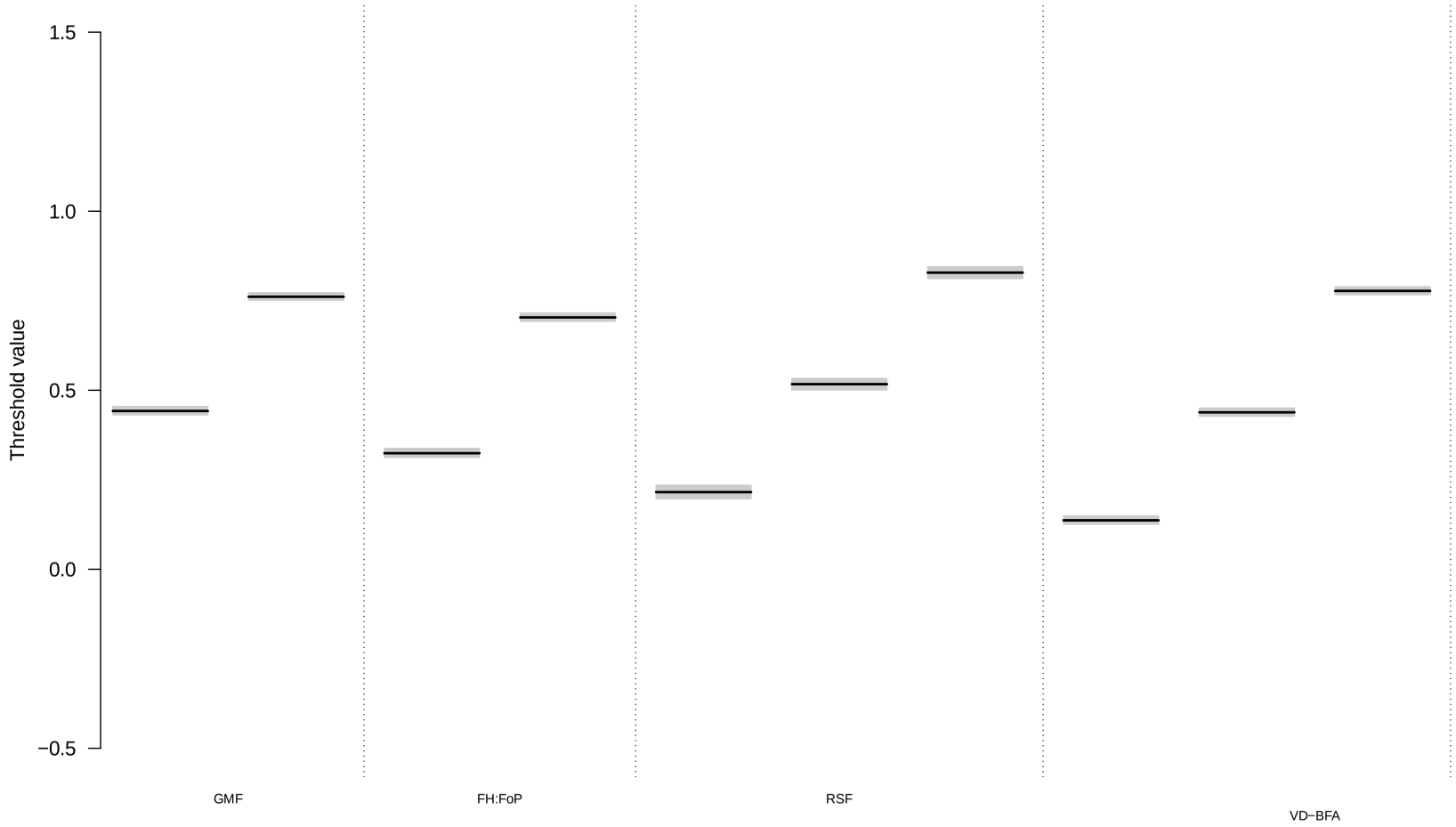


Figure 12: Eight Countries, IRT Specification w/ BFA Scores

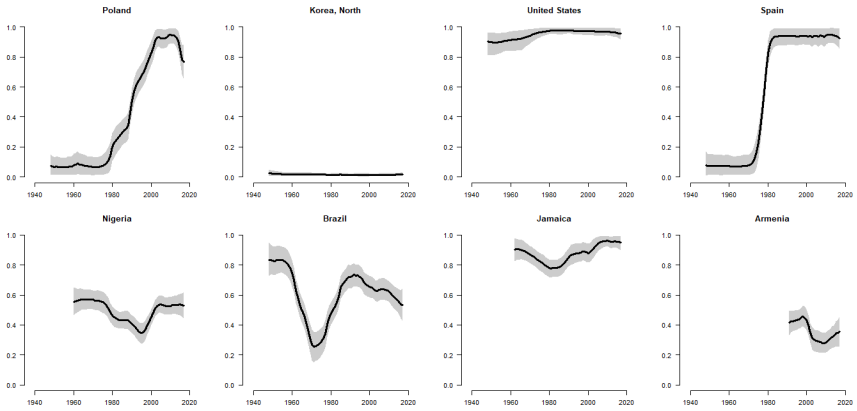


Figure 13: Eight Countries, IRT Specification w/ FEASI Index

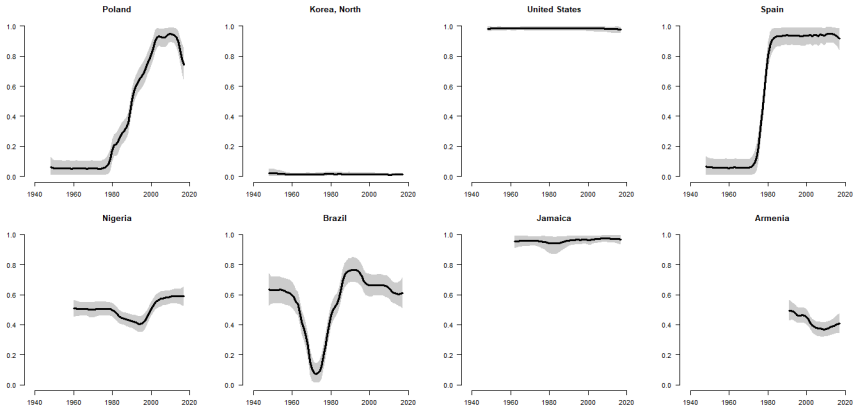
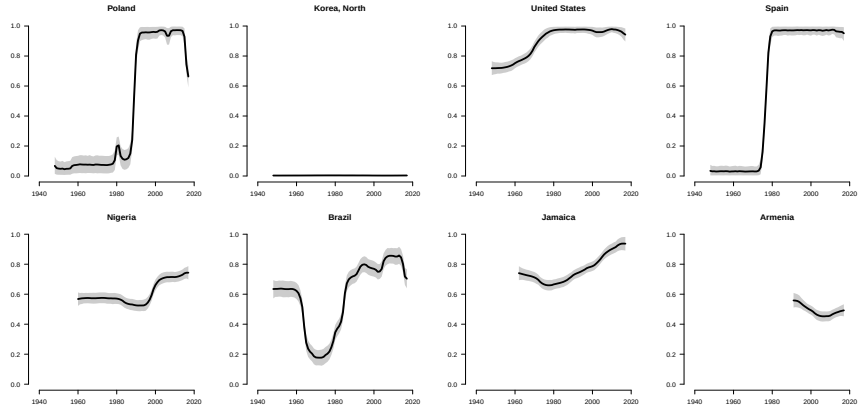


Figure 14: Eight Countries, Manuscript MSF Score



Note: Figure 4 in manuscript; Manifest variable jitter plots removed.

F.2 Incorporate V-Dem Uncertainty

An important assumption of the IRT model is that raters do not perfectly measure the latent variable. The model presents a statistically principled way to synthesize variables that attempt to measure the same or similar concepts. Therefore, the measurement model would adequately handle variation among the manifest variables in measuring the laten variable.

However to test this presumption, we utilize V-Dem's uncertainty measurement for the ordinal variables. For Version 8 of the V-Dem measure, the raters include *code high* and *code low* variables. These variables are exactly one standard deviation above and below the point estimate of the variable's distribution that generates the ordinal variables. In total, these variables cover 70% of the measurement's distribution (Coppedge et al. 2018c). From here, we use the range from these variables to randomly draw numbers for each country-year from a uniform distribution 10 times to create 10 random datasets.[30] We note that *non*-V-Dem variables remain unchanged in all 10 new datasets.

We then run these 10 datasets through our measurement model. Next, we pull the 10 alternative MSF scores including measure uncertainty and compare them to our original MSF score via correlation analysis. In table 12 below we show the Pearson's correlation coefficient of each of their 10 datasets. The results indicate that even when we include uncertainty, these data highly correlate with the manuscript's output. Again, this makes sense given the model assumes raters do not perfectly evaluate media freedom and estimates the model accordingly. We take these results to show the measurement highly valid even when considering V-Dem's uncertainty measures.

[30] Because each draw produces a continuous number within the specified range, we round up or down at .5 to create an integers for our graded, IRT model.

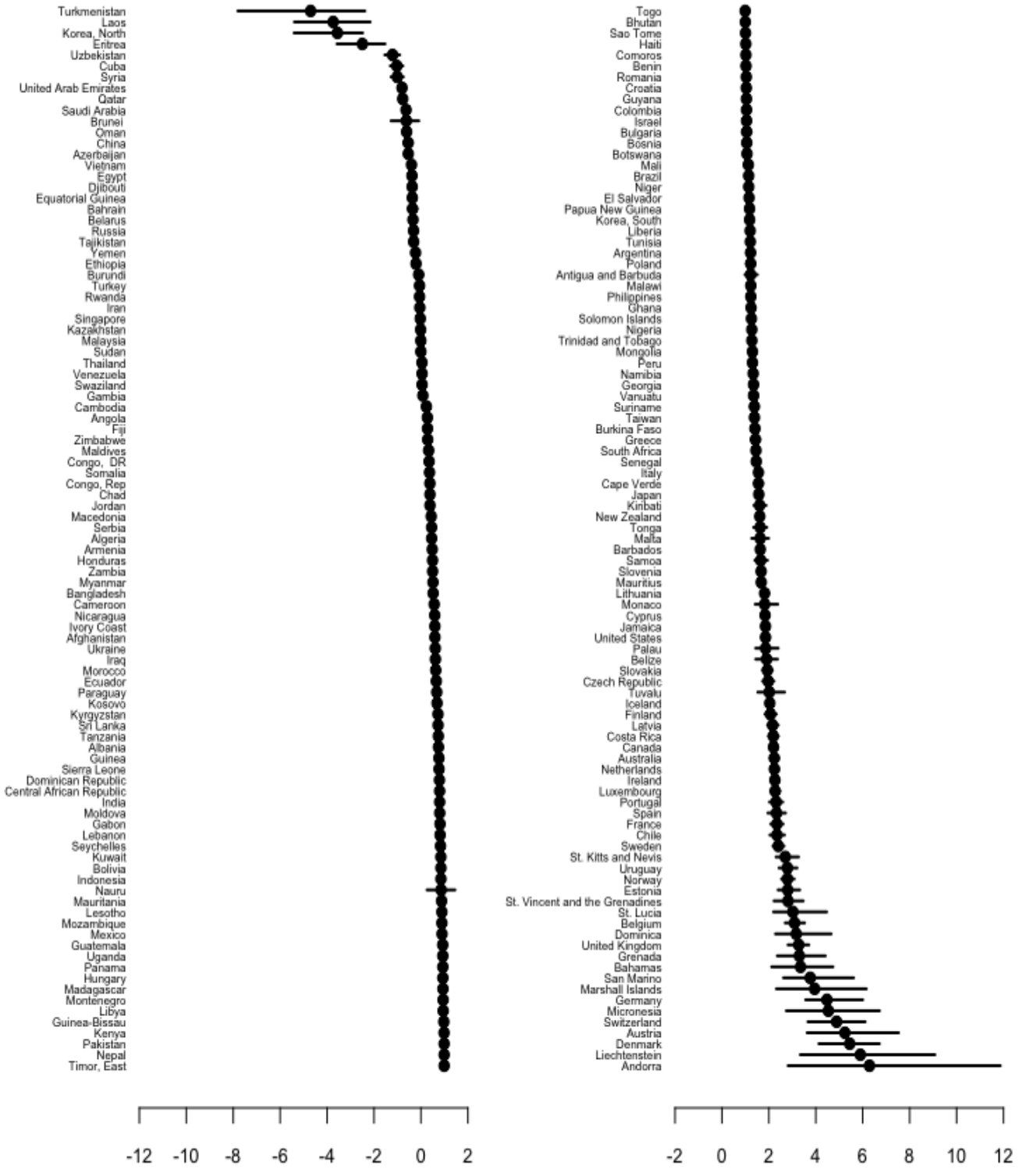
Table 12: Compare MSF Score to 10 Datasets w/ Random Draws from V-Dem Measure of Uncertainty

Dataset	Correlation Coefficient
Random Draws 1	0.9967
Random Draws 2	0.9967
Random Draws 3	0.9966
Random Draws 4	0.9968
Random Draws 5	0.9968
Random Draws 6	0.9968
Random Draws 7	0.9968
Random Draws 8	0.9967
Random Draws 9	0.9967
Random Draws 10	0.9967

F.3 Unconstrained Model

We also fit an unconstrained standard normal prior, thereby letting values vary freely along the latent dimension. Notably, while we see a few minor shifts in country rankings of freedom compared to our original constrained model. We also do not see any substantive shifts in levels of media freedom—Pearson's correlation coefficient is .9023. For example, when we compare the cross-section estimate from the constrained model in the manuscript's figure 2, of the five *least* free counties in the constrained model (North Korea, Turkmenistan, Oman, Uzbekistan, and Laos), four are the same in the unconstrained model (Turkmenistan, Laos, North Korea, and Uzbekistan). Note that the unconstrained model did not include Oman, and instead included Eritrea, which was ranked sixth from the bottom of least free. Similarly for the other extreme, where of the five most free countries (Germany, Austria, Belgium, Switzerland, and Denmark; from 5th to most free), three are in the top five of media freedom in the unconstrained iteration (Switzerland, Austria, and Denmark), with Germany ranked seventh and Belgium ranked 14. While small variance in ranking would exist each time the model is fit—such is the nature of Bayesian and Markov Chain Monte Carlo processes (drawing from distributions), we suggest that the constrained and unconstrained models reveal corroborating trends. Ultimately, this demonstrates that we are not seeing substantive differences across countries or deviations from the two specifications, but rather merely different values for media freedom (hence the *un*constrained). Given this corroborating evidence, we opt to keep our constrained version of the model in the manuscript—in line with past approaches doing a similar approach (for example, Linzer and Staton (2015)).

Figure 15: Unconstrained Media System Freedom (MSF) Scores, 2016: Lower and Upper Bounds



Note: Unconstrained score correlate to constrained model at .9023.

F.4 Compare Subsets of IRT Models: Discrimination Parameters

While we find that V-Dem manifest variables had larger discrimination parameters than *Freedom House*, *Reporters Without Borders*, and *Global Media Freedom* indicators, we fear that V-Dem's overrepresentation in the measurement model might overestimate discrimination parameter sizes. To check for the robustness of this result, we compare the discrimination parameters from the main IRT analysis to various subsets of the main analysis below. We first present the results from the main analysis from the manuscript (pg. 19), then the results from running the IRT model when only including one V-Dem variable at a time. We present the discrimination parameters from these various models below in table 13. Consistent with our initial findings, the V-Dem variables yield higher betas than the *Freedom House*, *Reporters Without Borders*, and the *Global Media Freedom* variables in each subset model.

In some cases, V-Dem's betas in the subsets vary from the betas in the full model. These include *government censorship efforts*, *media harassment*, and *media corruption*. *Media corruption* actual jumps from having the lowest beta in the full model to having the highest among the subsets. While this check provides further evidence that the V-Dem variables contribute more to the latent measure, we consider estimates in the full model that includes all 10 variables as a more principled reflection of their influence on the latent measure. Therefore, we highlight here that V-Dem remains higher in all subset analyses but hesitate to revise our initial inferences in the manuscript regarding individual variables.

Table 13: Compare Discrimination Parameters, Subsets of Main IRT Analysis

Manifest Variables	Main Analysis	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Subset 6	Subset 7
GMF	9.235226	9.903725	11.367539	10.879433	10.839149	10.921006	11.114432	10.358217
Freedom House	8.447979	9.285147	9.981942	9.666153	9.572078	9.397278	9.967244	9.42278
RSF	6.991127	7.1949	7.305385	7.849975	7.482926	7.460341	7.551094	7.418057
Vdem: Trd M.Cen	11.467051	12.421802						
Vdem: M.Critical	11.83162		11.649126					
Vdem: M.Perspectives	11.535464			11.571137				
Vdem: M.Harassment	11.348156				13.108052			
Vdem: M.Self-Cen	11.400918					11.664744		
Vdem: M.Bias	12.044261						12.072614	
Vdem: M.Corrupt	10.566969							13.704833

G Markov Chain Monte Carlo Convergence Diagnostics

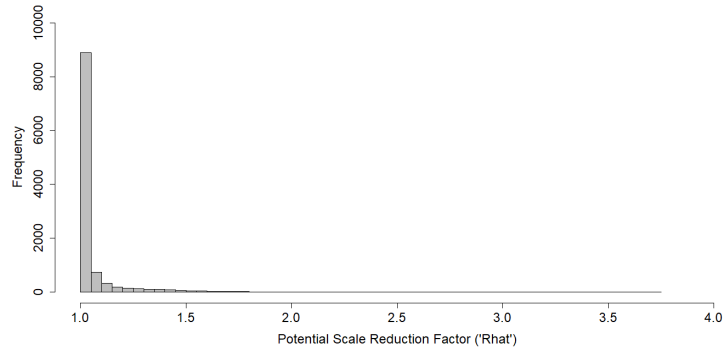
We check for convergence across every chain and for each quantity of interest by computing the potential scale reduction factor (or the $\hat{R}$ statistic) and present the results in figure 16a. In brief, $\hat{R}$ is computed by taking the average of both within and between distribution variance of the given chain. Commonly, $\hat{R}$ values ≈ 1 suggest these variances "mix" perfectly. Values greater than 1.1 suggest mixture is not fully realized (i.e., lack of convergence), requiring further simulation to strengthen estimates, and thus inferences (Gelman and Shirley, 2011). As such, note that in figure 16a, the vast majority of iterations have sufficiently low $\hat{R}$ values (i.e., $<$1.1), suggesting evidence of convergence across the many MCMC simulations. Intuitively, note also that the proportions of perfectly mixed iterations decreases as the threshold narrows. Specifically, between 1.0 and 1.09, $\hat{R} = 0.846$, between 1.0 and 1.05, $\hat{R} = 0.764$, and most narrowly, between 1.0 and 1.02, $\hat{R} = 0.604$. Therefore, with around 85% of 10,980 parameters yielding $\hat{R}$ values 1.1 or less, we are confident in the results from our MCMC procedure and that our model converges.

Though we are confident in this finding, we investigate $\hat{R}$ further by increasing the number of iterations per chain to 30,000. Under this specification, we find that 96.89% of iterations yield $\hat{R}$ values 1.1 or less. We again increase the number of iterations to 50,000 per chain. Under this specification, 98.37% of iterations yield $\hat{R}$ values 1.1 or less. Finally, we increase the number of iterations to 70,000 per chain. Under this specification, 99.41% of iterations yield $\hat{R}$ values less than 1.1.[31] We note that the results of these increased iterations result in a correlation of Media System Freedom (MSF) point estimates among all four models at $\rho = .99$, indicating the increased iterations do not affect the substantive outcomes of the MSF scores. The data also reveal that the largest $\hat{R}$ reduces from about 3.714 in the 2,000 iterations per chain model to about 1.242 in the 70,000 iterations per chain model. We graph these additional findings in figures 16b, 16c, and 16c.

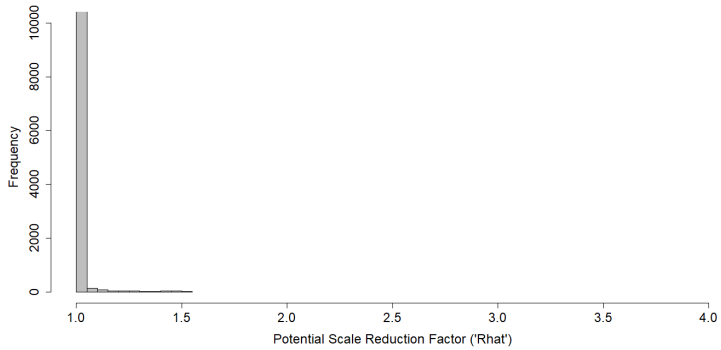
[31] Only 13 of the model's 10,980 parameters are $\approx$ 1.2, while all other ones are $\approx$ 1.1 or lower.

Figure 16: Histogram of Rhat Convergence Statistics, Various Specifications

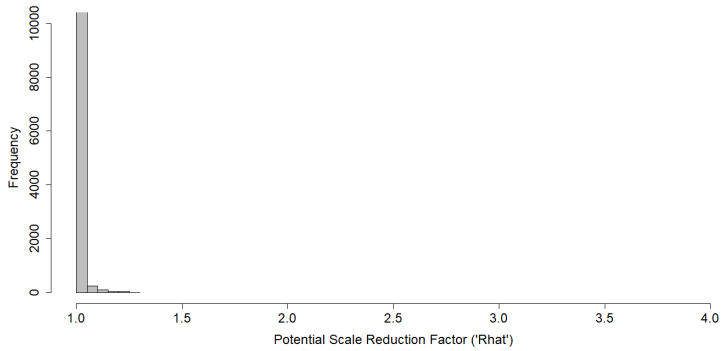
(a) 2,000 iterations per chain



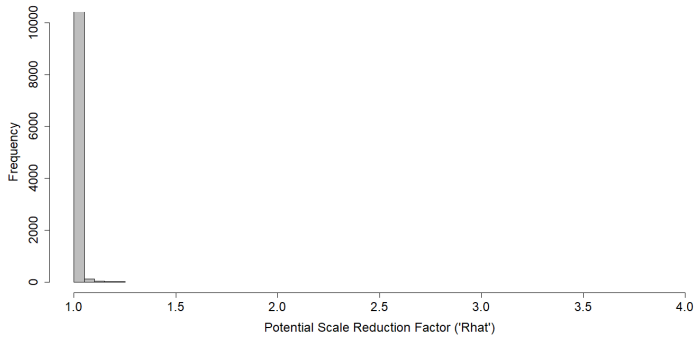
(b) 30,000 iterations per chain



(c) 50,000 iterations per chain



(d) 70,000 iterations per chain

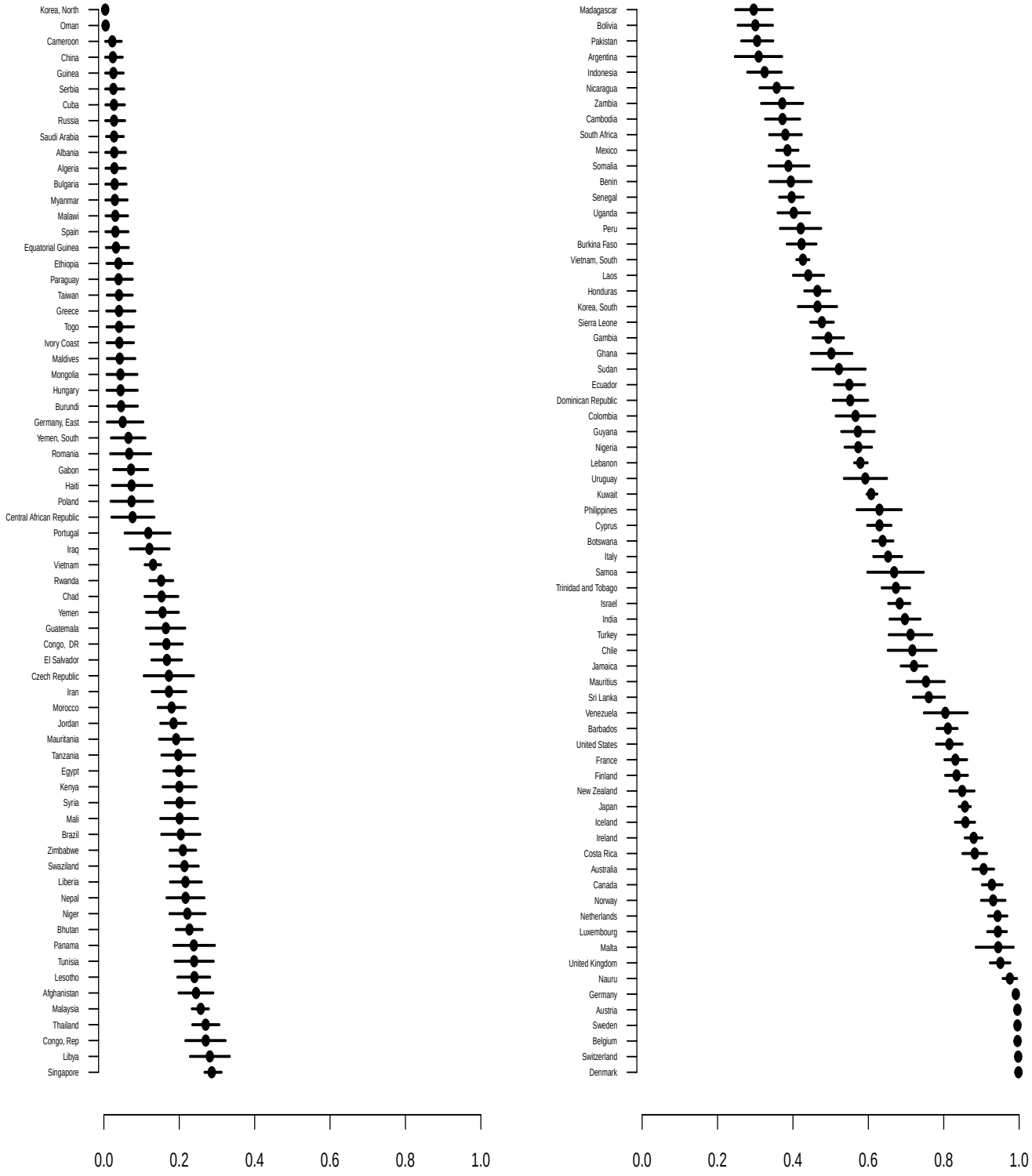


Note: Maximum $\hat{R}$ reduces as number of iterations increase from 3.714 (16a) to 1.535 (16b) to 1.272 (16c)to 1.242 (16d).

H Cross-sectional Results: MSF Scores in 137 countries, 1968

Figure 2 in the manuscript shows cross-sectional results for most countries in 2016. However, these figures do not show three countries with MSF data because they no longer existed that year. To provide results that include East Germany, South Yemen, and South Vietnam, we show cross-sectional results from the year 1968 in figure 17. South Yemen and East Germany rank near the very bottom, while the Republic of Vietnam ranks around the middle of all countries. Overall, the figure indicates that North Korea, Oman, Cameroon, China, and Guinea have the lowest media freedom scores in 1968, while Denmark, Switzerland, Belgium, Sweden, and Austria have the highest that same year. These results show similar countries at the very bottom and top of the MSF scores to the 2016 cross-sectional results. We note that in these figures, Russia indicates the Soviet Union, Serbia indicates Yugoslavia, Germany indicates West Germany, and Czech Republic indicates Czechoslovakia.

Figure 17: Media System Freedom (MSF) in 137 countries, 1968: Lower and Upper Bounds



Note: Media freedom point estimates along with associated error bars indicating 80% posterior credible intervals in 137 countries for 1968.

I Media System Freedom (MSF) Sample

In tables 14 and 15 below, we list each of the 197 countries in the Media System Freedom (MSF) dataset as well as the time each country covers. We note that in these figures, Russia includes the Soviet Union, Serbia includes Yugoslavia as well as Serbia and Montenegro, Germany includes West Germany, Yemen includes North Yemen, Vietnam includes North Vietnam, and Czech Republic includes Czechoslovakia

Table 14: Sample Summary: Countries and Time Coverage (Afghanistan-Liberia)

Afghanistan (1948-2017)	Chad (1960-2017)	Grenada (1974-2017)
Albania (1948-2017)	Chile (1948-2017)	Guatemala (1948-2017)
Algeria (1962-2017)	China (1948-2017)	Guinea (1958-2017)
Andorra (1993-2017)	Colombia (1948-2017)	Guinea-Bissau (1973-2017)
Angola (1975-2017)	Comoros (1975-2017)	Guyana (1966-2017)
Antig. and Barb. (1981-2017)	Congo, DR (1960-2017)	Haiti (1948-2017)
Argentina (1948-2017)	Congo, Rep (1960-2017)	Honduras (1948-2017)
Armenia (1991-2017)	Costa Rica (1948-2017)	Hungary (1948-2017)
Australia (1948-2017)	Croatia (1991-2017)	Iceland (1948-2017)
Austria (1948-2017)	Cuba (1948-2017)	India (1948-2017)
Azerbaijan (1991-2017)	Cyprus (1960-2017)	Indonesia (1948-2017)
Bahamas (1973-2016)	Czech Repub. (1948-2017)	Iran (1948-2017)
Bahrain (1971-2017)	Denmark (1948-2017)	Iraq (1948-2017)
Bangladesh (1971-2017)	Djibouti (1977-2017)	Ireland (1948-2017)
Barbados (1966-2017)	Dom. Repub. (1948-2017)	Israel (1948-2017)
Belarus (1991-2017)	Dominica (1979-2017)	Italy (1948-2017)
Belgium (1948-2017)	Ecuador (1948-2017)	Ivory Coast (1960-2017)
Belize (1981-2017)	Egypt (1948-2017)	Jamaica (1962-2017)
Benin (1960-2017)	El Salvador (1948-2017)	Japan (1952-2017)
Bhutan (1949-2017)	Eq. Guinea (1968-2017)	Jordan (1948-2017)
Bolivia (1948-2017)	Eritrea (1993-2017)	Kazakhstan (1991-2017)
Bosnia (1992-2017)	Estonia (1990-2017)	Kenya (1963-2017)
Botswana (1966-2017)	Ethiopia (1948-2017)	Kiribati (1979-2016)
Brazil (1948-2017)	Fiji (1970-2017)	Korea, North (1948-2017)
Brunei (1984-2017)	Finland (1948-2017)	Korea, South (1948-2017)
Bulgaria (1948-2017)	France (1948-2017)	Kosovo (2005-2017)
Burkina Faso (1960-2017)	Gabon (1960-2017)	Kuwait (1961-2017)
Burundi (1962-2017)	Gambia (1965-2017)	Kyrgyzstan (1991-2017)
Cambodia (1953-2017)	Georgia (1991-2017)	Laos (1953-2017)
Cameroon (1960-2017)	Germ., East (1948-1990)	Latvia (1991-2017)
Canada (1948-2017)	Germany (1949-2017)	Lebanon (1948-2017)
Cape Verde (1975-2017)	Ghana (1957-2017)	Lesotho (1966-2017)
Cen. Af. Rep. (1960-2017)	Greece (1948-2017)	Liberia (1948-2017)

Table 15: Sample Summary: Countries and Time Coverage (Libya-Zimbabwe)

Libya (1951-2017)	P. New Guinea (1975-2017)	Suriname (1975-2017)
Liechtenstein (1990-2017)	Pakistan (1948-2017)	Swaziland (1968-2017)
Lithuania (1991-2017)	Palau (1994-2016)	Sweden (1948-2017)
Luxembourg (1948-2017)	Panama (1948-2017)	Switzerland (1948-2017)
Macedonia (1991-2017)	Paraguay (1948-2017)	Syria (1948-2017)
Madagascar (1960-2017)	Peru (1948-2017)	Taiwan (1949-2017)
Malawi (1964-2017)	Philippines (1948-2017)	Tajikistan (1991-2017)
Malaysia (1957-2017)	Poland (1948-2017)	Tanzania (1961-2017)
Maldives (1965-2017)	Portugal (1948-2017)	Thailand (1948-2017)
Mali (1960-2017)	Qatar (1971-2017)	Timor, East (2002-2017)
Malta (1965-2017)	Romania (1948-2017)	Togo (1960-2017)
Marsh. Isls. (1986-2016)	Russia (1948-2017)	Tonga (1979-2017)
Mauritania (1960-2017)	Rwanda (1962-2017)	Trin. and Tobago (1962-2017)
Mauritius (1968-2017)	S. Africa (1948-2017)	Tunisia (1956-2017)
Mexico (1948-2017)	São Tome (1975-2017)	Turkey (1948-2017)
Micronesia (1986-2016)	Samoa (1962-2017)	Turkmenistan (1991-2017)
Moldova (1991-2017)	San Marino (1992-2016)	Tuvalu (2000-2016)
Monaco (1993-2016)	Saudi Arabia (1948-2017)	UAE (1971-2017)
Mongolia (1948-2017)	Senegal (1960-2017)	Uganda (1962-2017)
Montenegro (2006-2017)	Serbia (1948-2017)	Ukraine (1991-2017)
Morocco (1956-2017)	Seychelles (1976-2017)	United Kingdom (1948-2017)
Mozambique (1975-2017)	Sierra Leone (1961-2017)	United States (1948-2017)
Myanmar (1948-2017)	Singapore (1965-2017)	Uruguay (1948-2017)
Namibia (1990-2017)	Slovakia (1993-2017)	Uzbekistan (1991-2017)
Nauru (1968-2016)	Slovenia (1991-2017)	Vanuatu (1980-2017)
Nepal (1948-2017)	Sol. Islands (1978-2017)	Venezuela (1948-2017)
Netherlands (1948-2017)	Somalia (1960-2017)	Vietnam (1949-2017)
New Zealand (1948-2017)	Spain (1948-2017)	Vietnam, South (1955-1975)
Nicaragua (1948-2017)	Sri Lanka (1948-2017)	Yemen (1948-2017)
Niger (1960-2017)	St. Kit./Nev. (1983-2017)	Yemen, South (1963-1990)
Nigeria (1960-2017)	St. Lucia (1979-2017)	Zambia (1964-2017)
Norway (1948-2017)	St. Vin./Gren. (1979-2017)	Zimbabwe (1964-2017)
Oman (1948-2017)	Sudan (1956-2017)	

J MSF Time-Series Estimates, 1948-2017 in 197 Countries

In figures 8-15 below, we include the time-series, MSF scores for all 197 countries in the sample. Higher scores indicate greater media freedom. Error bars indicate 80% posterior credible intervals. We note that in these figures, Russia includes the Soviet Union, Serbia includes Yugoslavia as well as Serbia and Montenegro, Germany includes West Germany, Yemen includes North Yemen, Vietnam includes North Vietnam, and Czech Republic includes Czechoslovakia.

Figure 18: MSF Time-Series Estimates, 1948-2017 from Afghanistan to Brunei

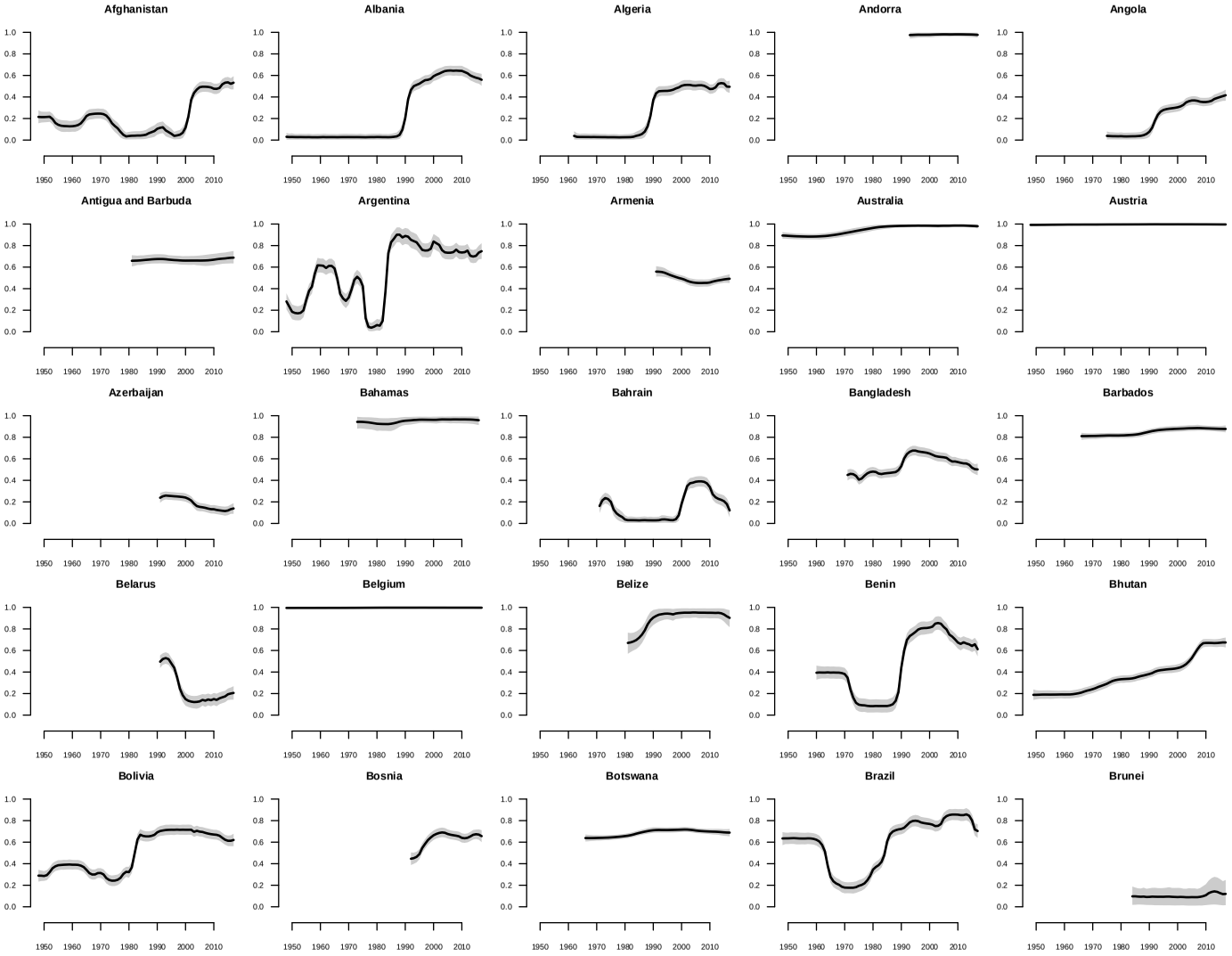


Figure 19: MSF Time-Series Estimates, 1948-2017 from Bulgaria to Ecuador

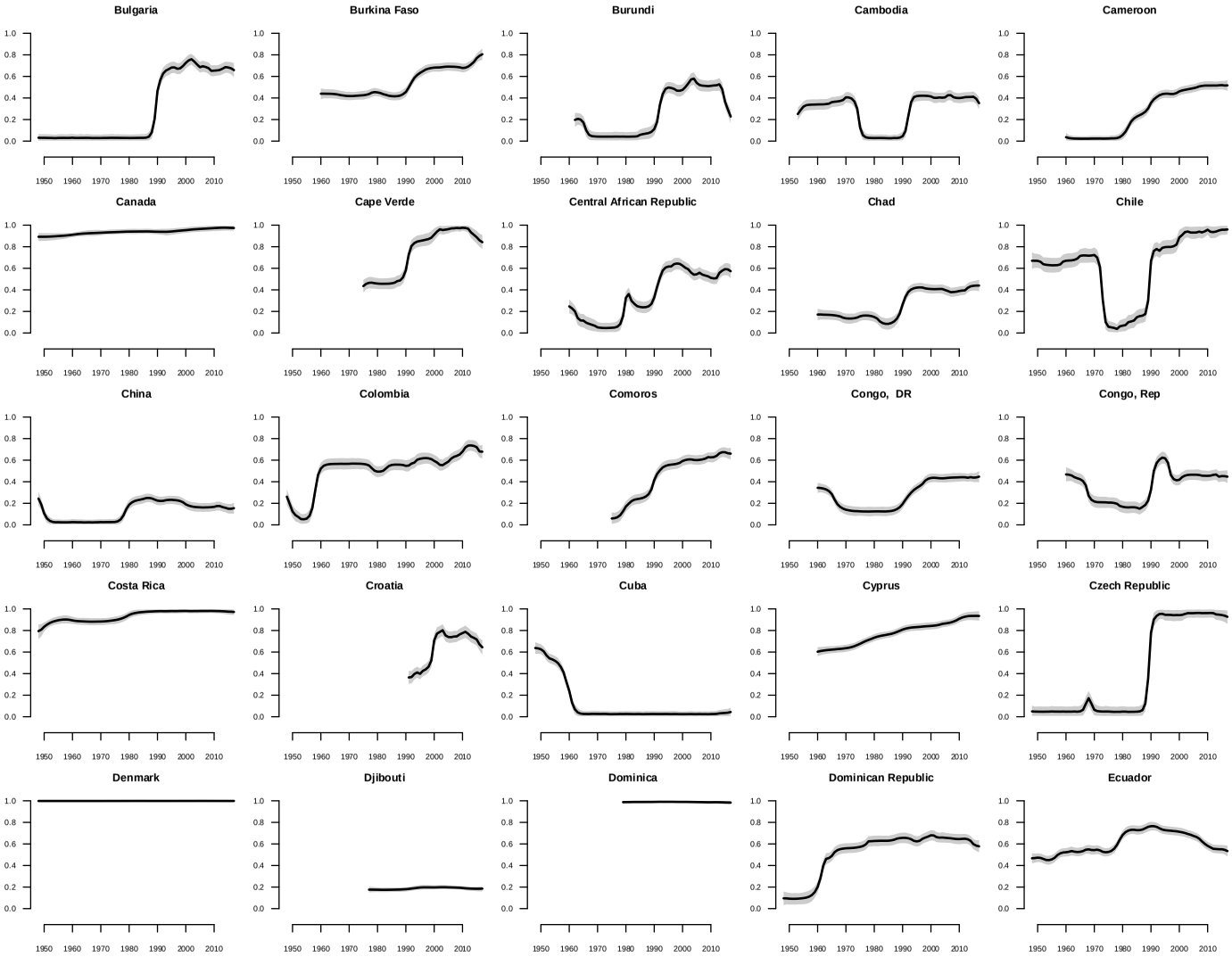


Figure 20: MSF Time-Series Estimates, 1948-2017 from Egypt to Iceland

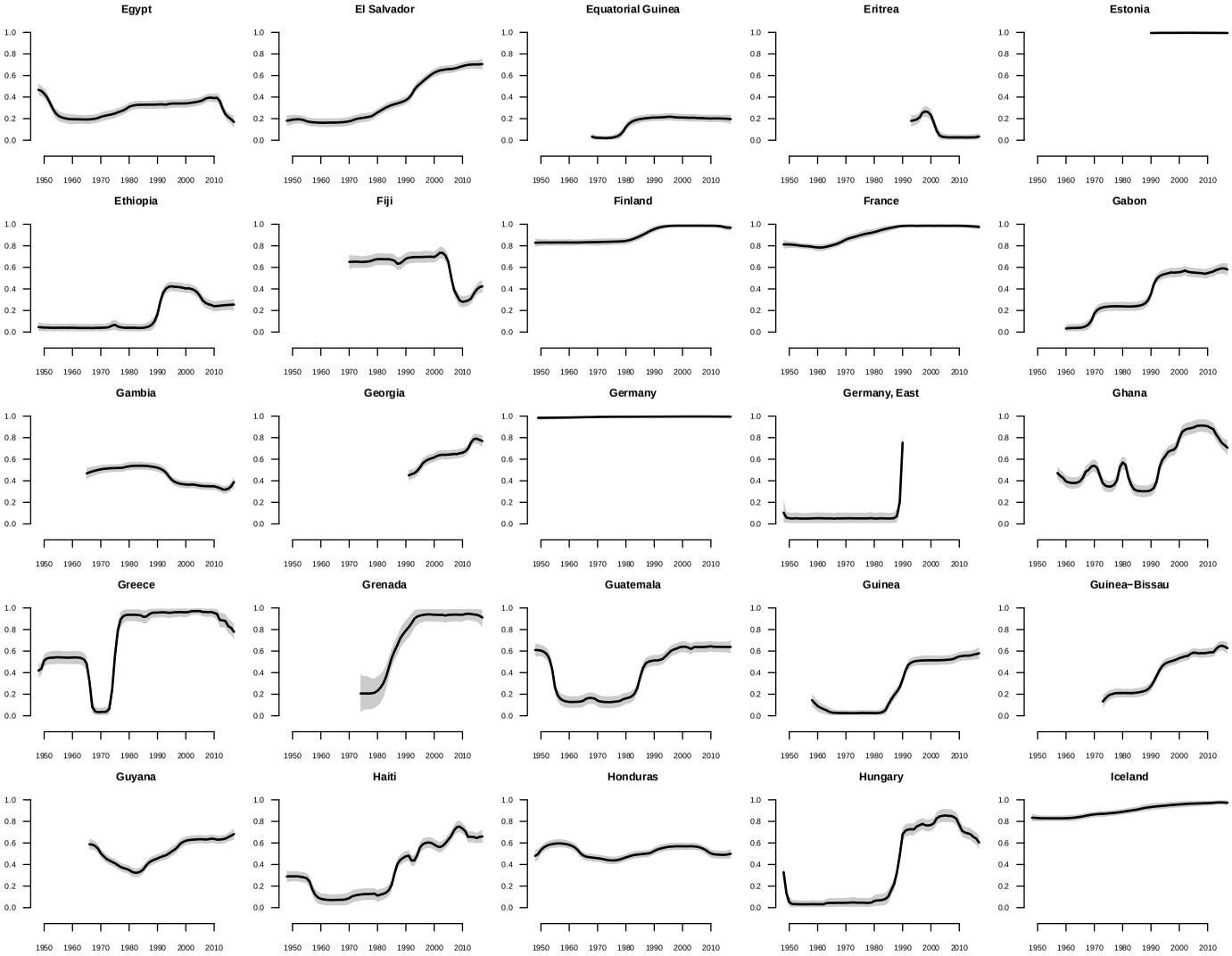


Figure 21: MSF Time-Series Estimates, 1948-2017 from India to Libya

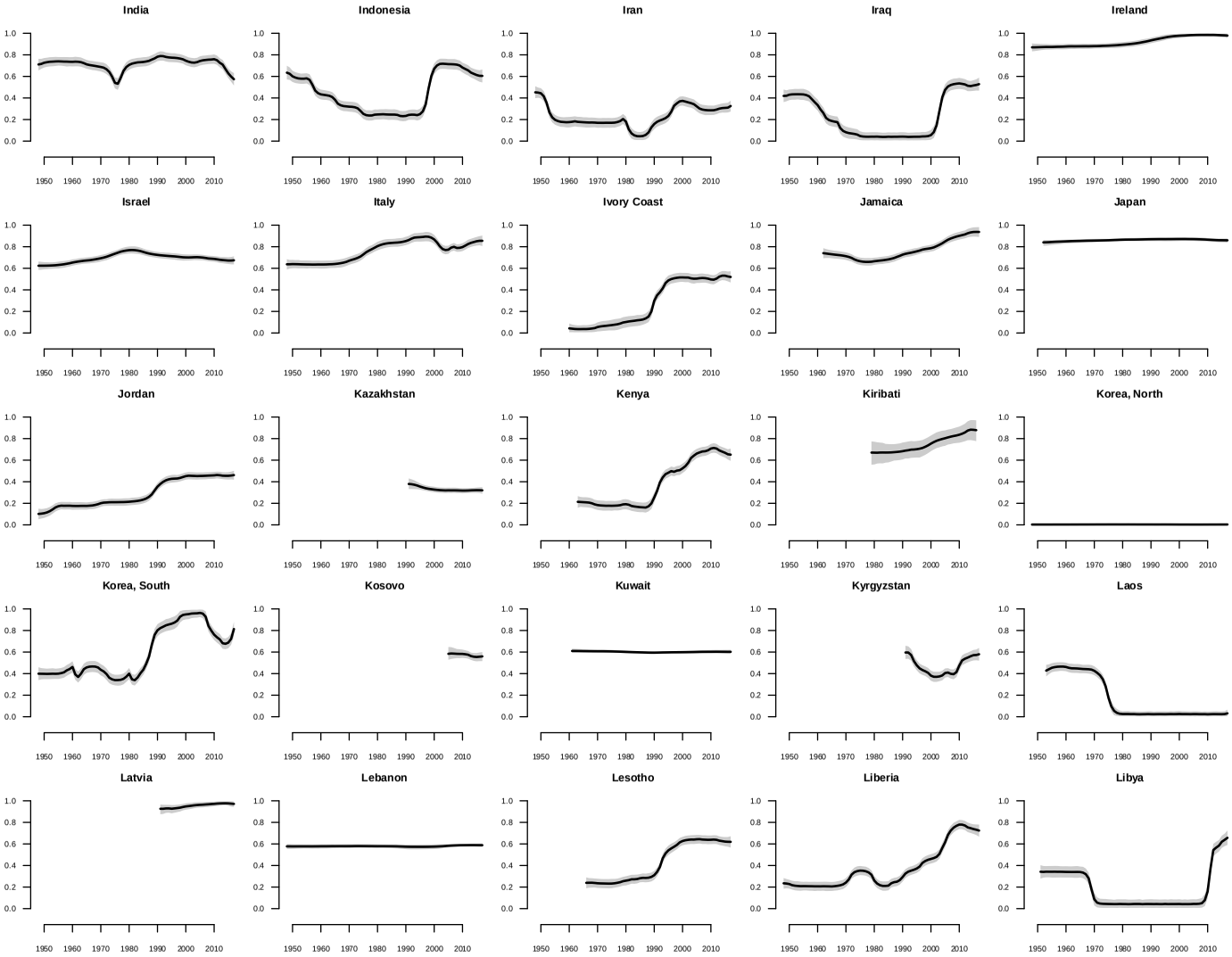


Figure 22: MSF Time-Series Estimates, 1948-2017 from Liechtenstein to Nepal

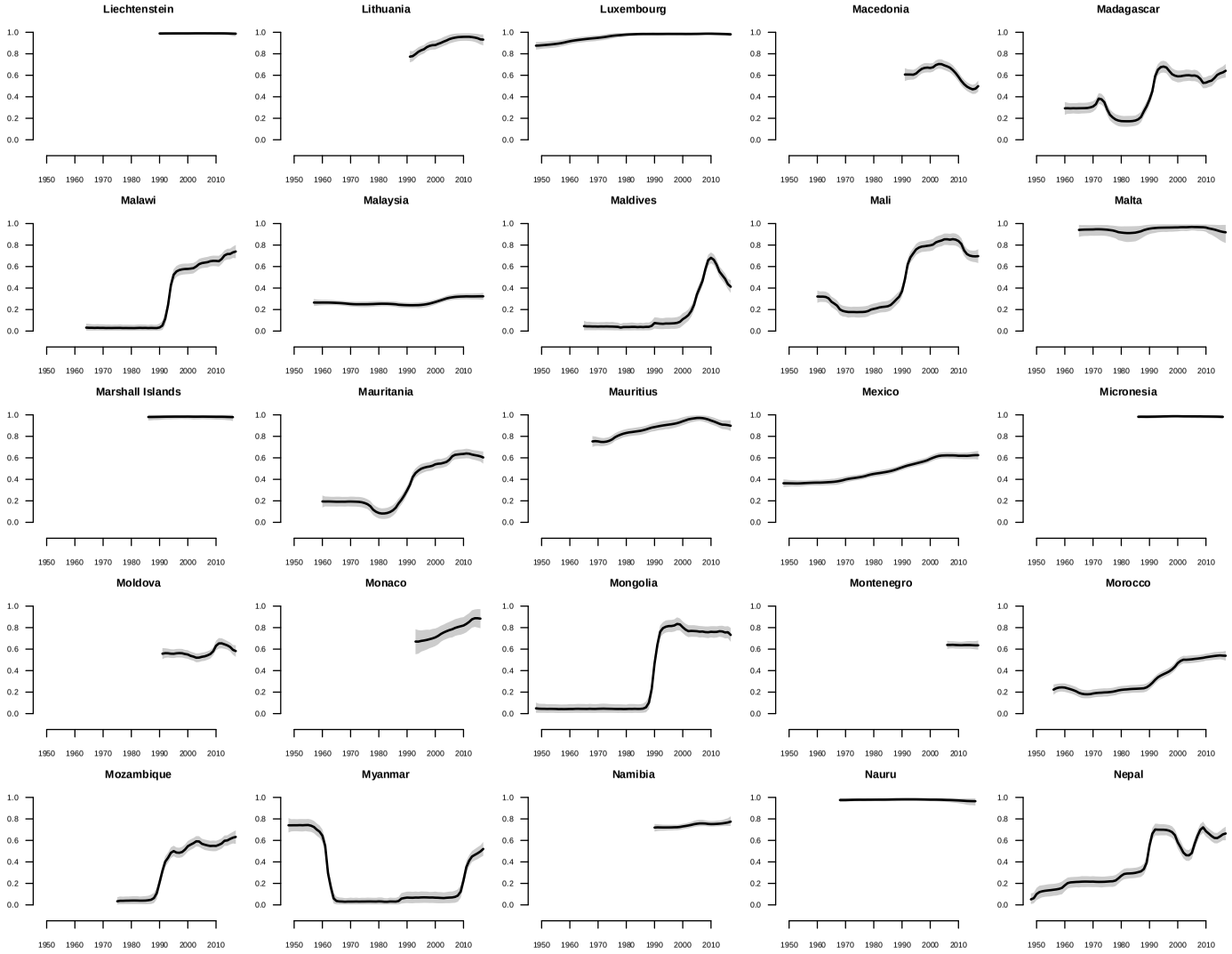


Figure 23: MSF Time-Series Estimates, 1948-2017 from Netherlands to Senegal)

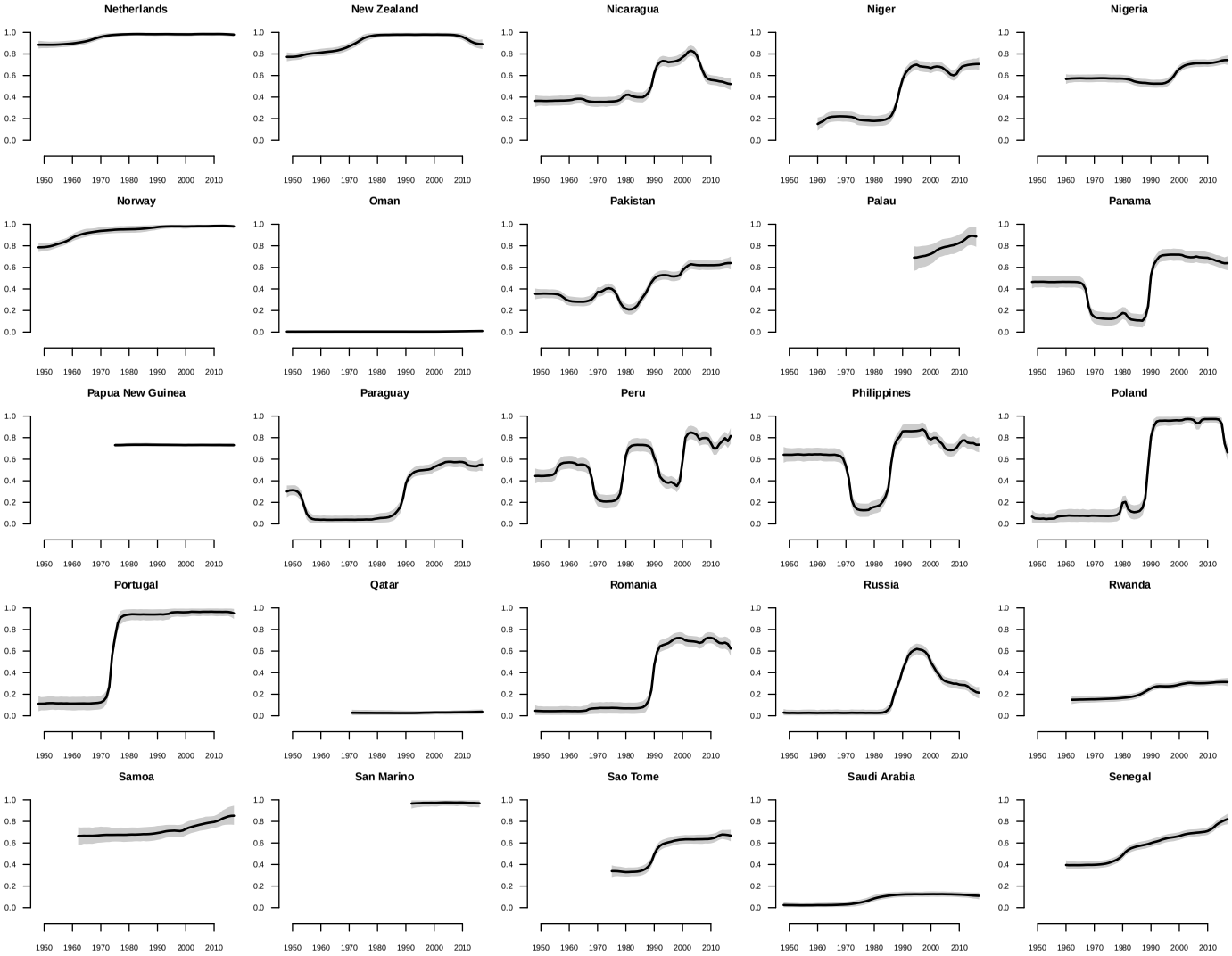


Figure 24: MSF Time-Series Estimates, 1948-2017 from Serbia (Yugoslavia) to East Timor

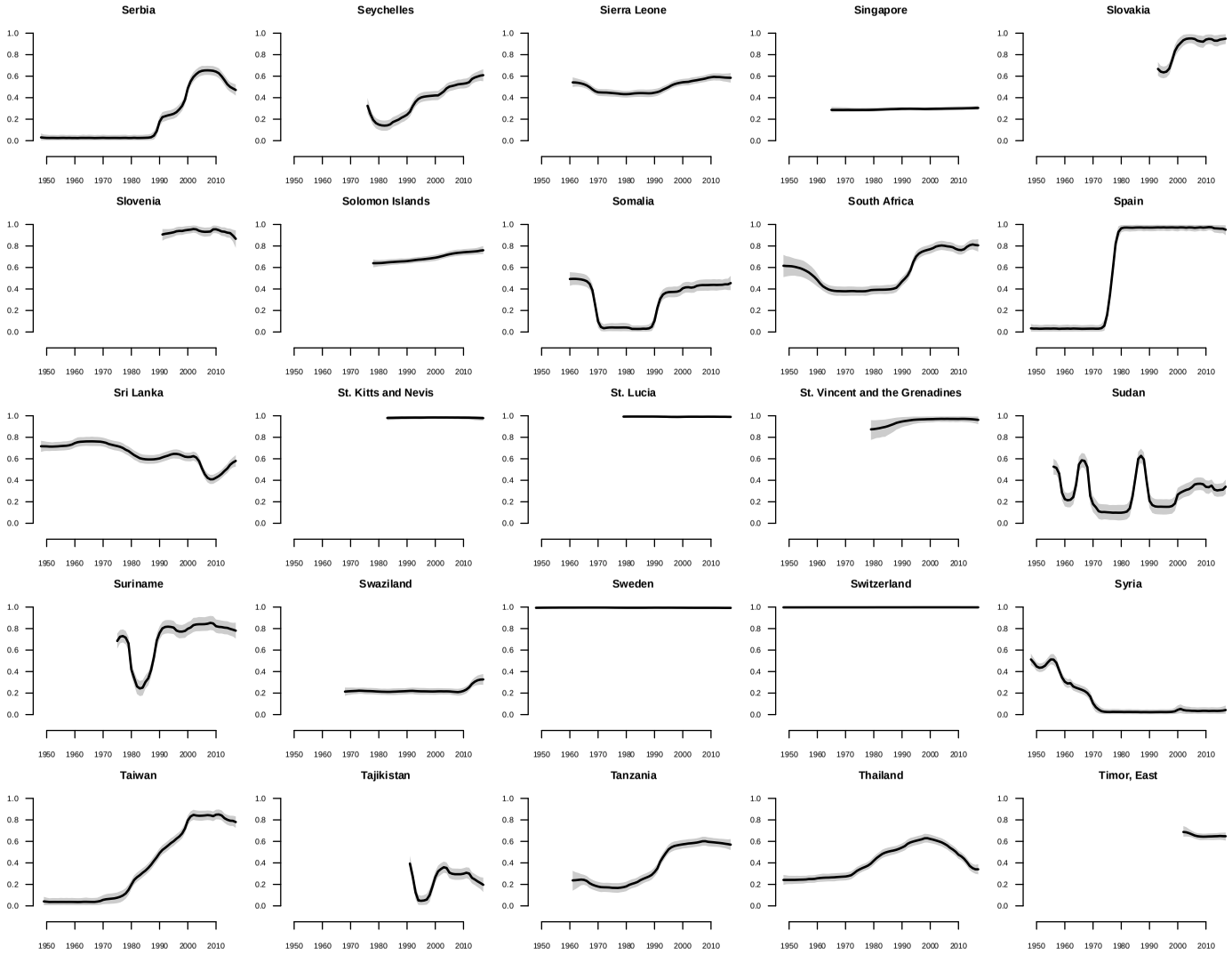
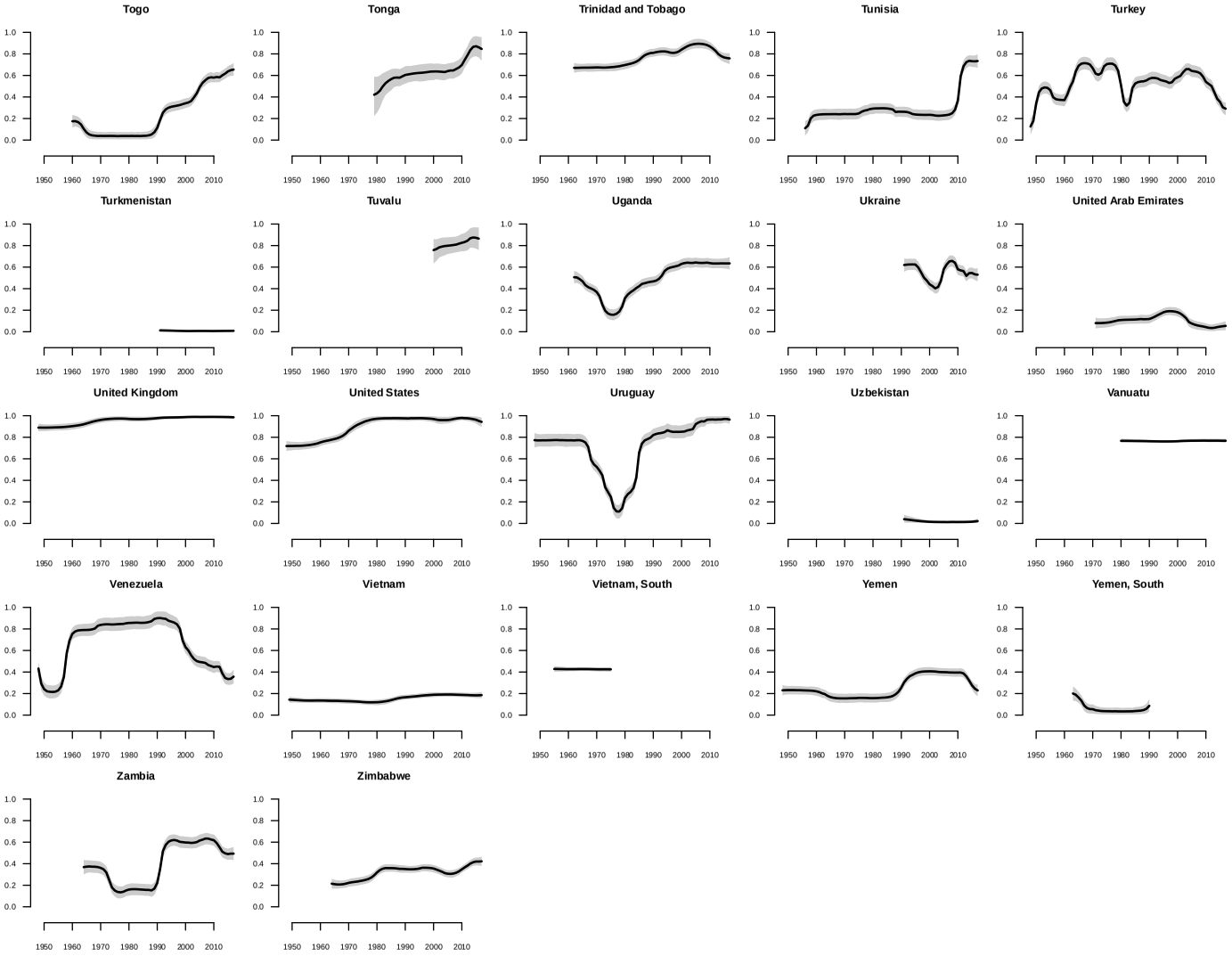


Figure 25: MSF Time-Series Estimates, 1948-2017 from Togo to Zimbabwe



References

Albuquerque, Afonso de. 2012. On Models and Margins. In *Comparing Media Systems Beyond the Western World*. Cambridge University Press pp. 72–95.

Baek, Mijeong. 2009. "A Comparative Analysis of Political Communication Systems and Voter Turnout." *American Journal of Political Science* 53(2):376–393.

Bajomi-Lázár, Péter. 2014. *Party Colonisation of the Media in Central and Eastern Europe: Modern Business Decision Making in Central and Eastern Europe*. Central European University Press.

Bolt, Jutta and Jan Luiten van Zanden. 2013. "The fFrst Update of the Maddison Project; Re-estimating Growth before 1820." *Maddison-Project Working Paper WP-4, University of Groningen, January* 5.

Brambor, Thomas, William Clark and Matt Golder. 2005. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14(1):63–82.

Brown, Bruce L., Suzanne B. Hendrix, Dawson W. Hedges and Timothy B. Smith. 2011. *Multivariate Analysis for the Biobehavioral and Social Sciences: A Graphical Approach*. John Wiley & Sons.

Brunetti, Aymo and Beatrice Weder. 2003. "A Free Press Is Bad News for Corruption." *Journal of Public Economics* 87(7):1801–1824.

Byman, Daniel and Jennifer Lind. 2010. "Pyongyang's Survival Strategy: Tools of Authoritarian Control in North Korea." *International Security* 35(1):44–74.

Chapman, Annabelle. 2017. "Pluralism Under Attack: The Assault on Press Freedom in Poland.".
URL: *https://freedomhouse.org/report/special-reports/assault-press-freedom-poland*

Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and dictatorship revisited." *Public choice* 143(1-2):67–101.

Cingranelli, David L. and David L. Richards. 2006. "The Cingranelli-Richards (CIRI) Human Rights Dataset.".
URL: *Available at: http://www.humanrightsdata.org (Accessed on January 4, 2007).*

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Luhrmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Aksel Sundtrom, Eitan Tzelgov, Luca Uberti, Yi ting Wang, Tore Wig and Daniel Ziblatt. 2018c. "V-Dem Codebook v8." *Varieties of Democracy (V-Dem) Project* .

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Agnes Cornell, Sirianne Dahlum, Haakon Gjerlow, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Luhrmann, Kyle Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Moa Olin, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Johannes von Romer, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Natalia Stepanova, Aksel Sundstrom, Eitan Tzelgov, Yi ting Wang, Tore Wig, Steven Wilson and Daniel Ziblatt. 2018a. "V-Dem [Country-Year/Country-Date] Dataset v8." *Varieties of Democracy (V-Dem) Project* .

Curry, Jane Leftwich. 2003. Status of the Media in Poland. In *Encylopedia of International Media and Communications, vol. 3*, ed. Donald H. Johnston. San Diego, CA: Academic Press pp. 495–504.

Deacon, Philip. 1999. The Media in Modern Spanish Culture. In *The Cambridge Companion to Modern Spanish Culture*, ed. David T. Gies. Cambridge University Press pp. 309–317.

Dunteman, George H. 1989. *Principal Components Analysis*. Number 69 Sage.

Egorov, Georgy, Sergei Guriev and Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103(4):645–668.

Eribo, Festus and William Jong-Ebot. 1997. *Press Freedom and Communication in Africa*. Africa World Press: Trenton New Jersey.

Freedom House. 2017. "Freedom of the Press Dataset.".
URL: *https://freedomhouse.org/report-types/freedom-world*

Freille, Sebastian, M. Emranul Haque and Richard Kneller. 2007. "A Contribution to the Empirics of Press Freedom and Corruption." *European Journal of Political Economy* 23:838–62.

Gelman, Andrew. 2008. "Scaling Regression Inputs by Dividing by Two Standard Deviations." *Statistics in Medicine* 27(15):2865–2873.

Gelman, Andrew and Kenneth Shirley. 2011. Inference from Simulationsand Monitoring Convergence. In *Handbook of Markov Chain Monte Carlo*, ed. Steve Brooks, Andrew Gelman, Galin Jones and Xiao-Li Meng. Vol. Ch. 6 CRC press.

Gleditsch, Nils Petter, Mikael Eriksson, Margareta Sollenberg and Havard Strand. 2002. "Armed Conflict 1946–2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.

Golder, Matt. 2005. "Democratic Electoral Systems around the World, 1946–2000." *Electoral Studies* 24(1):103–121.

Harbom, Lotta and Peter Wallensteen. 2012. "Armed Conflicts, 1946–2011." *Journal of Peace Research* 49(4):565–575.

Hawken, Angela and Gerardo L. Munck. 2009. "Do you know your data? Measurement validity in corruption research." *Unpublished typescript, Pepperdine University and University of Southern California, Malibu, CA, and Los Angeles* .

Hervieu, Benoît. 2013. "The Country of Thirty Berlusconis.".
URL: *https://rsf.org/en/reports/brazil-country-thirty-berlusconis*

Heston, Alan, Robert Summers and Bettina Aten. 2002. "Penn World Table Version 6.1." *Center for International Comparisons at the University of Pennsylvania (CICUP)* . Accessed January 4, 2008.

Hirschman, Albert O. 1964. "The Paternity of an Index." *The American Economic Review* 54(5):761–762.

Jakubowicz, Karol. 1992. Media and the Terminal Crisis of Communism in Poland. In *Media, Crisis and Democracy: Mass Communication and the Disruption of Social Order*, ed. Marc Raboy and Bernard Dagenais. London–Newbury Park–New Delhi: Sage pp. 79–93.

Kailitz, Steffen. 2013. "Classifying Political Regimes Revisited: Legitimation and Durability." *Democratization* 20(1):39–60.

Kaiser, Henry F. 1958. "The Varimax Criterion for Analytic Rotation in Factor Analysis." *Psychometrika* 23(3):187–200.

Kalven, Harry. 1988. *A Worthy Tradition*. 1st ed. Harper and Row.

Kellam, Marisa and Elizabeth A. Stein. 2016. "Silencing Critics: Why and How Presidents Restrict Media Freedom in Democracies." *Comparative Political Studies* 49(1):36–77.

Kitschelt, Herbert, Zdenka Mansfeldova, Radoslaw Markowski and Gábor Tóka. 1999. *Post-communist Party Systems: Competition, Representation, and Inter-party Cooperation*. Cambridge University Press.

Lewis, Anthony. 2007. *Freedom for the Thought That We Hate: A Biography of the First Amendment*. Basic Books.

Marshall, Monty G. and Keith Jaggers. 2017. "POLITY IV Project: Political Regime Characteristics and Transitions, 1800-2016 Dataset Users Manual.".
URL: *URL: http://www.systemicpeace.org/inscrdata.html*

McMann, Kelly M., Daniel Pemstein, Brigitte Seim, Jan Teorell and Staffan Lindberg. 2016. "Strategies of Validation: Assessing the Varieties of Democracy Corruption Data." *V-Dem Working Paper* 23.

Mohammed, Abubakar and Fadeyi Taofiq James. 2017. "Media and Democratic Imperatives in Nigerias Fourth Republic." *Journal of African Studies and Development* 9(2):7–14.

Ogbondah, Chris W. 2011. A Critical Analysis of State-Press Relations in Nigeria, 1999–2005. In *Mass Media and Society in Nigeria*. Malthouse Press Limited: Lagos pp. 25–48.

Oso, Lai, Dele Odunlami and Tokunbo Adaja. 2011. Socio-historical Context of the Development of Nigerian Media. In *Mass Media and Society in Nigeria*. Malthouse Press Limited: Lagos pp. 1–24.

Ost, David. 2001. Poland: Parties, Movements, Groups, and Ambiguity. In *Political Parties and Interest Groups. Shaping Democratic Governance*, ed. Clive S. Thomas. Boulder–London: Lynne Rienner pp. 211–227.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi ting Wang, Joshua Krusell and Farhad Miri. 2018. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." *University of Gothenburg, Varieties of Democracy Institute* (Working Paper No. 21, 3d edition.).

Porto, Mauro. 2012. *Media Power and Democratization in Brazil: TV Globo and the Dilemmas of Political Accountability*. Routledge.

Powe, Lucas A. 1991. *The Fourth Estate and the Constitution: Freedom of the Press in America*. University of California Press.

Reporters Without Borders. 2018. "Press Freedom Index.".
URL: *https://rsf.org/en/world-press-freedom-index*

Ross, Michael L. 2011. "Oil and Gas Data, 1932–2011.". Accessed July 16, 2013.
URL: *http://dvn.iq.harvard.edu/dvn/dv/mlros*

Sapiro, Virginia and W. Philips Shively. 2003. "Comparative Study of Electoral Systems, 1996–2001: Module 1 Full Release [dataset]." *Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor]* .

Schneider, Laura. 2014. "Media Freedom Indices. What They Tell Us–and What They Don't." *Deutsche Welle Akademie: Bonn* .

Schoonvelde, Martijn. 2014. "Media Freedom and the Institutional Underpinnings of Political Knowledge." *Political Science Research and Methods* 2(2):163–78.

Smith, Anne-Marie. 1997. *A Forced Agreement: Press Acquiescence to Censorship in Brazil*. University of Pittsburgh Press.

Stier, Sebastian. 2015. "Democracy, Autocracy and the News: The Impact of Regime Type on Media Freedom." *Democratization* 22(7):1273–1295.

Strand, Havard, Joachim Carlsen, Nils Petter Gleditsch, Havarad Hegre, Christin Ormhaug, Lars Wilhelmsen, Peter Wallensteen, Margareta Sollenberg, Mikael Eriksson, Lotta Harbom, Halvard Buhaug and Jan Ketil Rod. 2005. "Armed Conflict Dataset Codebook Version 3–2005. International Peace Research Institute, Oslo: Department of Peace and Conflict Research, Uppsala University. Department of Sociology and Political Science. Department of Geography Norwegian University of Science and Technology.".

Tejumaiye, Adepoju and Omowale Abelabu. 2011. Mass Media and Governance: Issues and Challenges in Contemporary Nigeria. In *Mass Media and Society in Nigeria*. Malthouse Press Limited: Lagos pp. 63–76.

Teorell, Jan, S. Dahlberg, S. Holmberg, B. Rothstein, A. Khomenko and R. Svensson. 2017. "The Quality of Government Standard Dataset, 2017." *University of Gothenburg: The Quality of Government Institute* .

Van Belle, Douglas A. 1997. "Press Freedom and the Democratic Peace." *Journal of Peace Research* 34(4):405–414.

Whitten-Woodring, Jenifer. 2009. "Watchdog or Lapdog? Media Freedom, Regime Type, and Government Respect for Human Rights." *International Studies Quarterly* 53(3):595–625.

Whitten-Woodring, Jenifer and Douglas A. Van Belle. 2017. "The Correlates of Media Freedom: An Introduction of the Global Media Freedom Dataset." *Political Science Research and Methods* 5(1):179–188.

Williamson, Edwin. 2009. *The Penguin History Of Latin America*. Penguin UK.

World Bank. 2017. "World Development Indicators.".
URL: *https://data.worldbank.org/data-catalog/world-development-indicators*