

Online Appendix I: Robustness analyses

This document reports diagnostics and robustness tests for the analysis presented in:

Sarah Shair-Rosenfield, Arjan H. Schakel, Sara Niedzwiecki, Gary Marks, Liesbet Hooghe, Sandra Chapman-Osterkat. 2021. "Language difference and regional authority" *Regional and Federal Studies* 31 (1): 73-97 (anniversary issue). Replication datasets are available on <https://garymarks.web.unc.edu>

I. EXPLAINING THE LIKELIHOOD OF A LANGUAGE REGION	2
Descriptives	2
Table A.1: Basic descriptives	2
Table A.2: Correlation matrix	2
Robustness analyses (for Table 1 and Figure 6)	3
Table A.3: Explaining the likelihood of a language region: robustness analyses	5
II. THE EFFECT OF LANGUAGE DISTINCTIVENESS ON REGIONAL AUTHORITY: GENERAL MODEL	6
Descriptives	6
Table A.4: Basic descriptives	6
Table A.5: Correlation matrix	6
Robustness analyses (for Table 2)	7
Table A.6: Explaining regional authority: robustness analyses (Table 2 in paper)	8
III. MECHANISM I: EFFECT OF LANGUAGE DISTINCTIVENESS ON A REGION'S AUTHORITY	9
Analysis supporting Figure 7	9
Table A.7: The effect of language difference for a region's authority conditional on democracy ..	10
IV. MECHANISM II: SYSTEM-WIDE EFFECT OF LANGUAGE DISTINCTIVENESS ON A REGIONAL TIER'S AUTHORITY	11
Descriptives	11
Table A.8: Descriptives (country-level)	12
Table A.9: Correlation matrix (country-level)	13
Analysis supporting Figure 8	14
Table A.10: Democracy, language regions, and their country-wide effect	14
V. DISTRIBUTION OF LANGUAGE REGIONS IN THE SAMPLE	15
Figure A.11: Distribution of language regions	15
VI. REGIONS BY COUNTRY: ROKKAN CHARACTERISTICS, REGIONAL VARIABLES, COUNTRY VARIABLES	16
Table A.12: Regions by country: Rokkan characteristics, regional variables, country variables	16

I. EXPLAINING THE LIKELIHOOD OF A LANGUAGE REGION

Descriptives

This section provides supplementary material on the first part in the causal model: explaining the incidence of language regions.

Table A.1: Basic descriptives

	Mean	median	Min	Max	sd	N
Language region	0.20	0.00	0.00	1.00	0.40	1767
Former state core	0.11	0.00	0.00	1.00	0.32	1767
Tribal	0.29	0.00	0.00	1.00	0.46	1767
Island-region	0.03	0.00	0.00	1.00	0.16	1767
Distance (km)	512.91	272.00	1.00	7772.00	749.75	1767
Distance (km, log 10)	2.33	2.43	0.00	3.89	0.75	1767
Settlers	43.90	20.00	0.00	100.00	44.67	1767
Linguistic fractionalization	0.29	0.22	0.00	0.84	0.24	1767

Table A.2: Correlation matrix

	Language region	Former state core	Tribal	Island-region	Distance	Distance (log10)	Settlers
Language region	1.00						
Former state core	0.22	1.00					
Tribal	0.15	-0.05	1.00				
Island-region	0.17	0.06	0.05	1.00			
Distance	0.13	0.07	0.18	0.20	1.00		
Distance (log10)	0.22	0.13	0.15	0.14	0.61	1.00	
Settlers	-0.25	0.08	-0.15	0.00	0.07	-0.07	1.00
Linguistic fractionalization	0.47	0.19	0.08	0.05	0.08	0.08	-0.28

Robustness analyses (for Table 1 and Figure 6)

Table A.3 reports robustness analyses alongside the models reported in Table 1 of the paper, and to aid comparison, we reprint Model 1 (Table 1) here.

Model 2 assesses whether the patterns we discover are driven by particular countries. A jackknife analysis that drops one country at a time reveals that the odds ratios for the geographical and political variables of interest remain robust.

Model 3 imposes dummies for worldregions to account for omitted variables that capture systematic variation across Asia, Europe, the New World, and Latin America. The reference category is Europe (which includes here Russia and Turkey). The explanatory model gains sharpness, suggesting that omitted worldregion-specific factors affect the baseline for comparison (change the constant). However, this does not invalidate the model, on the contrary, and additional analyses show that the model holds in each worldregion when we partition the sample into the four world regions.

Model 4 reprises the analysis without the small states that do not have regional government in 2018, without substantive changes in the coefficients compared to model 1.

Model 5 reprises the analysis for ex-colonies, and model 6 adds colonial empire dummies. Variables picking up political insulation strengthen.

Model 7 employs a more selective classification of language region as a region where a majority speaks the same non-core language. This reduces the number of language regions from 345 to 231. For example, no linguistically distinctive regions in Papua New Guinea, only six of eight regions with language difference in Pakistan, and just two of four in Mexico are still classified as language regions. The model does a credible job in explaining language difference, though the explained variance is lower.

Further diagnostics show also that Model 1 is efficient at predicting the likelihood of a language region. The area under the ROC curve, the concordance statistic, is 0.88, which is well above the 0.7 benchmark for improvement over random chance. The Hosmer-Lemeshow χ^2 which tests the observed against the expected number of responses is not significant ($\chi^2=2.07$ and $p=0.56$) indicating that one cannot reject the model.

Table A.3: Explaining the likelihood of a language region: robustness analyses

	Model 1 (Table 1 in the paper)		Model 2 jackknife (dropping one country at a time)		Model 3 controlling for worldregions		Model 4 excluding countries that had no regional government in 2018		Model 5 limiting to regions in ex-colonies		Model 6 Controlling for colonial empire		Model 6 restrictive language region: a majority speaks the same non-core language	
	Odds ratio	s.e.	Odds ratio	s.e.	Odds ratio	s.e.	Odds ratio	Odds ratio	Odds ratio	s.e.	Odds ratio	s.e.	Odds ratio	s.e.
Island-region	5.53***	2.29	5.53***	2.49	5.45***	2.16	5.56***	2.31	6.18°	5.87	5.40°	4.69	2.48	1.60
Distance	3.26***	1.04	3.26**	1.17	3.10***	1.00	3.18**	1.06	3.32**	1.24	3.46**	1.34	2.46***	0.85
Former state core	5.23***	1.36	5.23***	1.49	5.41***	1.35	5.23***	1.36	6.15***	2.76	6.09***	2.45	3.42***	0.91
Tribal	7.33***	2.92	7.33***	3.64	11.95***	6.42	7.37***	2.95	8.42***	3.49	18.26***	10.60	1.46	0.65
Settlers	1.00	0.00	1.00	0.00	0.99***	0.00	1.00	0.00	0.99	0.01	1.01	0.01	1.00	0.00
Tribal* Settlers	0.95***	0.00	0.95*	0.02	0.94***	0.02	0.95***	0.01	0.95**	0.02	0.93***	0.01	0.98*	0.01
Controls														
Linguistic fractionalization	42.82***	22.72	42.82***	28.60	34.92***	24.03	42.61***	22.68	43.96***	28.47	43.48***	19.43	38.87***	25.32
World region (ref = Europe)														
Asia					0.46*	0.14								
Latin America					0.32*	0.17								
New World					2.55	2.40								
Colonial empire (ref=British)														
French											1.14	0.43		
Dutch											0.74	0.27		
Portuguese											0.20	0.28		
Spanish											0.36	0.17		
Constant	0.00***	0.00	0.00***	0.00	0.01***	0.01	0.00***	0.00	0.02***	0.00	0.00***	0.00	0.00***	0.00***
Number of observations	1767		1767		1767		1749		992		992		1767	
McFadden R ²	0.37		0.37		0.38		0.36		0.42		0.44		0.25	
Tjur's D coefficient	0.40		0.40		0.41		0.40		0.47		0.49		0.23	
Wald Chi ² (df)	228.8 (7)				173.6 (10)		225.5(7)		198.1(7)		570.3(11)		74.9(7)	
AIC	1121.4		1121.4		1111.4		1120.6		671.8		664.4		1043.0	
BIC (df)	1165.2 (8)		1165.2 (8)		1171.7(11)		1164.3(8)		711.0(8)		723.2(12)		1086.9(8)	

Note: Logistic regression with cluster-robust standard errors (95 country clusters). Significance: *** p<.001 ** p<.01 * p<.05. °p<.10

II. THE EFFECT OF LANGUAGE DISTINCTIVENESS ON REGIONAL AUTHORITY: GENERAL MODEL

Descriptives

This section provides supplementary material on the second part of the causal analysis: how language regions affect regional authority.

Table A.4: Basic descriptives

	Mean	median	Min	Max	sd	N
RAI of region	11.08	9.00	1.00	28.00	7.00	1698
RAI difference	0.13	0.00	-18.00	19.00	2.08	1691
Language region	0.18	0.00	0.00	1.00	0.39	1698
Prior state core	0.11	0.00	0.00	1.00	0.32	1698
Region's population	3,165,296	761,322	1424	199,812,341	8,753,504	1698
Region's area (km ²)	53338	7284	14	3103200	194293	1698
Region's population (log 10)	5.90	5.88	3.15	8.30	0.66	1698
Region's area (km ² , log10)	3.96	3.86	1.15	6.49	0.78	1698
Democracy	0.51	0.52	0.07	0.89	0.25	1698
GDP per capita (US\$, constant)	22333	17382	1669	75916	16841	1698
GDP per capita (log10)	4.22	4.24	3.22	4.88	0.36	1698
Linguistic fractionalization	0.29	0.22	0.00	0.84	0.25	1698

Table A.5: Correlation matrix

	RAI of region	RAI difference	Language region	Prior state core	Population	Area	Democracy	GDP/capita
RAI of region	1.00							
RAI_difference	0.20	1.00						
Language region	0.02	0.06	1.00					
Prior state core	0.40	0.04	0.23	1.00				
Region's population (log)	0.31	0.01	0.07	0.18	1.00			
Region's area (log10)	0.40	0.04	0.03	0.17	0.40	1.00		
Democracy	0.38	0.06	-0.16	0.10	-0.14	0.02	1.00	
GDP per capita (log10)	0.41	0.07	-0.19	0.10	-0.12	0.11	0.64	1.00
Ling. Fraction.	0.02	-0.07	0.48	0.20	0.11	-0.04	-0.20	-0.22

Robustness analyses (for Table 2)

Table A.6 summarizes robustness analyses for Table 2 in the paper, with Model 2 (Table 2) reprinted as benchmark for evaluation.

Model 3 displays a jackknife analysis showing that the results of Model 3 are robust to country outliers. The analysis shows that the coefficients are stable. The only slight difference is that *Language region*, alone and interacted with democracy, is now significant at the 0.05 level.

Model 4 reveals that dummies for the four world regions do not reach significance suggesting that the model has causal power across world regions. Follow-up analyses that drop consecutively one of the four world regions reveal that the interaction between language and democracy appears sensitive to the inclusion of regions in Europe. Excluding Europe, we find that democracy is all-important as a predictor of regional authority, closely followed by regional population size and area size, and next by prior state core. Language region is not significant.

Model 5 reveals that our results are not sensitive to using the Polity2 measure of democracy in place of the Varieties of Democracy measure.

Model 6 shows that our results are virtually identical when we use a restrictive definition of language region, as a region in which a majority speaks the same distinctive non-core language.

Model 7 substitutes regional self-rule for regional authority. Regional self-rule refers to the authority of a regional government exercised in its territory (rather than authority co-exercised in the country as a whole). Results are robust.

Table A.6: Explaining regional authority: robustness analyses (Table 2 in paper)

	Model 2 (Table 2 in paper)		Model 3 jackknife (dropping one country at a time)		Model 4 controlling for world regions		Model 5 using Polity for democracy		Model 6 restrictive language region: a majority speaks the same non- core language		Model 7 regional self-rule as dependent variable	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Democracy facilitates voice												
Language region	-4.03*	1.54	-4.03°	2.19	-3.88**	1.33	-5.23**	1.47	-4.32*	1.71	-1.47	0.91
Democracy in country	5.08	5.18	5.08	7.13	4.51	4.95	0.13	0.13	5.25	5.14	5.15	2.70
Language * Democracy	9.02**	3.37	9.02*	4.40	9.12**	2.89	0.35**	0.10	10.81**	3.49	4.28*	1.87
Controls												
Former state core	5.56***	1.40	5.56**	1.69	5.46***	1.46	5.71***	1.53	5.32***	1.39	2.07*	0.94
Population size (region)	2.42***	0.62	2.42**	0.76	2.51***	0.60	2.46***	0.63	2.47***	0.63	1.74***	0.40
Area size (region)	2.16*	0.81	2.17°	1.11	1.93*	0.85	2.03*	0.81	2.16*	0.82	1.07*	0.46
GDP per capita (country)	5.09°	2.77	5.09	3.71	4.90°	2.59	6.52**	1.90	1.11°	2.80	3.08	1.61
Ling. fraction. (country)	1.56	2.23	1.56	2.90	1.53	2.44	0.17	2.44	1.43	2.16	0.41	2.08
World region (ref = Europe)												
Asia					-0.30	1.90						
Latin America					0.22	1.83						
New World					2.28	2.45						
<i>Constant</i>	-36.85**	10.77	-36.85**	14.02	-35.44**	11.23	-41.66***	8.60	-37.32**	10.98	-21.61**	6.60
Observations	1698		1698		1698		1698		1698		1698	
R ²	0.47		0.47		0.48		0.46		0.48		0.46	
F (df)	15.0 (8)		10.9 (8)		13.71 (11)		12.19 (8)		17.98 (8)		11.4 (8)	

Note: OLS regression with cluster-robust standard errors by country (74 clusters). Significance: *** p<.001 ** p<.01 * p<.05.

III. MECHANISM I: EFFECT OF LANGUAGE DISTINCTIVENESS ON A REGION'S AUTHORITY

Here we examine the first mechanism by which language may affect regional authority. Can we say that language regions are more likely to have different authority—more or less—than regions that do not have linguistic distinctiveness? The descriptives were provided in Table A.4 and Table A.5, so we go straight to analysis.

[Analysis supporting Figure 7](#)

Table A.7 reports the model statistics that produce Figure 7 in the paper. Like in prior analyses, the unit of the analysis is the individual region. The dependent variable here is *RAI difference*, which is calculated by subtracting from the RAI of a region the RAI of standard regions in the most authoritative general-purpose tier. The analysis excludes regions that are not part of a regional tier, such as the Indian Tribes in the US. Regions in Belgium and Bosnia-Herzegovina are also excluded because all regions are asymmetrical and so there is no standard tier. Hence the size of the sample is slightly smaller than in the previous section.

Model 1 shows the results that produce Figure 7. It shows that, as a country is more democratic, a language region is statistically more likely to have higher regional authority than a non-language region, but the substantive effect is small.

Model 2 tests whether these results are robust when including regions in Belgium and Bosnia and Herzegovina. To do this, we need to estimate a hypothetical standard RAI score to set against individual regions' scores. There is no self-evident standard, and so we compare several operationalizations. The one reported here seems the most conservative:

the RAI score of the “least asymmetrical region” in each country is used as a reference point, i.e. French community in Belgium and Federacie in BIH. Alternative standards might be a RAI score of zero for hypothetical standard regions on account that this is the true score, or using historical standard regions as a baseline, i.e. Belgian provinces or cantons in BIH.

Table A.7: The effect of language difference for a region’s authority conditional on democracy

	Model 1		Model 2	
	Base line model		Including regions in Belgium & Bosnia-Herzegovina	
	Coeff.	s.e.	Coeff.	s.e.
Language region	-0.10	0.29	-0.13	0.29
Democracy	-0.29	0.28	-0.28	0.28
Language region * Democracy	1.54*	0.60	1.64**	0.59
Controls (region)				
Former state core	0.30°	0.17	0.28°	0.17
Population	-0.04	0.09	-0.02	0.09
Area	0.11	0.07	0.10	0.07
Controls (country)				
GDP per capita	0.19	0.19	0.20	0.19
Linguistic fractionalization	-1.74***	0.24	-1.72***	0.24
<i>Constant</i>	-0.42**	0.87	-0.58	0.87
Observations		1691		1698
R ²		0.041		0.041
F (df)		8.91 (8)		8.88(8)

Note: OLS regression. Significance: *** p<.001 ** p<.01 * p<.05 ° p<.10

IV. MECHANISM II: SYSTEM–WIDE EFFECT OF LANGUAGE DISTINCTIVENESS ON A REGIONAL TIER’S AUTHORITY

Descriptives

Now we shift the analysis to the country-level, and the dependent variable becomes the regional authority (RAI) of the most authoritative tier in the country. Can we establish that language regions have a knock-on effect on regional authority in the country as a whole?

The dependent variable *RAI of the tier* is calculated in two ways:

- Unweighted average of the RAI of all regions – standard and non-standard – in a tier. So this excludes self-governing territorial communities that exist outside a tier, such as special arrangements for indigenous communities in Canada, the United States, and several Latin American countries. However, indigenous regions that exist in the same tier as other regions (e.g. the five comarcas in Panama) are included.
- Average of the RAI of all regions (see above) weighted by their population.
- We estimate *Language regions* and *Former state core regions* in the same way: as the percentage of regions that are language or former state core, and as the percentage of the population in a country that lives in language regions or in former state core regions.

The unit of the analysis is the country, and this analysis applies to 74 countries. This excludes the 18 countries that have no regions that meet our criteria, as well as three countries for which data on democracy or GDP are missing.

Table A.8: Descriptives (country-level)

	Mean	median	Min	Max	sd	N
RAI of tier (unweighted mean of RAI of regions)	10.82	9.04	0.91	27.00	7.23	74
RAI of tier (population-weighted mean)	10.97	8.80	0.53	27.00	7.43	74
Language regions (unweighted:% regions)	0.17	0.07	0.00	1.00	0.23	74
Language regions (weighted: % population)	0.15	0.02	0.00	0.98	0.22	74
Former state core (unweighted: % regions)	0.13	0.00	0.00	0.88	0.21	74
Former state core (weighted: % population)	0.14	0.00	0.00	0.98	0.25	74
Country population (log 10)	7.28	7.18	6.06	9.13	0.65	74
Country area (km², log10)	5.40	5.38	3.71	7.23	0.74	74
Democracy	0.58	0.63	0.07	0.89	0.24	74
GDP per capita (log10)	4.24	4.35	3.22	4.88	0.37	74
Linguistic fractionalization	0.29	0.22	0.00	0.84	0.23	74

Table A.9: Correlation matrix (country-level)

	RAI (u)	RAI (w)	Language (u)	Language (w)	State core (u)	State core (w)	Population	Area	Democracy	GDP
RAI of tier (unweighted)	1.00									
RAI of tier (population-weighted)	0.99	1.00								
Language regions (unweighted)	0.14	0.19	1.00							
Language regions (% population)	0.17	0.23	0.94	1.00						
Former state core (% regions)	0.65	0.69	0.37	0.36	1.00					
Former state core (% population)	0.65	0.70	0.35	0.36	0.97	1.00				
Country population (log 10)	0.38	0.40	0.33	0.29	0.29	0.27	1.00			
Country area (km², log10)	0.42	0.44	0.08	0.08	0.28	0.30	0.68	1.00		
Democracy	0.37	0.32	-0.26	-0.21	0.07	0.07	-0.25	-0.14	1.00	
GDP per capita (log10)	0.36	0.32	-0.26	-0.22	0.13	0.12	-0.12	-0.03	0.74	1.00
Linguistic fractionalization	0.14	0.19	0.72	0.71	0.39	0.42	0.13	0.05	-0.17	-0.19

Analysis supporting Figure 8

The table A.10 below reports the models that generated Figure 8 (based on model 1). Model 1 uses unweighted measures of RAI, language, and state core regions, and Model 2 uses population-weighted measures. We can see that the coefficients are very similar, and in either case, we can detect a systemic effect. In democratic regimes, the more language regions in a country, the higher the regional authority. The effect runs opposite in authoritarian regimes: the more language regions, the lower the regional authority.

Table A.10: Democracy, language regions, and their country-wide effect

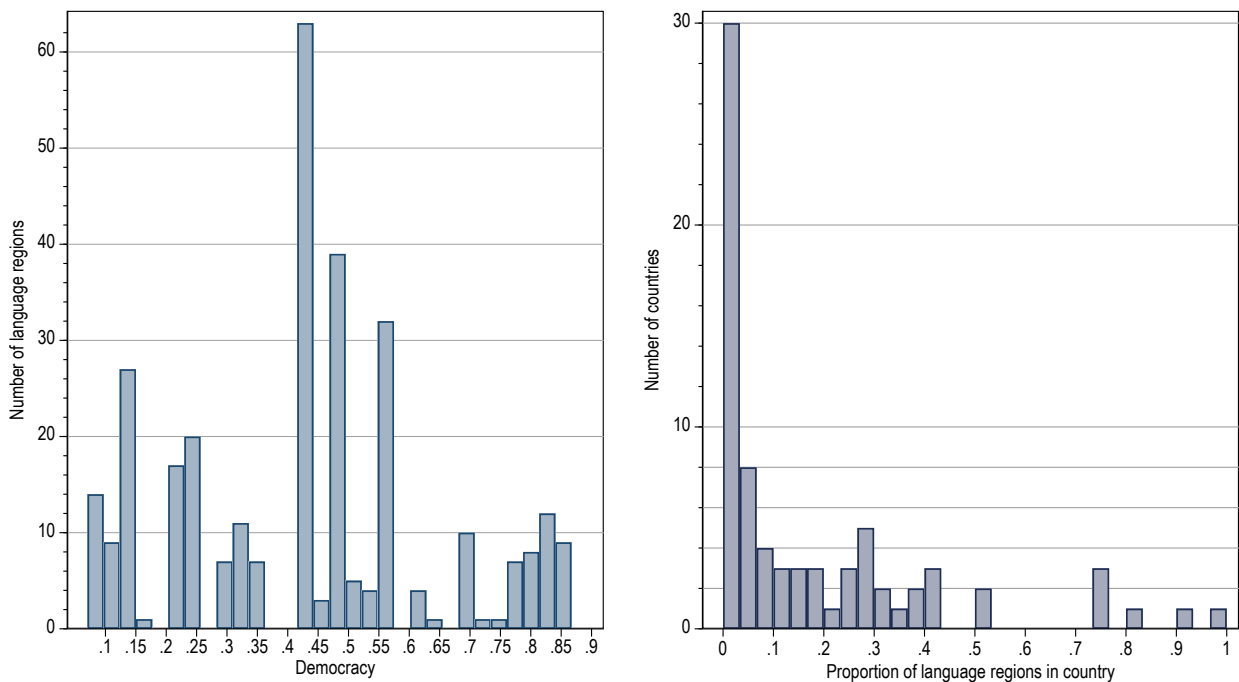
	Model 1 proportion		Model 2 Weighted by population	
	Coeff.	s.e.	Coeff.	s.e.
Language regions	-19.85**	7.08	-14.07°	7.20
Democracy	4.60	3.87	6.19	3.64
Language region * Democracy	35.80**	3.03	28.50**	10.41
Controls (country)				
Former state core regions	18.18***	3.03	17.14***	2.37
Population	2.51*	1.17	2.74*	1.13
Area	2.18*	0.97	1.96*	0.94
GDP per capita	1.53	2.07	0.91	2.02
Linguistic fractionalization	-0.11	3.23	-1.80	3.18
<i>Constant</i>	-30.49**	9.75	-29.07**	9.56
Observations	74		74	
R ²	0.69		0.72	
F (df)	17.94 (8)		20.69 (8)	

Note: OLS regression. Significance: *** p<.001 ** p<.01 * p<.05 ° p<.10.

V. DISTRIBUTION OF LANGUAGE REGIONS IN THE SAMPLE

The Figures below visualize the distribution of language regions across our sample. The figure on the left shows the distribution of 312 language regions in our data by degree of democracy in their country. The number of language regions presented here is lower than that in our dataset (see Table A.12) because data on democracy are missing for several countries, including no fewer than 20 language regions in Papua New Guinea. The figure on the right shows the incidence of language regions, measured as a proportion of regions at the same tier, across the 74 countries in the country-level analysis.

Figure A.11: Distribution of language regions



VI. REGIONS BY COUNTRY: ROKKAN CHARACTERISTICS, REGIONAL VARIABLES, COUNTRY VARIABLES

Table A.12: Regions by country: Rokkan characteristics, regional variables, country variables

Country	Number of regions	Language	Former state core	Over-arching governance	Island-region	Median distance from the capital (km)	Greatest distance from the capital (km)	Median population of region	Median area (km ²) of region	% European settlers in country	Linguistic fract. in country	Democracy (country)	GDPpc (country)
Albania	12	3	0	12	0	70	142	228912	2440	100	0.04	0.47	11941
Argentina*	24	0	13	6	1	926	2372	712240	96856	60	0.06	0.75	16467
Australia*	8	0	6	0	1	909	3133	1798677	892901	98	0.34	0.84	47393
Austria*	9	0	4	9	0	155	501	708220	9532	100	0.15	0.75	46103
Bahamas ^φ	1	0	0	0	0	1	1	395360	13880	10	0.19	.	22026
Bangladesh	65	4	1	62	0	147	1551	1946000	2073	0	0.09	0.24	3436
Barbados ^φ	1	0	0	0	0	1	1	285720	430	20	0.09	0.72	9273
Belgium*	5	4	0	5	0	1	119	3882152	13511	100	0.54	0.83	38341
Belize ^φ	1	0	0	0	0	1	1	374680	22970	20	0.63	.	6777
Bhutan ^φ	1	0	0	1	0	1	1	807600	46683	0	0.61	0.56	.
Bolivia	11	4	1	5	0	290	968	494178	63827	30	0.22	0.52	5838
Bosnia & Herz.*	2	1	1	2	0	62	123	1718664	25565	100	0.68	0.45	11573
Brazil*	27	0	0	0	0	1315	2498	3514952	224301	40	0.05	0.65	13813
Brunei ^φ	1	0	0	1	0	1	1	428700	5770	0	0.34	.	51703
Bulgaria	28	2	0	28	0	178	378	199924	3590	100	0.30	0.55	18640
Cambodia	23	0	0	23	0	148	328	592845	6158	0	0.21	0.17	3161
Canada*	15	4	9	0	0	1674	4139	746936	647797	99	0.58	0.80	42907
Chile	16	0	0	8	0	646	2190	548597	31069	50	0.19	0.81	22729
China	33	14	6	31	1	1126	2560	36900000	167000	0	0.13	0.07	13043
Colombia	35	5	1	0	1	406	1220	898490	23818	20	0.09	0.57	12797
Costa Rica	1	1	0	0	0	116	1	48500	3423	20	0.05	0.86	15044

Country	Number of regions	Language	Former state core	Over-arching governance	Island-region	Median distance from the capital (km)	Greatest distance from the capital (km)	Median population of region	Median area (km ²) of region	% European settlers in country	Linguistic fract. in country	Democracy (country)	GDPpc (country)
Croatia	21	0	0	21	0	130	389	160394	2640	100	0.08	0.63	22996
Cuba	16	0	0	0	1	291	811	564010	6663	25	0.05	0.13	.
Cyprus ^φ	1	0	0	1	0	1	1	855150	9250	100	0.40	0.79	29229
Czech Republic	14	0	0	14	0	107	278	590298	5247	100	0.32	0.76	31141
Denmark	7	2	1	7	2	215	3536	814492	8020	100	0.11	0.87	45875
Dominican Rep.	32	0	0	32	0	103	235	174784	1250	25	0.04	0.46	14781
East Timor ^φ	1	0	0	0	0	1	1	1267970	14870	0	0.53	0.63	.
Ecuador	25	1	0	11	1	172	1236	344044	8010	30	0.13	0.57	9880
El Salvador	15	0	0	9	0	54	153	319503	1240	20	0.18	0.56	6660
Estonia ^φ	1	0	0	1	0	1	1	1309600	45230	100	0.49	0.87	25939
Finland	19	1	0	19	1	234	705	183600	10663	100	0.14	0.83	39893
France	13	2	2	13	1	343	917	4984058	32082	100	0.12	0.81	39461
Germany*	16	1	6	16	0	276	580	3053524	20146	100	0.16	0.81	47556
Greece	14	0	2	14	2	267	379	562647	9327	100	0.03	0.77	23765
Guatemala	22	9	0	0	0	91	264	454989	2600	20	0.46	0.48	7256
Guyana ^φ	1	0	0	0	0	1	1	779000	214970	2	0.07	0.51	.
Haiti	10	0	0	10	0	121	187	507314	2109	0	0.06	0.30	1669
Honduras	18	1	0	0	1	141	391	315410	4271	20	0.05	0.31	4255
Hungary	20	0	0	20	0	120	207	403657	4201	100	0.03	0.49	22572
Iceland ^φ	1	0	0	1	0	1	1	335000	103000	100	0.08	0.82	41306
India*	36	27	19	30	2	1212	2481	14700000	54578	0	0.81	0.49	6422
Indonesia	34	25	10	22	5	950	3779	3589880	43223	0	0.77	0.55	10450
Ireland	31	0	0	31	0	110	260	130425	1859	100	0.03	0.80	75916
Israel	6	1	0	6	0	63	118	1194925	1048	100	0.55	0.63	36186
Italy	21	6	8	21	1	394	595	1816255	13613	100	0.12	0.83	38000
Jamaica ^φ	1	0	0	0	0	1	1	2890300	10990	1.9	0.11	0.75	6805

Country	Number of regions	Language	Former state core	Over-arching governance	Island-region	Median distance from the capital (km)	Greatest distance from the capital (km)	Median population of region	Median area (km ²) of region	% European settlers in country	Linguistic fract. in country	Democracy (country)	GDPpc (country)
Japan	47	0	0	47	0	390	1562	1774622	6087	0	0.02	0.77	40374
Kosovo ^φ	1	0	0	1	0	1	1	1845300	10910	100	0.11	0.45	.
Laos	18	9	3	18	0	319	570	298508	15393	0	0.64	0.10	6518
Latvia	5	2	1	5	0	41	203	301621	13596	100	0.58	0.80	24140
Lithuania	10	0	0	10	0	128	287	182500	6313	100	0.32	0.77	25992
Luxembourg ^φ	1	0	0	1	0	1	1	583000	2590	100	0.64	0.82	74605
Malaysia*	16	4	10	15	3	288	1624	1467208	8668	0	0.60	0.33	22776
Malta ^φ	1	0	0	1	0	1	1	430800	320	100	0.09	0.66	30831
Mexico*	32	4	3	17	0	481	2188	2452807	59128	15	0.15	0.62	17383
Mongolia	1	0	0	1	0	1	1	1154290	4700	0	0.37	0.55	10716
Montenegro ^φ	1	0	0	1	0	1	1	628960	13810	100	0.67	0.40	17889
Myanmar	21	12	4	14	0	352	719	1574079	30383	0	0.51	0.31	5730
Nepal*	7	2	1	4	0	185	481	4499272	175	0	0.72	0.53	2720
Netherlands	12	1	1	12	0	81	178	1152176	2501	100	0.51	0.83	46602
New Zealand	16	0	0	0	0	314	768	144453	13213	93	0.17	0.84	36538
Nicaragua	17	2	2	0	0	111	377	208523	4177	20	0.05	0.14	5210
North Macedonia	8	2	0	8	0	81	119	198738	3011	100	0.50	0.46	13397
Norway	20	0	0	20	1	278	2043	209688	15382	100	0.07	0.89	60590
Pakistan*	8	8	5	6	0	327	780	8673042	73521	0	0.72	0.34	5303
Panama	15	5	0	0	0	150	309	102465	4601	20	0.39	0.70	21928
Papua New Guinea	22	20	0	0	4	524	975	302441	14717	0	0.35	0.38	.
Paraguay	18	0	0	0	0	190	2069	179050	12540	25	0.60	0.52	8135
Peru	26	5	1	21	0	534	1019	800846	34537	30	0.34	0.69	11370
Philippines	83	63	7	17	4	466	1062	722620	3240	1	0.84	0.42	7679
Poland	16	0	0	16	0	253	454	2151025	18265	100	0.05	0.63	26637
Portugal	25	0	0	25	2	219	1447	247453	3238	100	0.02	0.84	24310

Country	Number of regions	Language	Former state core	Over-arching governance	Island-region	Median distance from the capital (km)	Greatest distance from the capital (km)	Median population of region	Median area (km ²) of region	% European settlers in country	Linguistic fract. in country	Democracy (country)	GDPpc (country)
Romania	42	2	0	42	0	222	448	443594	5550	100	0.17	0.52	24349
Russia*	85	8	13	81	1	1165	6439	1202444	71000	100	0.25	0.20	22581
Serbia	4	0	1	4	0	83	200	1747863	21895	100	0.22	0.34	14028
Singapore ^φ	1	0	0	1	0	1	1	5708840	710	4	0.38	0.36	67138
Slovakia	8	0	0	8	0	136	319	672452	6548	100	0.26	0.77	24924
Slovenia ^φ	1	0	0	1	0	1	1	2079980	20270	100	0.22	0.80	28235
South Korea	17	0	0	17	1	187	460	1891000	1846	0	0.00	0.83	36265
Spain	19	7	5	19	4	339	1746	1435680	10565	100	0.41	0.78	35696
Sri Lanka	9	2	3	5	0	101	309	1928655	7888	0	0.47	0.55	12697
Suriname ^φ	1	0	0	0	0	1	1	563400	163820	1	0.33	0.70	11815
Sweden	21	0	0	21	1	271	726	269280	11369	100	0.20	0.88	43376
Switzerland*	26	7	23	26	0	88	682	220533	873	100	0.54	0.86	64219
Taiwan	22	1	0	22	3	155	335	534870	1624	0	0.50	0.75	43211
Thailand	78	25	2	78	0	395	825	594187	5355	0	0.63	0.13	15740
Trinidad & Tobago	1	1	0	0	1	43	104	60874	300	40	0.13	0.73	39780
Turkey	81	16	3	81	0	415	978	479238	7626	0	0.22	0.24	27263
Ukraine	25	7	0	25	0	367	626	1331534	24587	100	0.47	0.32	9512
United Kingdom	4	0	2	12	1	194	534	3923740	17141	100	0.05	0.84	39128
United States*	53	2	15	0	3	1231	7772	4339367	137754	87.2	0.25	0.79	54795
Uruguay	19	0	0	0	0	173	500	84698	9529	60	0.08	0.83	19723
Venezuela*	25	1	0	0	0	301	657	715268	14800	20	0.07	0.17	7697
Vietnam	63	9	3	63	0	605	1320	1144988	4760	0	0.24	0.21	5846
Total/median	1767	345	198	1258	50	269	7772	732762	7388	25	0.22	0.52	173823

Note: ^φ states treated as regions; * federal countries.