Supporting Information:

In Search of Self-censorship

The Supporting Information includes the following materials:

Missingness and Country Coverage

Table A1: Country Samples and Question Coverage (1/2)

Country-Year	Q1 *conf.gov*	Q2 *humanrights*	Q3 *democracy*	Q4 *conf.tv*	Q5 *gen.trust*	Q6 *life.sat*	Obs
Algeria (2002)			NA				1282
Algeria (2013)							1200
Armenia (1997)			NA				2000
Azerbaijan (1997)			NA				2002
Azerbaijan (2011)							1002
Bahrain (2014)			NA				1200
Belarus (1996)			NA				2092
Belarus (2011)							1535
China (2001)			NA				1000
China (2007)							1991
China (2012)							2300
Croatia (1996)		NA	NA				1196
Egypt (2001)			NA				3000
Egypt (2013)							1523
Ethiopia (2007)							1500
Iran (2000)			NA	NA			2532
Iran (2007)							2667
Iraq (2012)							1200
Jordan (2001)			NA				1223
Jordan (2007)							1200
Jordan (2014)							1200
Kuwait (2014)			NA				1303
Kyrgyzstan (2003)			NA				1043
Malaysia (2006)							1201
Mexico (1995)		NA	NA				854
Mexico (1996)		NA	NA				1510

Note: Table shows all authoritarian country samples utilized in the analysis from the WVS. "NA" indicates where a question was not asked for the sample.

Table A1: Country Samples and Question Coverage (2/2)

Country-Year	Q1 *conf.gov*	Q2 *humanrights*	Q3 *democracy*	Q4 *conf.tv*	Q5 *gen.trust*	Q6 *life.sat*	Obs
Morocco (2001)			NA				1251
Morocco (2007)							1200
Morocco (2011)							1200
Nigeria (1990)		NA	NA				1001
Nigeria (1995)		NA	NA				1996
Nigeria (2000)			NA				2022
Pakistan (2001)			NA				2000
Peru (1996)		NA	NA				1211
Qatar (2010)			NA				1060
Russia (1995)		NA	NA				2040
Russia (2011)							2500
Saudi Arabia (2003)	NA		NA				1502
Singapore (2012)							1972
Vietnam (2001)		NA					1000
Vietnam (2006)							1495
Tanzania (2001)			NA				1171
Thailand (2007)							1534
Uganda (2001)			NA				1002
Uzbekistan (2011)		NA	NA				1500
Zimbabwe (2001)			NA				1002

Note: Table shows all authoritarian country samples utilized in the analysis from the WVS. "NA" indicates where a question was not asked for the sample.

Index Construction and Performance

The paper calculates the following index across the different country-year samples:

$$cens.ind.3q.3q_{ct} = \left(\frac{\sum_{j=1}^{m} \sum_{i=1}^{n} nonresponse_{ij}}{m \times n} \right)_{regime\ assessment} - \left(\frac{\sum_{k=1}^{m} \sum_{i=1}^{n} nonresponse_{ik}}{m \times n} \right)_{nonsensitive} \quad (1)$$

For each country c and sample at time t, for each regime assessment question j, we sum the item nonresponse indicator across all respondents n, sum this across all of the m questions, and divide the total by $m \times n$. This gives the mean item nonresponse for the regime assessment questions in the sample. We then subtract the mean item nonresponse for a set of nonsensitive political questions.

The questions used for the construction of the self-censorship index are shown in Box 1. The three regime assessment questions ask respondents to report their general confidence in government, and their perceptions of human rights and democracy in their country. The nonsensitive questions involve measures of interpersonal trust, life satisfaction, and confidence in the television industry.

Box 1: Questions for Falsification Index Construction
World Values Survey - Core Questionnaire

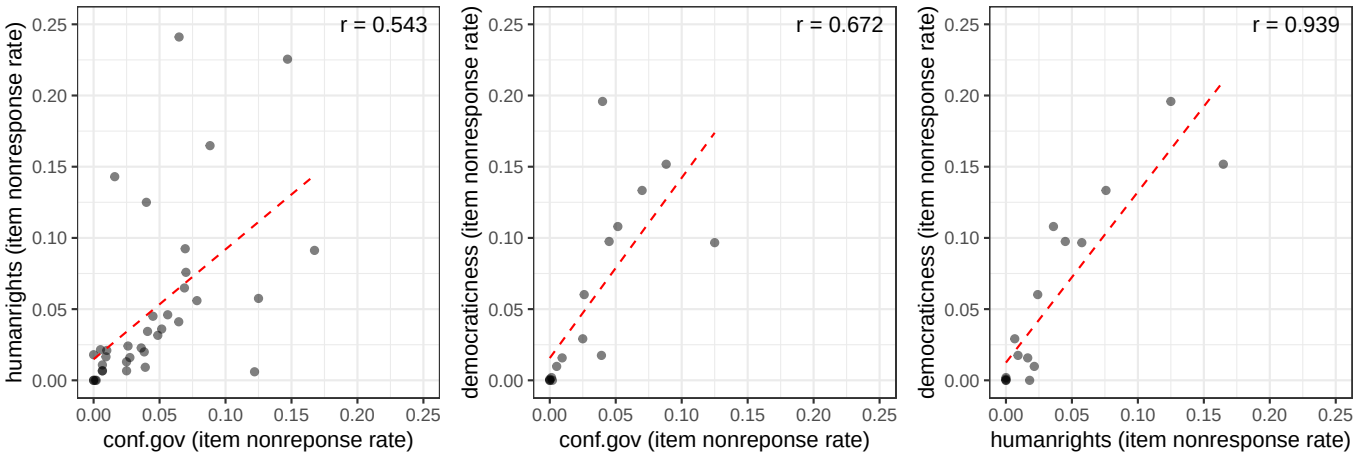
Regime Assessment Questions	Nonsensitive Questions
V115. I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all? The government (in your nation's capital) (*conf.gov*)	V111. I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all? Television (*conf.tv*)
V141. How democratically is this country being governed today? (*democracy*)	V24. Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? (*gen.trust*)
V142. How much respect is there for individual human rights nowadays in this country? (*humanrights*)	V23. All things considered, how satisfied are you with your life as a whole these days? (*life.sat*)

The choice of these questions is inherently arbitrary, and the index itself may be sensitive to this decision. In the paper and in the Supporting Information, we constructed different ver-

sions of the index. Our primary index used in the paper *cens.ind.3q.3q* uses all six questions but requires more imputation; a more minimalist approach (*cens.ind.1q.1q*) employs only two questions (*conf.gov* and *gen.trust*) but has greater coverage. We also explore an index that does not include any nonsensitive questions (*cens.ind*.3*q*.0*q*), which amounts to the mean nonresponse rate for sensitive questions. We also explore an version of the index that includes 20 nonsensitive questions included in the WVS sample (*cens.ind*.3*q*.20*q*).[1] The figures below show the correlations in item nonresponse rates for the different questions, as well as the correlations between the indices themselves. One multiple questions are included, the correlations in the different versions of the index are generally above 0.75.

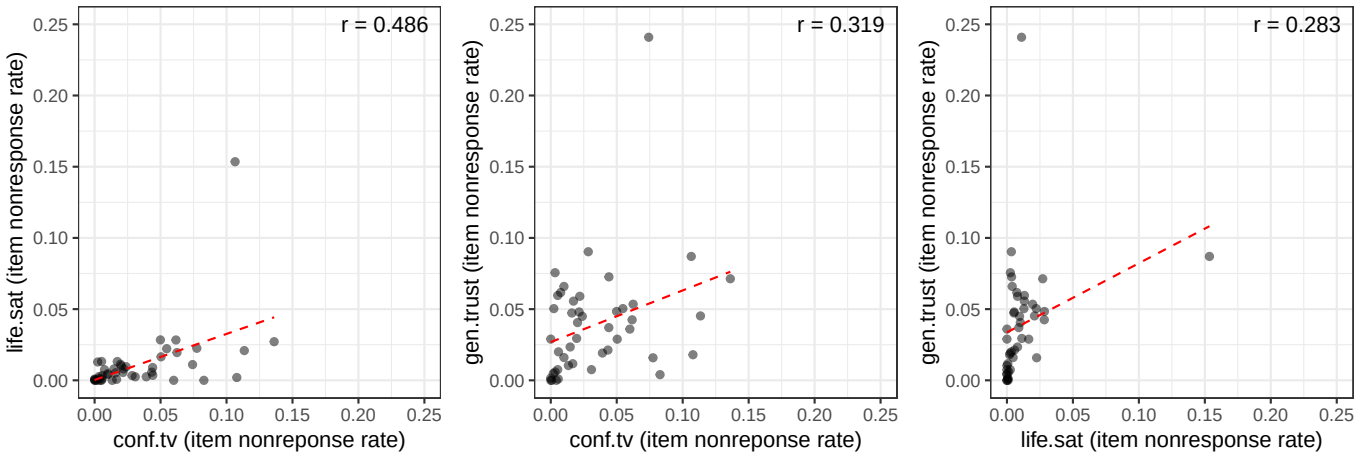
[1]The list of nonsensitive WVS questions used are: A001 - Important in life: Family; A002 - Important in life: Friends; A003 - Important in life: Leisure time; A005 - Important in life: Work; A008 - Feeling of happiness; A009 - State of health (subjective); A165 - Most people can be trusted; A170 - Satisfaction with your life; B002 - Increase in taxes if used to prevent environmental pollution; B008 - Protecting environment vs. Economic growth; C006 - Satisfaction with financial situation of household; D018 - Child needs a home with father and mother; D022 - Marriage is an out-dated institution; D023 - Woman as a single parent; D054 - One of main goals in life has been to make my parents proud; D055 - Make effort to live up to what my friends expect; D057 - Being a housewife just as fulfilling; D060 - University is more important for a boy than for a girl; E012 - Willingness to fight for country; E069.10 - Confidence: Television.

Figure A1a: Correlations in Item Nonresponse Rates for Sensitive Questions



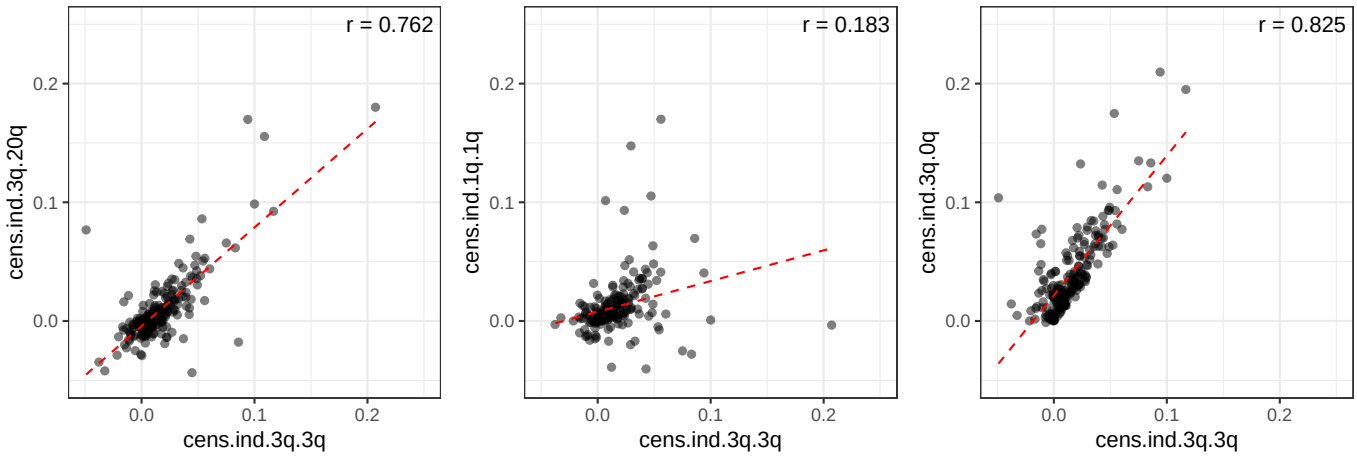
Note: Figure shows correlations in item nonresponse rates across the three sensitive questions (*conf.gov*, *democracy*, *humanrights*) used in the construction of *cens.ind*. All data is unweighted. Each point represents a different authoritarian country-year sample of the World Values Survey.

Figure A1b: Correlations in Item Nonresponse Rates for Nonsensitive Questions



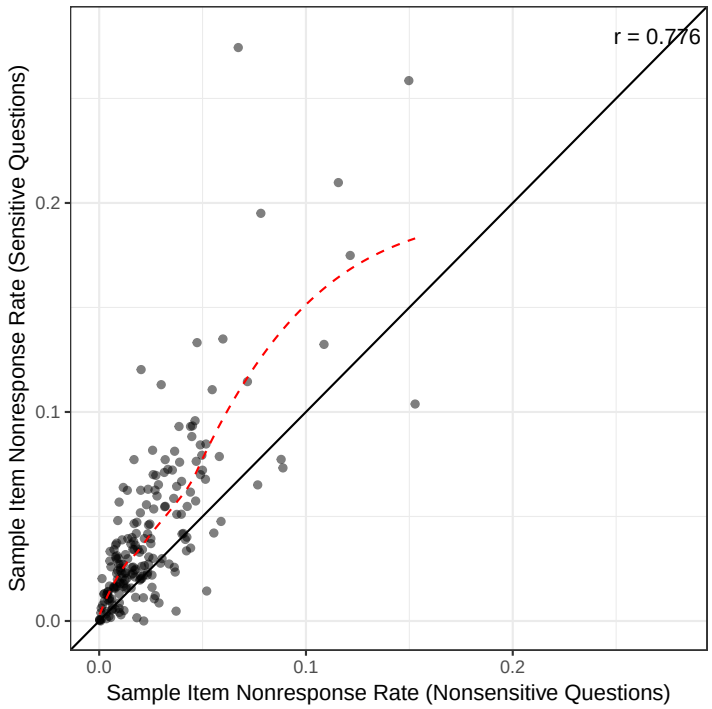
Note: Figure shows correlations in item nonresponse rates across the three nonsensitive questions (*conf.tv*, *gen.trust*, *life.sat*) used in the construction of *cens.ind*. All data is unweighted. Each point represents a different authoritarian country-year sample of the World Values Survey.

Figure A2: Correlations between Constructions of Self-censorship Index



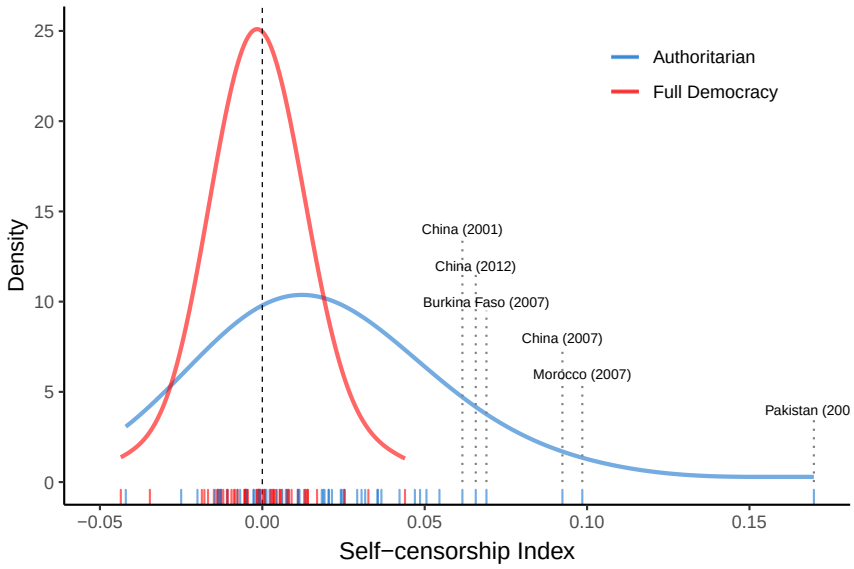
Note: Figure shows correlations between four different constructions of the self-censorship index (*cens.*3*q*.3*q*, *cens.ind.*3*q*.20*q*, *cens.ind.*1*q*.1*q*, *cens.ind.*3*q*.0*q*). Each point represents a different country-year sample of the World Values Survey. All data is unweighted.

Figure A3: Sensitive and Nonsensitive Nonresponse Rates



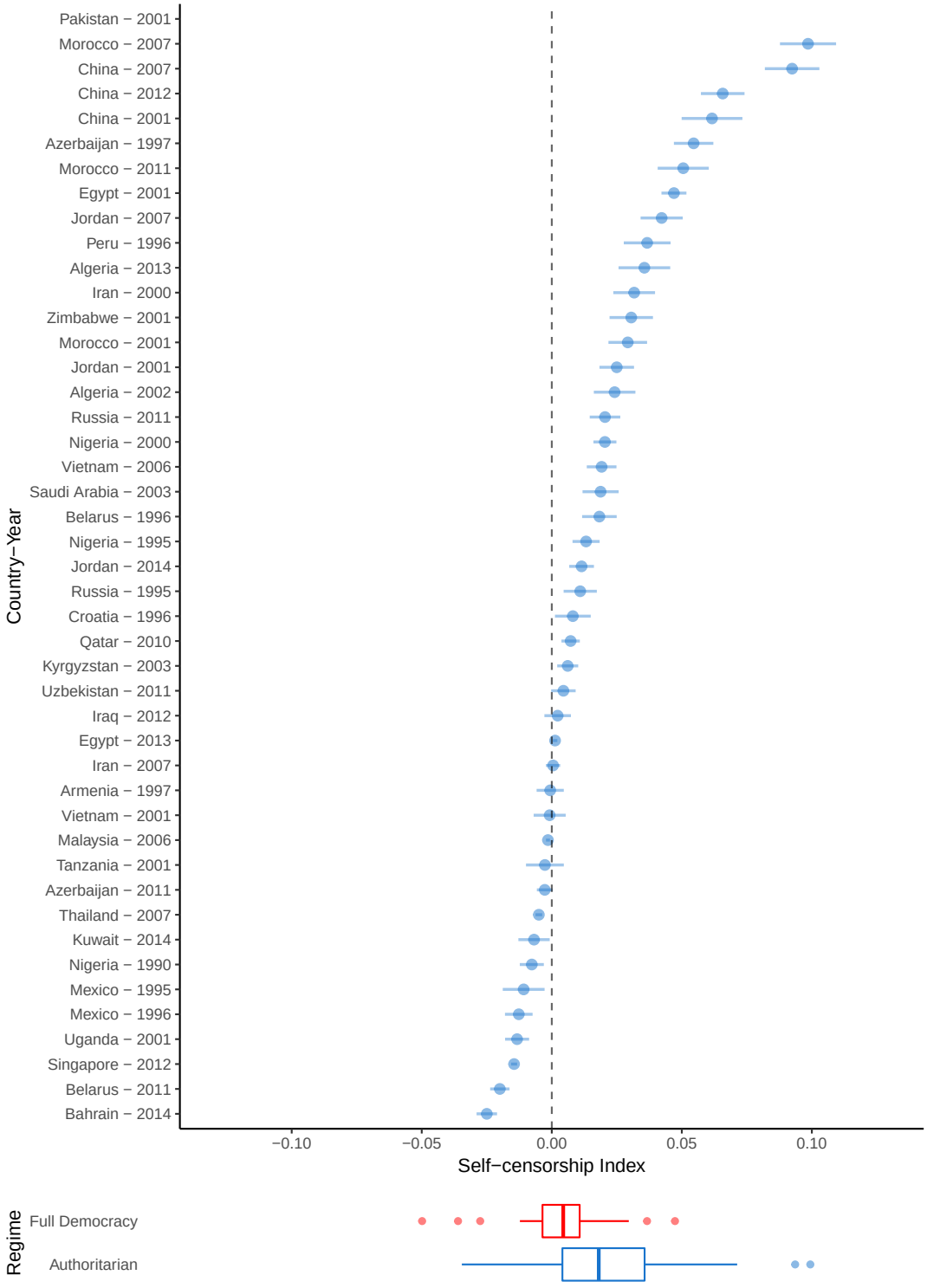
Note: Figure shows correlations between the sample average item nonresponse rates for the sensitive and nonsensitive questions. Each point represents a different country-year sample of the World Values Survey. All data is unweighted.

Figure A4: Self-censorship Index by Regime Type (alternative index)



Note: Figure shows distribution of self-censorship falsification index ($cens.ind.3q.20q_{ct}$) across democratic and authoritarian country-year samples.

Figure A5: Self-Censorship Index in Authoritarian Country-year Samples



Note: Figure shows the self-censorship index across different country-year samples of the the World Values Survey. Segments depict 95% confidence intervals. The boxplots show the distribution of self-censorship falsification index ($cens.ind.3q.20q_{ct}$) across democratic and authoritarian country-year samples.

Cross-National Analysis

Table A2: Testing Relationship Between Democracy and Self-censorship Index

Democracy Measure	Outcome					
	*cens.ind.*3*q*.3*q*		*cens.ind.*3*q*.0*q*		*cens.ind.*3*q*.20*q*	
	(1)	(2)	(3)	(4)	(5)	(6)
polity (continuous)	-0.00060*		-0.00111**		-0.00134***	
	(0.00042)		(0.00057)		(0.00044)	
democracy (dichotomized)		-0.0097**		-0.0133**		-0.0148***
		(0.0059)		(0.0079)		(0.0063)
Obs	197	197	197	197	197	197

Note: Table shows results from bivariate regressions of different versions of the self-censorship index on measures of democracy. The *polity* measure is the ordinal 20-point combined index from the Polity IV project. The dichotomous measure splits the sample at the a combined score of 5, with country-years below coded as authoritarian, and country years above coded as democratic. Analysis relies on multiple imputation for the construction of the index. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

What explains variation in the level of self-censorship in authoritarian systems? Below we conduct an exploratory analysis to probe several plausible correlates of interest from the authoritarian politics literature: the level of repression (*repression*); the level of political competition in the executive branch (*exec.comp*); the presence of multiple parties (*party.comp*); the level of military control in politics (*military*); and the duration of the regime spell (*duration*). The repression variable is drawn from the **?**'s Political Terror Scale, a five-point index where higher values correspond to higher levels of repression in society. All other variables are created from **?**'s Authoritarian Institutions dataset, which was extended by the authors to the present day to increase the coverage of the analysis.

Given the relatively small number of observations, we include only a few additional controls of theoretical interest: GDP per capita (*gdppc*), education levels (*edulevel*), urban population (*urbpop*), and oil rents (*oil*).

Figure XX shows the results of a "global sensitivity analysis" (**?**; **?**), which probes the distribution of coefficient estimates across different covariate sets and measurement strategies for the construction of the self-censorship index. Each point represents the coefficient estimate (and corresponding p-value from a two sided null hypothesis of no effect) from a different regression model of the self-censorship index on the independent variable of interest. Each figure shows all combinations of the six versions of the dependent variable (*cens.ind.3q.3q*, *cens.ind.3q.20q*, *cens.ind.1q.1q*, *cens.ind.3q.0q*) and four covariate sets: (bivariate; institutions; regime strength, and full). The covariate sets are as follows:

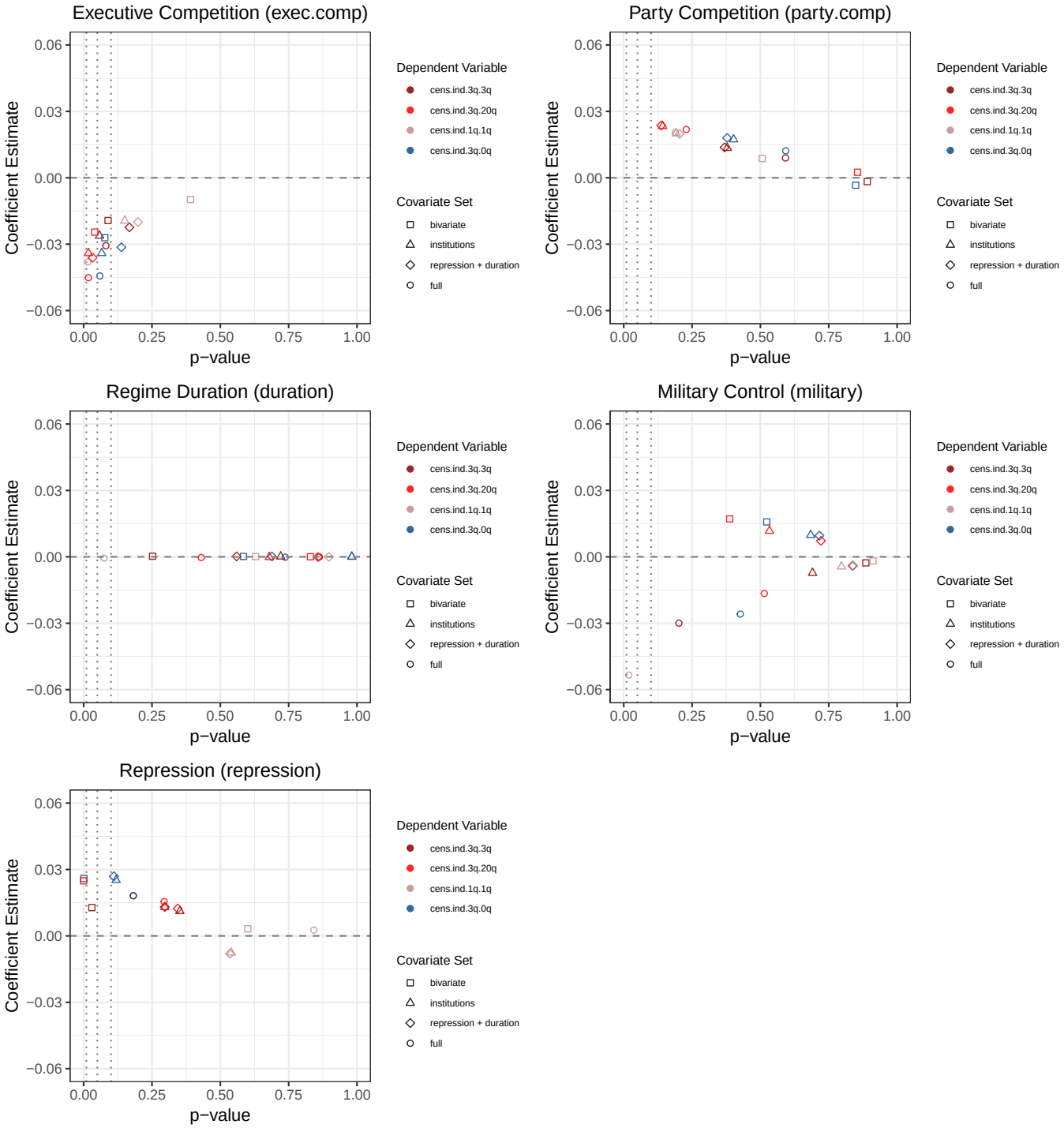
1. Bivariate
2. Institutions: *exec.comp*, *party.comp*, *military*
3. Regime Strength: *exec.comp*, *party.comp*, *military*, *repression*, *duration*
4. Full: *exec.comp*, *party.comp*, *military*, *repression*, *duration*, *gdppc*, *edulevel*, *oil*

This yields sixteen estimates for each independent variable: *exec.comp*, *party.comp*, *duration*, *military*, *repression*.[2]

The analysis shows that a lack of competition in the executive appears to be robustly associated with self-censorship. The *exec.com* indicator is coded "1" for any regime where the ruler is elected and "0" for all other regimes. This is meant to capture "electoral authoritarian" regimes (**?**; **?**; **??**). Paradigmatic cases in our dataset include Yeltsin's Russia, Mugabe's Zimbabwe, and Zedillo's Mexico. Regimes with elections appear to have a roughly 2-4 percentage lower score on our index.

[2]All models are estimated using OLS with robust standard errors. The imputation models are conducted using the Amelia II package in concert with Zelig, where the number of missing datasets is set to 50.

Figure A6: Determinants of Self-censorship Index in Authoritarian Systems



Note: Figure shows the distribution of coefficient estimates across different covariate sets and measurement strategies for the construction of the self-censorship index. Each point represents the coefficient estimate (and corresponding p-value from a two sided null hypothesis of no effect) from a different regression model. Each model shows all combinations of the four versions of the dependent variable and four covariate sets, yielding sixteen estimates for each independent variable of interest.

Table A3: Variable Descriptions for Cross-national Analysis

Concept	Variable	Description	Source
Repression	*repression*	Political Terror Scale, dichotomized (4,5 → 1; 1,2,3 → 0)	**?**
Executive competition	*exec.comp*	Electoral competition in the executive (elected by less than 75%, elected by more than 75% → 1; all others → 0)	**?**
Party competition	*party.comp*	Party competition (multiple parties → 1; all others → 0)	**?**
Military control	*military*	Military involvement in politics (direct, indirect → 1; civilian → 0)	**?**
Regime duration	*duration*	consecutive years in which the egime has been in power	**?**
Wealth	*gdppc*	GDP per capita (current USD)	World Development Indicators
Education	*edulevel*	Enrollment in secondary education, % of population	World Development Indicators
Urbanization	*urbpop*	People living in urban areas, % of population	World Development Indicators
Oil rents	*oil*	Oil rents as a percentage of GDP, dichotomized (greater than 5% → 1; less than 5% → 0)	World Development Indicators

Table A4: Results of Cross-national Analysis (1/2)

Independent Variable	Dependent Variable	Covariates	Estimate	SE	p-value
repression	*cens.ind.3q.3q*	bivariate	0.0128	0.0059	0.0301
repression	*cens.ind.3q.3q*	institutions	0.0112	0.012	0.352
repression	*cens.ind.3q.3q*	repression + duration	0.0131	0.0126	0.297
repression	*cens.ind.3q.3q*	full	0.0181	0.0136	0.1816
repression	*cens.ind.1q.1q*	bivariate	0.0032	0.0061	0.6006
repression	*cens.ind.1q.1q*	institutions	-0.0076	0.0124	0.5396
repression	*cens.ind.1q.1q*	repression + duration	-0.008	0.0129	0.5351
repression	*cens.ind.1q.1q*	full	0.0026	0.0131	0.842
repression	*cens.ind.3q.0q*	bivariate	0.0259	0.0078	0.0009
repression	*cens.ind.3q.0q*	institutions	0.0252	0.0162	0.1188
repression	*cens.ind.3q.0q*	repression + duration	0.027	0.0169	0.1111
repression	*cens.ind.3q.0q*	full	0.0181	0.0136	0.1816
repression	*cens.ind.3q.20q*	bivariate	0.0249	0.0061	0.0000
repression	*cens.ind.3q.20q*	institutions	0.013	0.0124	0.2973
repression	*cens.ind.3q.20q*	repression + duration	0.0124	0.0131	0.3442
repression	*cens.ind.3q.20q*	full	0.0155	0.0148	0.2946
party.comp	*cens.ind.3q.3q*	bivariate	-0.0017	0.0127	0.8909
party.comp	*cens.ind.3q.3q*	institutions	0.0134	0.0153	0.3789
party.comp	*cens.ind.3q.3q*	repression + duration	0.0137	0.0153	0.3689
party.comp	*cens.ind.3q.3q*	full	0.0089	0.0167	0.5917
party.comp	*cens.ind.1q.1q*	bivariate	0.0087	0.0131	0.5065
party.comp	*cens.ind.1q.1q*	institutions	0.02	0.0153	0.1912
party.comp	*cens.ind.1q.1q*	repression + duration	0.0198	0.0156	0.2047
party.comp	*cens.ind.1q.1q*	full	0.0205	0.0156	0.1893
party.comp	*cens.ind.3q.0q*	bivariate	-0.0034	0.0176	0.8485
party.comp	*cens.ind.3q.0q*	institutions	0.0173	0.0207	0.4021
party.comp	*cens.ind.3q.0q*	repression + duration	0.018	0.0205	0.3783
party.comp	*cens.ind.3q.0q*	full	0.0122	0.0227	0.5925
party.comp	*cens.ind.3q.20q*	bivariate	0.0025	0.0136	0.8554
party.comp	*cens.ind.3q.20q*	institutions	0.0233	0.0159	0.1425
party.comp	*cens.ind.3q.20q*	repression + duration	0.0236	0.0159	0.1373
party.comp	*cens.ind.3q.20q*	full	0.0218	0.0181	0.2288
military	*cens.ind.3q.3q*	bivariate	-0.0027	0.019	0.8857
military	*cens.ind.3q.3q*	institutions	-0.0073	0.0183	0.6909
military	*cens.ind.3q.3q*	repression + duration	-0.004	0.0197	0.8373
military	*cens.ind.3q.3q*	full	-0.0299	0.0235	0.202
military	*cens.ind.1q.1q*	bivariate	-0.0019	0.017	0.9122
military	*cens.ind.1q.1q*	institutions	-0.0044	0.0171	0.7966
military	*cens.ind.1q.1q*	repression + duration	-0.004	0.0199	0.8396
military	*cens.ind.1q.1q*	full	-0.0534	0.0227	0.0185
military	*cens.ind.3q.0q*	bivariate	0.0158	0.0247	0.523
military	*cens.ind.3q.0q*	institutions	0.0098	0.0241	0.6841
military	*cens.ind.3q.0q*	repression + duration	0.0096	0.0263	0.7158
military	*cens.ind.3q.0q*	full	-0.0259	0.0325	0.4263
military	*cens.ind.3q.20q*	bivariate	0.0171	0.0198	0.3876
military	*cens.ind.3q.20q*	institutions	0.0117	0.0188	0.5331
military	*cens.ind.3q.20q*	repression + duration	0.0072	0.0203	0.7216
military	*cens.ind.3q.20q*	full	-0.0165	0.0254	0.5143

Table A4: Results of Cross-national Analysis (2/2)

Independent Variable	Dependent Variable	Covariates	Estimate	SE	p-value
executive.comp	*cens.ind.3q.3q*	bivariate	-0.0193	0.0114	0.0893
executive.comp	*cens.ind.3q.3q*	institutions	-0.0262	0.0138	0.0582
executive.comp	*cens.ind.3q.3q*	repression + duration	-0.0224	0.0162	0.1675
executive.comp	*cens.ind.3q.3q*	full	-0.0307	0.0176	0.0815
executive.comp	*cens.ind.1q.1q*	bivariate	-0.0098	0.0114	0.3912
executive.comp	*cens.ind.1q.1q*	institutions	-0.0194	0.0135	0.1504
executive.comp	*cens.ind.1q.1q*	repression + duration	-0.02	0.0155	0.1987
executive.comp	*cens.ind.1q.1q*	full	-0.0381	0.0157	0.0153
executive.comp	*cens.ind.3q.0q*	bivariate	-0.0271	0.0153	0.0776
executive.comp	*cens.ind.3q.0q*	institutions	-0.0341	0.0185	0.0662
executive.comp	*cens.ind.3q.0q*	repression + duration	-0.0314	0.0212	0.1382
executive.comp	*cens.ind.3q.0q*	full	-0.0444	0.0235	0.059
executive.comp	*cens.ind.3q.20q*	bivariate	-0.0246	0.012	0.0406
executive.comp	*cens.ind.3q.20q*	institutions	-0.0341	0.0144	0.0178
executive.comp	*cens.ind.3q.20q*	repression + duration	-0.0361	0.0169	0.0329
executive.comp	*cens.ind.3q.20q*	full	-0.0451	0.019	0.0176
duration	*cens.ind.3q.3q*	bivariate	0.0003	0.0002	0.2522
duration	*cens.ind.3q.3q*	institutions	0.0001	0.0003	0.721
duration	*cens.ind.3q.3q*	repression + duration	0.0002	0.0003	0.5593
duration	*cens.ind.3q.3q*	full	-0.0001	0.0004	0.8622
duration	*cens.ind.1q.1q*	bivariate	0.0001	0.0002	0.6291
duration	*cens.ind.1q.1q*	institutions	0	0.0003	0.9818
duration	*cens.ind.1q.1q*	repression + duration	0	0.0003	0.8972
duration	*cens.ind.1q.1q*	full	-0.0006	0.0003	0.0747
duration	*cens.ind.3q.0q*	bivariate	0.0002	0.0003	0.5849
duration	*cens.ind.3q.0q*	institutions	0	0.0004	0.9807
duration	*cens.ind.3q.0q*	repression + duration	0.0002	0.0004	0.6888
duration	*cens.ind.3q.0q*	full	-0.0002	0.0005	0.7374
duration	*cens.ind.3q.20q*	bivariate	0.0001	0.0003	0.8298
duration	*cens.ind.3q.20q*	institutions	-0.0001	0.0003	0.6795
duration	*cens.ind.3q.20q*	repression + duration	-0.0001	0.0003	0.8572
duration	*cens.ind.3q.20q*	full	-0.0003	0.0004	0.4301