

## Online Appendix Tables

### 1) Descriptive statistics, time coverage, and data sources

Variable	No of obs.	Mean	Std. Dev.	Min	Max	Years of data	Source
Proportion of population excluded	7138	0.1576534	0.225158	0	0.98	1945-2005	Wimmer, Cederman, and Min 2009
Change in the size of the excluded population during the next five years	6365	-0.0005764	0.1874977	-0.98	0.98	1945-2000	Wimmer, Cederman, and Min 2009
Number of ethnopolitically relevant groups	7138	4.133091	6.410085	0	57	1945-2005	Wimmer, Cederman, and Min 2009
Size of largest politically relevant ethnic group (in %)	7138	0.5239445	0.3243058	0	0.988	1945-2005	Wimmer, Cederman, and Min 2009
Proportion of years under imperial rule since 1816	7155	0.4749055	0.3144366	0	1	1945-2005	Wimmer and Min 2009
Former Spanish dependency	7155	0.1677149	0.373639	0	1	1945-2005	Wimmer and Feinstein 2010
Former Hapsburg dependency	7155	0.0600978	0.2376846	0	1	1945-2005	Wimmer and Feinstein 2010
Former Ottoman dependency	7155	0.1185185	0.3232437	0	1	1945-2005	Wimmer and Feinstein 2010
Former Russian dependency	7155	0.0377358	0.19057	0	1	1945-2005	Wimmer and Feinstein 2010
Former French dependency	7155	0.1587701	0.3654871	0	1	1945-2005	Wimmer and Feinstein 2010
Former British dependency	7155	0.2665269	0.4421738	0	1	1945-2005	Wimmer and Feinstein 2010
Former Portuguese dependency	7155	0.0283718	0.1660442	0	1	1945-2005	Wimmer and Feinstein 2010
Former dependency of other empires	7155	0.1861635	0.3892658	0	1	1945-2005	Wimmer and Feinstein 2010
Number of memberships in IGOs	7151	48.02808	22.4178	0	134	1945-2005	Pevehouse et al. 2004
Democracy, lagged	7041	0.3469678	0.4760392	0	1	1945-2005	Polity II
Democratic transition during next five years	6932	0.076457	0.2657471	0	1	1945-2005	Polity II
Democratic transition during the past ten years	6932	0.0692441	0.2538871	0	1	1945-2005	Polity II
Fully proportional systems	3468	0.3638985	0.4811893	0	1	1946-2002	Gerring and Thacker 2008
Fully parliamentary systems	3465	0.4836941	0.4998062	0	1	1946-2002	Gerring and Thacker 2008
% literates among the adult population	7155	64.30083	30.41654	1.3	99	1945-2005	Wimmer and Feinstein 2010
% literates among the adult population in 1900	6931	25.99416	29.33812	0	96	1900, fixed	Wimmer and Feinstein 2010
Number of associations per capita	4628	0.0097242	0.0204929	0.0000214	0.1735801	1970-2005	Schofer and Langhofer 2011
Linguistic fractionalization	7155	0.3812028	0.2843958	0.001	0.9250348	1960, fixed	Fearon and Laitin 2003
Length of railway tracks (km) per 1000km2	7155	19.27059	29.55708	0	153.6556	1945-2005	Wimmer and Feinstein 2010
Length of railway tracks (km) per 1000km2 in 1900	7155	13.09411	26.75258	0	137.8248	1900, fixed	Wimmer and Feinstein 2010
Proportion of pop. governed by states before colonization	3778	0.5321575	0.4131228	0	1	pre-1900, fixe	Müller 1999
Proportion of excluded population at first year of data	7155	0.1549871	0.2254832	0	0.98	Varying	Wimmer, Cederman, and Min 2009
Number of years since 1816 with constant borders	6709	106.86	56.6638	0	190	1945-2005	Wimmer and Feinstein 2010
GDP per capita in thousands of dollars, lagged	6990	5.968217	7.292209	0.0278672	110.3153	1946-2005	Penn World Table
Difference between highest and lowest elevation (in meters)	7046	3214.903	2007.652	53	9002	Fixed	Fearon and Laitin 2003
Number of ethno-national wars fought between 1816 and first year of data	6833	2.789843	3.184423	0	22	1945-2005	Wimmer and Min 2006
Political instability (change of pol. regime during past 3 years)	7155	0.1220126	0.3273232	0	1	1945-2005	Polity II

## 2) Table 2, Model 2: Count of INGOs (lingopc) instead of IGOs

Generalized linear models No. of obs = 5109  
(Std. Err. adjusted for 154 clusters in cowcode)

---

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
lingopc	-.1318473	.0802475	-1.64	0.100	-.2891295	.025435
groups	.0434243	.0222454	1.95	0.051	-.0001758	.0870245
maxpop	.1409329	.3574505	0.39	0.693	-.5596573	.8415231
_Syear1	.0056577	.0141641	0.40	0.690	-.0221033	.0334188
_Syear2	-.0057757	.0115748	-0.50	0.618	-.028462	.0169105
_cons	-13.48365	28.04286	-0.48	0.631	-68.44664	41.47934

---

## 3) Table 2, Model 3: Democracy and ethno-political inclusion: Full sample

Generalized linear models No. of obs = 6969  
(Std. Err. adjusted for 155 clusters in cowcode)

---

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
democl	-.9460328	.2193976	-4.31	0.000	-1.376044	-.5160214
groups	.0324582	.0153138	2.12	0.034	.0024436	.0624727
maxpop	.215526	.3668696	0.59	0.557	-.5035252	.9345773
_Syear1	.0079686	.0073232	1.09	0.277	-.0063846	.0223218
_Syear2	-.0031918	.0075995	-0.42	0.674	-.0180866	.0117031
_cons	-17.40036	14.41485	-1.21	0.227	-45.65295	10.85223

---

## 4) Table 2, Model 3: PR system (pr), parliamentarism (fullparl2), and ethno-political exclusion: Usir dataset

Generalized linear models No. of obs = 4045  
(Std. Err. adjusted for 154 clusters in cowcode)

---

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
pr	-.0000654	.0002122	-0.31	0.758	-.0004813	.0003506
groups	.034444	.0155476	2.22	0.027	.0039713	.0649166
maxpop	.3873037	.361452	1.07	0.284	-.3211292	1.095737
_Syear1	.008753	.0311605	0.28	0.779	-.0523204	.0698264
_Syear2	-.0011355	.022851	-0.05	0.960	-.0459225	.0436515
democl	-.8682988	.2625135	-3.31	0.001	-1.382816	-.3537818
_cons	-19.18387	61.50296	-0.31	0.755	-139.7275	101.3597

---

Generalized linear models No. of obs = 4180  
(Std. Err. adjusted for 154 clusters in cowcode)

---

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
fullparl2	-.21814	.2619788	-0.83	0.405	-.731609	.295329
groups	.032043	.0159273	2.01	0.044	.000826	.0632599
maxpop	.5893172	.3790728	1.55	0.120	-.1536518	1.332286
_Syear1	-.0145297	.029571	-0.49	0.623	-.0724877	.0434283
_Syear2	.0108239	.0216516	0.50	0.617	-.0316126	.0532603
democl	-.4979894	.2462309	-2.02	0.043	-.980593	-.0153858
_cons	27.07891	58.36285	0.46	0.643	-87.31016	141.468

---

5) Table 2, Model 3: Parliamentary systems (parliamentiaed) and PR systems (propiaep) measured by the IAEP dataset

Generalized linear models No. of obs = 3730  
(Std. Err. adjusted for 148 clusters in cowcode)

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
democl	-.8144472	.2520323	-3.23	0.001	-1.308421	-.320473
propiaep	.0100726	.2669819	0.04	0.970	-.5132024	.5333476
groups	.0354641	.0244186	1.45	0.146	-.0123956	.0833238
maxpop	.0347939	.3452725	0.10	0.920	-.6419277	.7115155
_Syear1	.0277593	.026667	1.04	0.298	-.024507	.0800257
_Syear2	-.0171608	.0195853	-0.88	0.381	-.0555473	.0212257
_cons	-56.50911	52.62305	-1.07	0.283	-159.6484	46.63016

Generalized linear models No. of obs = 4035  
(Std. Err. adjusted for 154 clusters in cowcode)

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
democl	-.8778607	.2476084	-3.55	0.000	-1.363164	-.3925571
parliamentiaed	-.1004006	.6652717	-0.15	0.880	-1.404309	1.203508
groups	.0339564	.0152626	2.22	0.026	.0040423	.0638705
maxpop	.2154869	.3498683	0.62	0.538	-.4702423	.9012161
_Syear1	.0165063	.0244007	0.68	0.499	-.0313182	.0643308
_Syear2	-.0088911	.0182114	-0.49	0.625	-.0445848	.0268025
_cons	-34.3559	48.15663	-0.71	0.476	-128.7412	60.02937

6) Table 2, Model 3: Parliamentary systems and PR systems: Restricting observations to democracies

Generalized linear models No. of obs = 2279  
(Std. Err. adjusted for 98 clusters in cowcode)

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
fullparl	-.5856844	.3751614	-1.56	0.118	-1.320987	.1496185
fullprop	.3511789	.3949419	0.89	0.374	-.4228931	1.125251
groups	.1044723	.0579691	1.80	0.072	-.0091451	.2180898
maxpop	.0820008	.4492789	0.18	0.855	-.7985698	.9625713
_Syear1	.0217465	.0126619	1.72	0.086	-.0030704	.0465634
_Syear2	-.021846	.0139411	-1.57	0.117	-.04917	.005478
_cons	-45.39446	24.82527	-1.83	0.067	-94.05111	3.262179

7) Table 2, Model 4 with 5-year time lag (recentdemoc)

Linear regression Number of obs = 6176  
(Std. Err. adjusted for 154 clusters in cowcode)

futureexcl	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
recentdemoc	-.0245555	.0164051	-1.50	0.136	-.0569652	.0078542
groups	.0012925	.0008122	1.59	0.114	-.0003121	.0028971
maxpop	.005861	.008874	0.66	0.510	-.0116704	.0233924
_Syear1	.0003134	.0008422	0.37	0.710	-.0013504	.0019772
_Syear2	-.0007417	.0009485	-0.78	0.435	-.0026157	.0011322
_cons	-.6177331	1.657534	-0.37	0.710	-3.892341	2.656875

8) Table 2, Model 5 with 10 year time lag (futuredemoc1) and in a fixed effects specification

Logistic regression Number of obs = 4439  
(Std. Err. adjusted for 122 clusters in cowcode)

futuredemoc1	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
rexclpop	-.915189	.4666583	-1.96	0.050	-1.829822	-.0005556
groups	-.0185953	.0129054	-1.44	0.150	-.0438895	.0066989
maxpop	.9460149	.4286592	2.21	0.027	.1058583	1.786172
_Syear1	-.0169985	.0072411	-2.35	0.019	-.0311907	-.0028063
_Syear2	-.0094503	.0114242	-0.83	0.408	-.0318413	.0129408
_cons	31.3754	14.23777	2.20	0.028	3.469885	59.28091

Conditional fixed-effects logistic regression Number of obs = 3447  
 Group variable: cowcode Number of groups = 70

democ	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
rexclpop	-1.817067	.5654987	-3.21	0.001	-2.925424	-.7087103
lgdpcapl	.1621129	.2099433	0.77	0.440	-.2493683	.5735942
groups	.0464264	.0450175	1.03	0.302	-.0418064	.1346591
maxpop	-2.982966	1.581084	-1.89	0.059	-6.081832	.1159012
_Syear1	.0123901	.0111923	1.11	0.268	-.0095465	.0343267
_Syear2	.101602	.0108513	9.36	0.000	.0803338	.1228702

9) Literacy and “willingness to fight for country” (fightcountry) based on World Value Survey data

Source	SS	df	MS	Number of obs = 77		
Model	.382998906	7	.054714129	F( 7, 69) = 3.03		
Residual	1.24502503	69	.018043841	Prob > F = 0.0077		
				R-squared = 0.2353		
				Adj R-squared = 0.1577		
Total	1.62802394	76	.021421368	Root MSE = .13433		

fightcountry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
literacy	.000183	.0009362	0.20	0.846	-.0016846	.0020506
gdpcapl	-.0085715	.0032587	-2.63	0.011	-.0150725	-.0020706
democ1	-.0150502	.0394823	-0.38	0.704	-.0938153	.0637149
cinc	-.3687816	1.155498	-0.32	0.751	-2.673937	1.936374
oilpc	.0086757	.0042826	2.03	0.047	.0001323	.0172192
popavg	1.37e-07	1.53e-07	0.89	0.374	-1.68e-07	4.42e-07
ethfrac	-.0121291	.062872	-0.19	0.848	-.1375553	.113297
_cons	.809407	.0846328	9.56	0.000	.6405692	.9782449

10) Using government expenditure per capita as a proxy for state centralization for a sub-sample of independent countries of Europe and the Americas

a) With gov. exp. in 1920 (stateexp1920), only countries that lack data on precolonial state centralization

Generalized linear models No. of obs = 2140

(Std. Err. adjusted for 39 clusters in cowcode)

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
stateexp1920	-3.237166	1.280355	-2.53	0.011	-5.746616	-.7277161
groups	.638886	.1748914	3.65	0.000	.2961052	.9816668
maxpop	-2.19354	1.010987	-2.17	0.030	-4.175038	-.2120419
_Syear1	-.0032977	.0069835	-0.47	0.637	-.016985	.0103897
_Syear2	.0039849	.0054476	0.73	0.464	-.0066923	.0146621
_cons	4.796091	14.02521	0.34	0.732	-22.69281	32.28499

a) With gov. exp. in 1935 (stateexp1935), all countries with data available

Generalized linear models No. of obs = 3204

(Std. Err. adjusted for 68 clusters in cowcode)

exclpop	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
stateexp1935	-1.739544	.8403788	-2.07	0.038	-3.386656	-.0924314
groups	.0139081	.0090906	1.53	0.126	-.0039092	.0317255
maxpop	-.0637239	.5962196	-0.11	0.915	-1.232293	1.104845
_Syear1	.0026321	.0093004	0.28	0.777	-.0155963	.0208606
_Syear2	-.0006165	.0102846	-0.06	0.952	-.0207741	.019541
_cons	-6.386456	18.14118	-0.35	0.725	-41.94252	29.16961