

Ethnic Regional Autonomies Database (ERAD) Codebook

Version 2-2017

Compiled by Petr Panov

ERAD has been developed in the framework of the research project «Securing a balance in interethnic relations: regional autonomies, the state integrity and the rights of ethnic minorities». The project is supported by the Russian Science Foundation under Grant №15-18-00034 and carried out by the group of researchers of Perm State National Research University (Political Science Department) and Perm Federal Research Center of the Ural Branch of the Russian Academy of Sciences (Political Institutions and Processes Department)

Nadezhda Borisova, Lidia Borodina, Inna Vakina, Eleonora Minaeva, Maria Nazukina, Petr Panov, Oleg Podvinstev, Andrey Semenov, Konstantin Sulimov, Lyubov Fadeeva.

Content

The concept of ethnic regional autonomy (ERA).....	5
The list of ethnic regional autonomies.....	8
Ethnic Regional Autonomies Database (ERAD)	9
Units of observation	9
Structure of the ERAD	10
Missing data: coding rules.....	10
Projection of the data.....	11
1. General characteristics of both ERA and respective (host) country.....	12
Notes and comments	13
Post-colonialism: coding rules	13
Post-colonialism: coding rules	14
2. Ethnic composition of the population.....	16
Notes and comments	18
Dominant and titular ethnic groups: definition	18
Dominant ethnic groups: comments on some cases.....	19
Titular ethnic groups: comments on some cases	20
Linguistic and religious distance between the dominant and the titular ethnic groups	22

3. Autonomy in the context of the host country.....	25
Notes and comments	26
Economic indicators	26
Demographic indicators.....	30
4. Political characteristics of the host country	40
Notes and comments	42
Missing data.....	42
Ethno-regional political parties	43
Power-sharing in national government.....	44
Electoral statistics	45
5. Political characteristics of an autonomy	47
Notes and comments	48
Missing data.....	48
Ethno-regional political parties	49
Power-sharing in regional government	49
6. Preferential policies.....	51
Notes and comments	53
Index RAI: content and missing data	53
Linguistic policy	54
7. Conflicts.....	64
Notes and comments	65

Variables from ACD2EPR	65
Variables on the basis of Conflict Barometer Annual Reports	66
References	69
Appendix. The list of ethnic regional autonomies	71
Core list.....	71
Border-line list	75
Notes and comments	76

The concept of ethnic regional autonomy (ERA)

In the framework of the project, ethnic regional autonomy (ERA) is defined on the ground of essential characteristics (attributes) – territorial autonomy and ethnic character of a sub-national unit.

ERA is an administrative-territorial unit of the first sub-national level, which is based on an ethnic ground and has a sufficiently high degree of political self-government within the sovereignty of the national state.

Empirical criteria (necessary conditions) and operational indicators for defining of sub-national unit as ERA:

1. Autonomy

Territorial autonomy is assumed that sub-national units enjoy self-government (self-rule); at the same time, autonomous sub-national units are the parts of their host (principle) states, not independent entities. The scope, the content, and the form of territorial units' self-governance differ significantly from case to case. Therefore two criteria are introduced for the definition of territorial autonomy:

1.1. Region has political autonomy

Operational indicator of political autonomy - the presence in the regions its own legislature, and its own powers and jurisdiction

This distinguishes political autonomy from administrative autonomy, which is a result of delegation of executive / administrative authority to lower levels of government (Benedikter, 2009: 11; Wolff, 2010: 10).

1.2. Region is included in national politics and the host state is able and capable to control the sub-national units

Operational indicator – institutionalized representation of a region in the national legislature

That is why we do not consider as autonomies 'associated states' – numerous unincorporated organized territories of the USA, the British overseas territories, etc., which are independent in their domestic affairs (Lapidot, 1997: 54-55). Similarly, the so-called 'de-facto states' and regions of 'failed states' do not meet with this criteria even if these units are formally (legally) recognized as the parts of the respective state by the international community (like, for example, Abkhazia in Georgia).

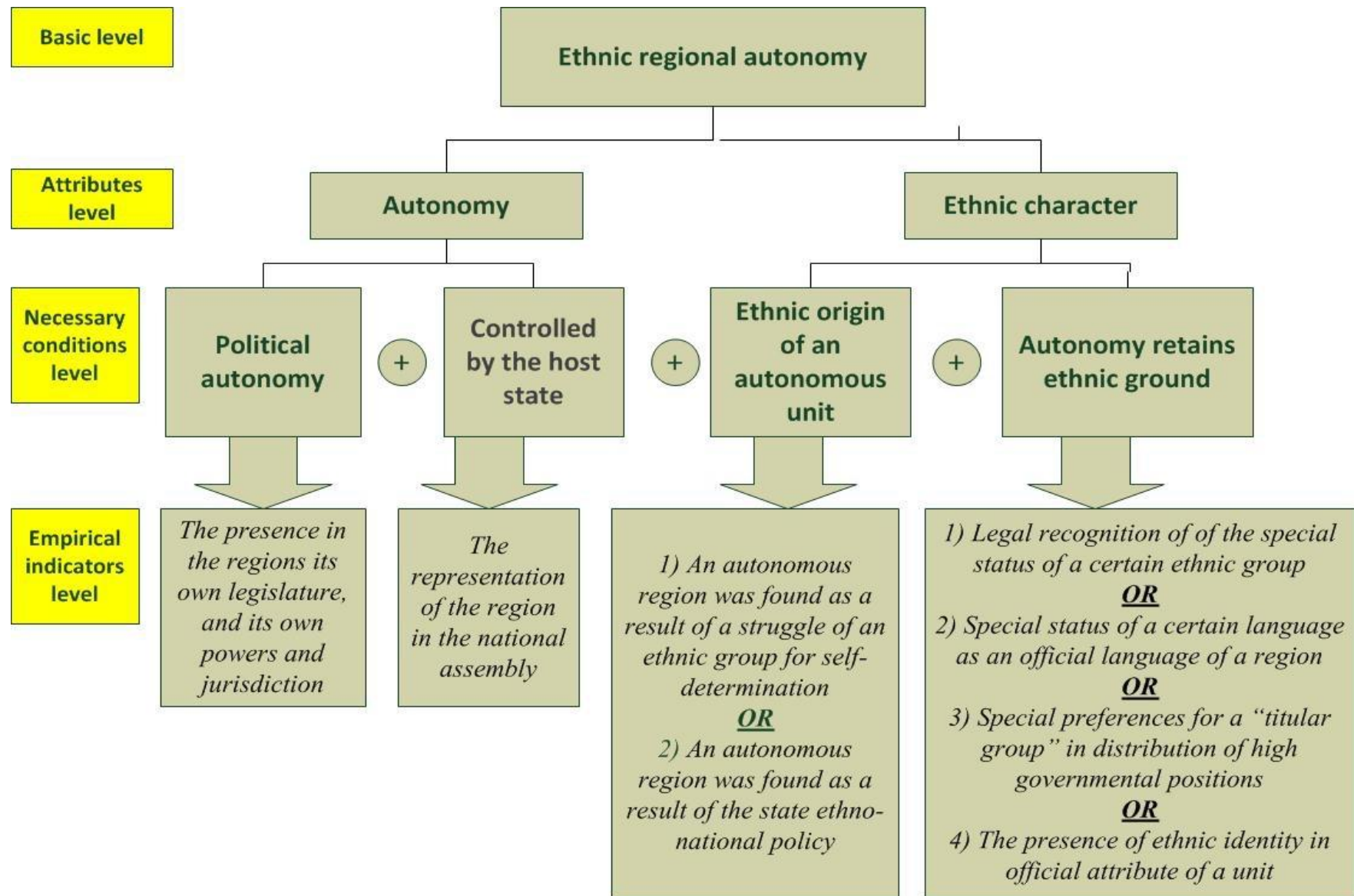


Figure 1. Conceptual structure of ‘ethnic regional autonomy’

2. Ethnic character of a sub-national unit

Ethnicity is understood in social constructivist way as a social categorization based on subjective belief in “common origin” or “common descent” of group’s members, regardless of what feature provide such a belief (Wimmer 2013: 7; Ghai 2000: 4; Horowitz 1985: 17-18). Two criteria are introduced for the definition of Ethnic character of a sub-national unit:

2.1. Ethnic origin of an autonomous unit: historically, ethnicity was a ground for granting autonomy to the region (Anderson, 2016: 18; Ghai and Woodman, 2013: 2).

Two empirical indicators have been developed in order to capture this criterion:

- *ERA is the result of the struggle of an ethnic group for self-determination. In other words, autonomous status represents a compromise between self-determination claims and maintenance of the state integrity;*
- *ERA is the result of the implementation of top-down “ethnonational policy”*

These two indicators may be complementary to each other in a supply-demand fashion that is self-determination claims (demand) meet respective top-down policy (supply) in the course of interactions between the central authorities and spatially concentrated ethnic group.

2.2. Ethnic character of autonomy remains up to date.

Empirically, current ethnic status of an autonomous region can be identified by one of the following indicators:

- *legal recognition of the special status of a certain ethnic group as a “titular ethnic group” in the region.*
- As a rule, “titular ethnic group” is recognized as “distinct nationality” (Ganguly and MacDuff, 2003: 3-4) while autonomous region is still perceived as “the homeland of the definite ethnic group” (Roeder, 2014: 93);
- *the presence of ethnic identity in official attributes of a unit (ethnonym in the title of a unit, ethnic symbols in flag, emblem, etc.);*
 - *special status of a certain language as an official language of a region;*
 - *special preferences for a “titular group” in distribution of high governmental positions.*

The term “**titular ethnic group**”, although not official, quite successfully grasps the ethnic character of the constitution of autonomy, especially in relation to another term - Staatsvolk, "state-forming" or "**dominant in the state**" **ethnic group** (O’Leary, 2001: 285).

The list of ethnic regional autonomies

Political autonomy of the subnational units can be found both in federal and unitary states, hence, it is useful to think of some unitary states as the “federacies” in Daniel Elazar’s terms where “a larger power and a smaller polity are linked asymmetrically in a federal relationship” (Elazar 1994: xvi). Consequently, empirical criteria of ERA are applied for unites of the first sub-national level with special status in unitary states and all units (entities) of the federations. In contrast to federacies, in federations, the principle of federalism is applied to all entities of the federations. Therefore, federation’s provinces have political autonomy by definition; all we need in these cases is to check for their ethnic character. In this regards a useful distinction can be made between “territorial” and “ethnic” federations, where “at least one constituent territorial governance unit is intentionally associated with a specific ethnic category” (Hale, 2004: 167-168).

In sum, the full list of ethnic regional autonomies (140 cases), which exist at present or ceased to exist in the beginning of XXI century, has been composed – *see Appendix*.

It has to be taken into account that ethnic boundaries and the very nature of ethnicity is a socially changeable and contentious phenomenon. Since ethnicity is a very strong sentiment, actors struggle over which social boundaries should be considered as ethnic lines, especially if ethnicity becomes politically salient. Consequently, though in most of cases one can observe any dominant and legitimate perception of ethnic categories and ethnic boundaries shared by the members of a society there are some cases without conventional perception of what social features are considered as ethnic lines. The second source of ambiguity is rooted in complexity of the process of the creation of some autonomies. Undoubtedly, granting autonomy is caused by the combination of the reasons; and in some cases our empirical indicators for the identification of ethnic origin of an autonomous unit – struggle of an ethnic group for self-determination and the implementation of top-down “ethno-national policy” – don’t allow making a clear decision. Therefore the list of ERA is divided in to two parts – “core list” and “border-line list”. While the ERAs from the “core list” fully accord to all the criteria of ethnic regional autonomies, 16 autonomies included in the “border-line list” are ambiguous cases.

Ethnic Regional Autonomies Database (ERAD)

The Ethnic Regional Autonomies Database (ERAD) contains a wide range of information on all ethnic regional autonomies around the world, which exist at present or ceased to exist in the beginning of XXI century. In contrast to other datasets (UCDP / PRIO Armed conflict datasets; Minorities at Risk - MAR; Ethnic Power Relations – EPR, etc.), the ERAD focuses on ethnic autonomous regions, not on ethnic groups or ethnic conflicts. The ERAD is a unique dataset in the sense that although it partially accumulates some relevant data from other datasets, it contains dozens of variables which were extracted from original sources. The data is available for 2001-2015 timespan and presented in autonomy-year format.

Units of observation

The list includes all ethnic regional autonomies around from the list of ERAs (except Telangana that was founded only in 2014) – 139 autonomies from 34 countries. Some of them exists only part of the time interval of 2001 - 2015 (marked in the list of ERAs). Rodrigues was granted autonomy in 2002, Aceh – in 2005, autonomy of Kurdistan was reestablished in 2005. South Sudan existed as autonomous region in of Sudan in 2005-2011 and then became an independent state. On the other hand, 6 autonomous okrugs of the Russian Federation were eliminated during administrative reforms in 2005-2008. In Crimea case, there was an irredentist secession in 2014.

In order to make it easier to use the ERAD⁶ for those ERAs that emerged after 2001 or ceased to exist until 2015, 72 "missed years" are left in the ERAD, but are highlighted by a dark color. Thus, there are 2085 rows in the database, but 2013 units of observations.

Additionally, there were some reorganizations of ERAs in 2001-2015. Until 2010, Curaçao, Bonaire, Sint Maarten, Sint Eustatius and Saba constituted The Netherlands Antilles as one of the three self-governing entities of Kingdom of the Netherlands, so that they are presented in the list as a single ERA in spite of that in 2010 the reorganization took place, and only two islands (Curaçao and Sint Maarten) get the status of self-governing entities. Since Curaçao is the greatest of the Netherlands Antilles, all the data in the ERAD took for Curaçao.

The second reorganization is Papua in Indonesia: in 2003, it was decided to divide it into two provinces with a special status, but this process is still incomplete, and Papua is presented in the list as a single region.

Structure of the ERAD

The ERAD consists of 148 variables accumulating the data on wide range of aspects of ethnic regional autonomies. The database is divided into 7 sections:

- 1) General characteristics of both ERA and respective (host) country;
- 2) Ethnic composition of the population;
- 3) ERA in the background of the country: economic and demographic statistics;
- 4) Political characteristics of the country;
- 5) Political characteristics of the ERA;
- 6) Preferential policies;
- 7) Ethno-political conflicts in ERA

Below, for each section, tables are presented, which describe the variables: list, type, data sources, coding rules, notes. In those cases (marked with a *), when additional information is required to characterize the variable, justify encoding rules for the change, and present data sources, it is given in the comments to the section.

Missing data: coding rules

The data are coded as "not applicable" and respective cells are marked with "NA" in the following cases:

- 1) This indicator is not applicable to the unit of observation, for example, the variable "year of the termination of an autonomy" can take value only if the ERA was terminated, in other cases - "not applicable"; the variables characterizing the upper house of parliament can take values only cases where the upper chamber exists, in others cases - "not applicable", etc.
- 2) The data is taken from already existing databases, where they are presented in a systematic form, but there are no data for individual units in this database.

The lists of cases, where the data are coded as "not applicable", are presented in the comments to each section. Some exceptions to this rule with their justification are also presented in the comments.

Another type of the missing data is the cases, where no information was available. In these cases, the respective cells were left blank. In the comments to individual sections, a list of these cases is given.

Projection of the data

The specificity of some data is that they do not appear annually, but at regular intervals, for example, the results of censuses, the results of parliamentary elections. Since the unit of observation in ERAD is "autonomy - year", in these cases, the data was projected "forward", that is, prospectively, for the next period of time (before the appearance of new data). In some cases, when earlier data are not available, such a projection was made "backward", that is, retrospectively, for the previous period of time. All the cases of projection of the data are presented in the comments to individual sections.

1. General characteristics of both ERA and respective (host) country

Description of the variable	Label	Type	Sources	Comments
1. General characteristics				
Autonomy name	regions_name	String	Original list of ERAs	
Year of observation	year	String		
Country name	country_name	String		
Region	region_name	String and some binominal		
1.1. General characteristics of autonomy				
Autonomy area (km ²)	area_aut	Interval	Statoids project (http://www.statoids.com/ldf.html)	
Autonomy population (thousands of people)	pop_aut	Interval	National censuses	From the census the data are projected forward and backward
Island territory	island	Binominal	Geographic maps	(0) no (1) yes
Post-colonial territory*	postcolon	Binominal		(0) no (1) yes
Kin-state*	kin_state	Binominal		(0) no (1) yes
Year of the creation of autonomy in its present form*	start_year	Interval		
Year of the creation of an autonomy (historically)*	start_year_1	Interval		
Legal act that establishes an autonomy	legal_act_estab	String		
Legal act that regulate an autonomy at present	legal_act_current	String		
Year of the termination of an autonomy	end_year	String		
The gap in the existence of autonomy	discontinuity	Binominal		(0) no (1) yes

Years of the gap in the existence of autonomy	discont_years	String		
Number of years of the gap in the existence of autonomy	discont_period	Interval		
1.2. General characteristics of country				
Country area (km ²)	area_country	Interval	World Development Indicators http://data.worldbank.org/indicator	Data on 23.04.2017
Country population (thousands of people)	pop_country	Interval	World Development Indicators http://data.worldbank.org/indicator	Data on 23.04.2017
GDP per capita (current \$)	gdp_pc_country	Interval	World Development Indicators http://data.worldbank.org/indicator	Data on 23.04.2017
Share of urban population	urban_country	Interval	World Development Indicators http://data.worldbank.org/indicator	Data on 23.04.2017
Gini inequality Index	Gini_country	Interval	World Development Indicators http://data.worldbank.org/indicator	Data on 23.04.2017 (if the data absent, the data are projected forward and backward)
The spread of the Internet	internet_country	Interval	World Development Indicators http://data.worldbank.org/indicator	Data on 23.04.2017
Human development index	HDI_country	Interval	United Nations Development Program http://hdr.undp.org/en/data	Data on 01.05.2017
Globalization index	glob_country	Interval	KOF Index of Globalization, ETH Zurich http://globalization.kof.ethz.ch/	

Notes and comments

Post-colonialism: coding rules

This variable was coded as "1" if the territory of the ERA was the periphery of the colonial empire. After the collapse of empires, there are probably four routes. The first, when a periphery received independence, does not concern ERA. The other three routes led to the creation of an ERA:

- 1) The periphery remained in the former metropolitan state, receiving an autonomous status (Netherlands Antilles, French overseas territories, Russian ERAs (Tuva is a “postponed case”), etc.
- 2) Ethnically specific parts of the periphery, which were the imperial provinces, get autonomy as part of a new independent state, and there is a good reason to consider them as ERAs. Postcolonial India is an exemplary case. Originally it inherited numerous features of the British colonial administrative system; and some states were created on the basis of ethnic colonial provinces (West Bengal, Jammu and Kashmir, Orissa). At the same time, it is possible that a unit was initially constituted on non-ethnic grounds but transformed into an ERA later. Thus, a number of other Indian states (Assam, Tamil Nadu, Karnataka, Punjab) were constituted on the basis of non-ethnic colonial provinces, however in the 1950s, the government started the reform of the administrative-political composition upon the linguistic principle (The States Reorganisation Act, 1956). In course of this reform, non-ethnic units were transformed into ethnic autonomies based on definite titular ethno-linguistic groups.
- 3) Ethnically specific imperial province gaining independence jointed the other former imperial province on the basis of autonomous arrangements (often with the active mediation of the former metropolitan country) (Zanzibar, Sarawak and Sabah).

Kin-state: coding rules

Kin-state is assumed as the state a dominant ethnic group of which is the same as a titular ethnic group of ERA. This variable was coded as "1" if ERA has such a kin-state (15 cases out of 139).

Year of the creation of autonomy: coding rules

In some cases, the dating of the creation of an ERA can be interpreted ambiguously. Therefore ERAD contains two variables. The variable "start_year" indicates the year of the creation of the ERA in its current form, while the variable "start_year_1" - in historical perspective. In 47 ERAs, the values of these variables do not coincide for the following reasons:

- 1) in 12 cases, there was a gap in the existence of autonomy, as it was eliminated in a definite time and then reestablished. Therefore, "start_year_1" indicates when the ERA was created for the first time, and "start_year" - the year of its reestablishment after the gap.
- 2) in some cases, the state, in which the territory had an autonomous status, collapsed, and the new state was created in its place, within which this territory retained an autonomous status (post-Soviet countries - Russia, Georgia, Uzbekistan, Tajikistan). The "start_year_1" indicates when the ERA was established within the former state, and "start_year" - the year of its "confirmation" within the "new state".

3) in a few cases, a territory had a status similar to autonomy ("proto-autonomy") within the framework of imperial / post-imperial polity, but later it was substantially transformed into modern autonomy (for example, Scotland, Wales, Quebec). The "start_year_1" indicates when the ERA appeared as a "proto-autonomous entity", and "start_year" - the year of the creation of this ERA in the modern form.

2. Ethnic composition of the population

Description of the variable	Label	Type	Sources	Comments
2. Ethnic composition of the population				
2.1. Ethnic composition of the mother country				
Fractionalization indices <ul style="list-style-type: none"> ethnic language religion 	Ethnicfrac Langfrac Religfrac	Interval	Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." Journal of Economic Growth 8: 155-94.	
Polarization index Fragmentation index	RQETHPOL RQETHFRAG	Interval	Montalvo, José G., and Marta Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict, and Civil Wars." American Economic Review, 95(3): 796-816.	
Segregation index <ul style="list-style-type: none"> ethnic language religion 	Ethnicseg Langseg Religseg	Interval	Alesina, Alberto, and Ekaterina Zhuravskaya. 2011. "Segregation and the Quality of Government in a Cross Section of Countries." American Economic Review, 101(5): 1872-1911.	The variant with taking into account of 'other groups' is used
Spatial inequality index	spatial_inequality	Interval	Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. 2016. "Ethnic Inequality." Journal of Political Economy 124 (2): 428-488.	The variable 'ad1_a_lum00pc', which indicates luminosity per capita in 2000, 2006 and 2012 based on first-level administrative units indicates, is used
Ethnic inequality index-1	ethnic_inequality_1	Interval	Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. 2016. "Ethnic Inequality." Journal of Political Economy 124 (2): 428-488.	The variable 'v2g_a_lum00pc', which indicates luminosity per capita in 2000, 2006 and 2012 based on perturbed ethnic homelands (Thiessen polygons) as depicted by GREG, is used
Ethnic inequality index-2	ethnic_inequality	Interval	Alesina, Alberto, Stelios	The variable 'v2g_a_lum00pc', which

	y_2		Michalopoulos, and Elias Papaioannou. 2016. "Ethnic Inequality." Journal of Political Economy 124 (2): 428-488.	indicates luminosity per capita in 2000, 2006 and 2012 based on perturbed ethnic homelands (Thiessen polygons) as depicted by Ethnologue, is used
Number of politically relevant ethnic groups (Cederman)	rlvt_groups_count	Interval	EPR project	
Number of ethnic groups in power (Cederman)	egip_groups_count	Interval	EPR project	
Index nstar (Cederman)	nstar	Interval	GROWup Cederman, L., and L. Girardin. 2007. Beyond fractionalization: Mapping ethnicity onto nationalist insurgencies. American Political Science Review 101(1):173–85.	
2.2. Dominant ethnic group in the country and titular ethnic group in the autonomy				
Dominant ethnic group in the country*	dom_group_name	String		
Number of the dominant ethnic groups in the country	dom_group_num	Interval		
Titular ethnic group in the autonomy*	titul_group_name	String		
Number of titular ethnic groups in the autonomy	titul_group_num	Interval		
Ethnonym in the name of an autonomy	ethnonim	Ordinal		(0) no (1) the name of autonomy contents the name of ethnic group but the name of ethnic group is based on toponym, not original (Corsicans, Sicilians, etc.) (2) yes
Language of a dominant ethnic group	lang_dom	String	GROWup https://growup.ethz.ch Ethnolog https://www.ethnologue.com/	
Language of a titular ethnic	lang_tit	String	GROWup https://growup.ethz.ch	

group			Ethnolog https://www.ethnologue.com/	
Linguistic distance between dominant and titular groups	dist_lang	Ordinal	Calculated on the basis of GROWup и Ethnolog	(3) languages belong to different linguistic families, (2) languages belong to the same family but different linguistic groups, (1) languages belong to the same group; (0) no language difference
Religion of a dominant ethnic group	religion_dom	String	GROWup https://growup.ethz.ch and other sources	
Religion of a titular ethnic group	religion_tit	String	GROWup https://growup.ethz.ch and other sources	
Religious distance between dominant and titular groups	dist_religion	Ordinal	Calculated on the basis of GROWup	(2) religions belong to different classes; (1) religions belong to the same classes but different sub-classes; (0) religions belong to the same sub-classes or no religious difference

Notes and comments

Dominant and titular ethnic groups: definition

Dominant ethnic group in the country is a “Staatsvolk” that is “a single national (or ethnic) people that dominates a state demographically and electorally” (O’Leary 2001: 285). As a rule, it is that ethnic group whose cultural characteristics became the main ones in the process of homogenization of the political and cultural space in the formation of a civil nation in a given state. The empirical indicator of this is, first and foremost, the official language of the state.

Titular ethnic group in ERA is an ethnic group that was the ground of the establishment of ERA.

In most of cases the distinguishing of both dominant and titular groups is fairly clear, but in some cases comments are required.

Dominant ethnic groups: comments on some cases

Iraq	In Iraq, not only language but also religion delimits ethnic boundaries, but since the Kurds as a titular group differ from Shiite Arabs and Sunni Arabs by language, it is the Arabs who are generally identified as the dominant group in the country (Shiite Arabs and Sunni Arabs are summed up when calculating the corresponding variables).
Indonesia	Due to the complexity of the ethnic composition of the country, different way to define the dominant group is possible. Here we follow the most conventional way and consider Javanese (the main population of the main island of Java; distinguished by language) as the dominant ethnic group.
Papua New Guinea	Many Papuan tribes live in the country, but since the "Bougainvilleans" are defined as the title group in Bougainville (see below), there is a good reason to consider all Papuan tribes, i.e. the entire population of Papua New Guinea, except Bougainville, as the dominant ethnic group.
Philippines	Due to the complexity of the ethnic composition of the country, different way to define the dominant group is possible. Here we follow the most conventional way and consider Tagalog people (the main population of the central part of the country; distinguished by language) as the dominant ethnic group.
Switzerland	German Swiss are considered as the dominant ethnic group.
Tanzania	Due to the complexity of the ethnic composition of the country, different way to define the dominant group is possible, but since the "Zanzibaris" are defined as the title group in Zanzibar (see below), there is a good reason to consider all residents of mainland ("Mainland Africans") as the dominant ethnic group ¹ .

In some cases, first of all in full ethno-federations, there are some dominant ethnic groups, and it is possible that a dominant group is also a titular group in its "own" entity of the federation.

Belgium	Two dominant ethnic groups: Flemings and Walloons. Each of them is a titular group in own entity of the federation.
Bosnia and Herzegovina	Three dominant ethnic groups: Bosniaks, Croats, and Serbs. Serbs is a titular group in one entity of the federation (Republika Srpska) while Bosniaks and Croats are titular ethnic groups in the other entity of

¹ Such interpretation is used, for instance, in Ethnic Power Relations Project (<https://icr.ethz.ch/data/epr/>).

	the federation (Federation of Bosnia and Herzegovina).
Ethiopia	Two dominant ethnic groups: Amhara and Tigrayans, since they are commonly recognized as “Staatsvolk peoples” in the historical Ethiopian state. Both of them are titular ethnic groups in their “own” entity of the federation.
Fiji	Two dominant ethnic groups: Fijians indigenous and Indo-Fijians.
Nigeria	Three ethnic groups are recognized as “Staatsvolk people” (Hausa, Igbo, and Yoruba). The states, where these groups are titular ones, are not considered as ERAs.
Pakistan	As it was Punjab that was a “core” in the formation of Pakistani state, there is a good reason to consider Punjabis as a dominant ethnic group. At the same time, Punjabis is a titular ethnic group in Punjab (one of autonomous provinces of Pakistan).

Titular ethnic groups: comments on some cases

Tripura (India)	Bengalis, who massively migrated here after 1947 and clearly dominate in politics of this state at present, are considered as the titular group. Moreover, the status of the state was given to Tripura only in 1971. In this context, Indigenous people Kokborok don’t accord with the concept of “titular ethnic group”.
Sabah and Sarawak (Malaysia)	Indigenous “bumiputera” peoples (Kadazans and Dayaks respectively) are considered as titular ethnic groups.
Jura (Switzerland)	French speaking Jurassians, not French speaking Swiss are considered as a titular ethnic group.

4 cases: there are no clear that a titular group is an ethnic group, as the strength of the ethnic grounds for the self-identification there remains unclear. These cases are included in “border-line list of ERAs”.

Bougainville (Papua New Guinea)	The processes of ethnogenesis in Bougainville are far from over; about two dozen language groups stand out among the island's population. Nevertheless, in the struggle for self-determination, common Bougainville identity has been constructed. Sometimes, it is interpreted as ethnic, and Bougainvilleans are slightly different from the Papuans (residents of the main part of the country) in a racial sense (darker skin).
Zanzibar	Unlike the mainland, the history of Zanzibar is closely related to the Arabs, who in the Middle Ages made the

(Tanzania)	island the center of the slave trade. The centuries-old contacts between the Arabs and the indigenous population led to the Islamization of the island's population, as well as the emergence of Shirazi as a special group of people: having African appearance, they do not have “their own language” but speak in Swahili (the official language of Tanzania) and profess Islam (99% of the population of Zanzibar are Muslims while only 35% - in Mainland).
Sicily (Italy)	Sicilians: language differences are often considered as "internal" dialect of the Italian language, but the Sicilian identity is sufficiently developed.
Ajara (Georgia)	Historically, Ajarans were under the rule of the Ottoman Empire and accepted Islam. After that Adjara was incorporated into Georgia, religious differences, which can be considered ethnic, became the basis for granting Ajaria autonomy. Nevertheless, over time the significance of religious differences was significantly reduced.

In 8 cases, there are some titular ethnic groups

Federation of Bosnia and Herzegovina	Bosniaks and Croats
Benishangul-Gumuz (Ethiopia)	Berta and Gumuz
Meghalaya (India)	Khasi, Garo, and Jaintia
Benue (Nigeria)	Tiv and Idoma
Kabardino-Balkariya (Russia)	Kabardians and Balkars
Karachaevo-Cherkessiya (Russia)	Karachays and Circassians
Khanty-Mansi AD (Russia)	Khanty and Mansi
Taymyr AD (Russia)	Dolgans and Nenets

In 4 cases, “the group of peoples” is usually considered in the literature as a titular ethnic group. Therefore the variable «titul_group_num» is coded as «group of peoples»².

Gambela (Ethiopia)	Nilotic peoples (Nuer people Anuak people are the most significant in their size).
--------------------	--

² Due to internal heterogeneity of ethnic group, in some cases it is also very hard to define whether the group should be considered as a “single ethnic group” or a “group of peoples” (Moro people in Philippines, Bouganvilleans in Papua New Guinea, Naga people in India, etc.). However, in the ERAD we follow conventional view on them as a “single ethnic group”.

Southern Nations, Nationalities and Peoples Region (Ethiopia)	Southern Peoples (more than 40 ethnic groups).
South Sudan (Sudan)	Southern Peoples (Dinka people and Nuer people are the most significant in their size).
Dagestan (Russia)	Informally, 14 peoples are considered as the “main peoples” Dagestan (http://valerytishkov.ru/cntnt/publikacii3/kollektivn/na_puti_k/respublika2.html). Though 3 of them are Russians, Azeri, and Chechens, who have their own states or ERA. Therefore there is a good reason to consider the group from 11 peoples as a “titular group”. Avars, Dargins, Lezgins, and Kumyks people are the most significant in their size.

In 5 cases, it is quite hard to define ethnic group. These cases are included in “border-line list of ERAs”. In ERAD the variable «titul_group_num» is coded as «NA”.

Arunachal Pradesh (India)	No definite titular group, the largest groups are Nyishi, Adi
Cross River (Nigeria)	No definite titular group, the largest groups are Efik, Ejagham
Nasarawa (Nigeria)	No definite titular group, the largest group is Eggon people
Plateau (Nigeria)	No definite titular group, the largest group is Berom people
Taraba (Nigeria)	No definite titular group, the largest groups are Mumuyes, Jukun, Tiv

Linguistic and religious distance between the dominant and the titular ethnic groups

To determine the linguistic distance between the dominant and the titular ethnic groups, the classification of languages into families and groups from the World Languages Register (<http://lingvarium.org/start.shtml>) and information from GROWup (<https://growup.ethz.ch>) and Ethnologue (<https://www.ethnologue.com/>) is used.

To determine the religious distance between the dominant and the titular ethnic groups, the information from GROWup was transformed into a scale similar to the classification of languages: all the religious were divided into some classes and then some of classes were divided into sub-classes (see Table).

Classes	Sub-classes
ARC Christianity	ARC1 Roman Catholic <u>RomCath</u> ARC2 Eastern Orthodox <u>EastOrth</u> ; <u>GeorgOrth</u> ; <u>RusOrth</u> ; <u>UkrOrth</u> ARC3 Oriental Orthodox <u>EthiopOrth</u> ARC4 Protestantism <u>Prot</u> ; <u>Presbiterianism</u> ; <u>Baptism</u> ARC5 Anglican <u>Anglican</u>
ARI Islam	ARI 1 <u>Sunni</u> ARI 2 <u>Shia</u>
ERH Hinduism	
ERB Buddhism	
ERS Sikhism	
EARS Chinese	
ANI Animism	
ARJ Judaism	

Coding rules in special cases:

- 1) when a group has several different religions / languages (in the ERAD they are indicated through the slash), the maximum distance is taken for the coding
- 2) when a dominant group is one, and there are several titular groups, which have different religions / languages (in the ERAD they are indicated through the slash), the maximum distance is taken for the coding
- 3) when a title group is one, and there are several dominant groups, which have different religions / languages (in the ERAD they are indicated through the slash), the maximum distance is taken for the coding
- 4) when a dominant group / group in the country is simultaneously a titular group in the ERA, the rules for these cases are as follows:

Belgium	Two dominant ethnic groups: Flemings and Walloons. Each of them is a titular group in own entity of the federation. Therefore, they are compared to each other.
Bosnia and Herzegovina	Three dominant ethnic groups: Bosniaks, Croats, and Serbs. Serbs is a titular group in one entity of the federation (Republika Srpska) while Bosniaks and Croats are titular ethnic groups in the other entity of the federation (Federation of Bosnia and Herzegovina). Therefore, Serbs are compared to Bosniaks and Croats while Bosniaks and Croats as titular ethnic groups in FBG – to Serbs..
Ethiopia	Two dominant ethnic groups: Amhara and Tigrayans, and both are titular ethnic groups in their “own” entity of the federation. However, there are 7 other ethnic regions (ERAs). Therefore, titular groups of these 7 ERAs are compared to Amhara and Tigrayans while Amhara and Tigrayans as titular egroups in their own regions are compared to each other.
Pakistan	Punjabis are a dominant group in the county and a titular group in Punjab. Therefore, the titular groups of all three other ERAs are compared to the Punjabis, and the Punjabis a titular group in Punjab - to the others.

3. Autonomy in the context of the host country

Description of the variable	Label	Type	Sources	Comments
3. Autonomy in the context of the host country				
Share of autonomy in the area of the country	share_area_aut	Interval	area_aut / area_country	
Share of autonomy in economy of the country*	share_econ_aut	Interval	grp_aut / gdp_country Calculated on the basis of the data of national statistic agencies	
The level of economic development of autonomy in the country's context*	level_econ_aut	Interval	grp_pc_aut / gdp_pc_country Calculated on the basis of the data of national statistic agencies	
Share of autonomy in the population of the country*	share_pop_aut	Interval	pop_aut / pop_country Calculated on the basis of the data of national censuses	
Share of dominant group in population of the country*	share_dom_pop_country	Interval	Calculated on the basis of the data of national censuses	
Share of titular group in population of the country*	share_titul_pop_country	Interval	Calculated on the basis of the data of national censuses	
Share of dominant group in population of the autonomy*	share_dom_pop_aut	Interval	Calculated on the basis of the data of national censuses	
Share of titular group in population of the autonomy*	share_titul_pop_aut	Interval	Calculated on the basis of the data of national censuses	
Concentration of a titular group in population of the autonomy*	conc_titul_pop_aut	Interval	titul_pop in autonomy / titul_pop in country Calculated on the basis of the data of national censuses	

Notes and comments

Economic indicators

When calculating the share of ERAs in the country's economy and the level of economic development of ERAs, compared to the level of development of the country as a whole, 4 indicators were used:

- 1) GDP of the country
- 2) GDP per capita in the country
- 3) GRP of autonomy
- 4) GRP per capita in the autonomy

For 12 countries and their ERAs, the data are taken from the OECD database, the only which contains relevant information on both countries and regions of the first sub-national level³.

1. France (only for Corsica)
2. Spain
3. Belgium
4. Canada
5. Italy
6. Switzerland
7. Great Britain
8. Finland
9. China
10. India
11. Russia
12. Indonesia⁴

³ In the database there are also Denmark, France and the Netherlands, but their island autonomies are not included in this database.

⁴ Since the data are presented for the two states of Papua separately, their regional GRP is summed up, and the value of GRP per capita is taken as average.

Data extracted on 02 May 2017 02:11 UTC (GMT) from OECD.Stat
 GDP and Regional GDP, Millions USD, constant prices, constant PPP, base year 2010
 GDP per head and Regional GDP per head, USD, constant prices, constant PPP, base year 2010

Based on these data, the share of ERAs in the country's economy and the level of economic development of ERAs, compared to the level of development of the country as a whole, were calculated.

Since the original data in the OECD database for some countries is not available for each year of the period of 2001-2015, the missing values were projected prospectively and retrospectively.

For all other cases, the data were taken from national statistical services; in some cases, where there are no available official statistics, expert evaluations were used (see below). All these data are not comparable to each other, because are available for different years, are given in different currencies and using different calculation methods. However, they were selected in such a way that the national and regional levels were comparable. This made it possible to calculate the share of ERAs in the country's economy and the level of economic development of ERAs compared to the level of development of the country.

Since these data are not available for each year of the period of 2001-2015, the missing values were projected prospectively and retrospectively.

1. Bosnia_Herzegovina	National statistical service http://www.bhas.ba/tematskibiltenti/GDP_2005-2013_001_01_eng.pdf GDP in millions USD, current prices; GDP per capita in USD, current prices Values on the separate entities are in the same measurement: Federation of Bosnia and Herzegovina – http://www.fzs.ba/SG2015.pdf ; Republika Srpska – http://www.rzs.rs.ba/front/article/1778/
2. Denmark	World Development Indicators (data on Greenland, Faroe Islands, and Denmark are given separately): GDP and GRP in millions USD, current prices; GDP per capita and GRP per capita in USD, current prices
3. Ethiopia	National statistical service http://www.csa.gov.et/ GDP and GDP per capita in 2005 – 2009 in Ethiopian Birrs
4. Fiji	National statistical service www.statsfiji.gov.fj GDP and GDP per capita in 2011 – 2015 in Fijian Dollars, Current Market Price
5. France (for New Caledonia and French	United Nations Statistics Division https://unstats.un.org (data on New Caledonia, French Polynesia and France are given separately): GDP and GRP in millions USD, current prices; GDP per capita

Polynesia)	and GRP per capita in USD, current prices
6. Georgia	National statistical service http://www.geostat.ge GDP at current prices, mil. GEL и GDP per capita (at current prices), GEL Georgia's Regions Gross Value Added (current prices, mil. GEL) http://www.geostat.ge/index.php?action=page&p_id=119&lang=eng
7. Iraq	National statistical service http://cosit.gov.iq/en/home GDP at basic current prices (Billion US\$) и GDP per capita at current prices (US\$) in 2010 – 2013 Kurdistan – expert evaluations from «Calculating the Gross Regional Product of the Kurdistan Region – Iraq» https://www.rand.org/pubs/research_reports/RR1405.html and from official web-site of The Board of Investment http://www.kurdistaninvestment.org/economy.html
8. Malaysia	National statistical service https://www.statistics.gov.my/ GDP and GRP – constant prices 2005 in national currency (millions RM); GDP per capita and GRP per capita - current prices 2005 in national currency (RM)
9. Mauritius	National statistical service http://statsmauritius.govmu.org GDP at market prices in mln rupees and GDP per capita at market prices in rupees
10.Moldova	National statistical service http://statbank.statistica.md GDP and GRP in current prices, thousand Lei
11.Myanmar	National statistical service http://www.myanmar.cm/myanmardata2007/14.htm GDP and GDP per capita in kyats
12.Netherlands	United Nations Statistics Division https://unstats.un.org (data on Aruba, Curacao and Netherlands are given separately): GDP and GRP in millions USD, current prices; GDP per capita and GRP per capita in USD, current prices
13.Nicaragua	The only source that contains comparable data on Nikaragua and its ERAs is: http://www.unfpa.org.ni/wp-content/uploads/2014/09/indh_costa_caribe_nic_2005.pdf GDP and GRP in 2000, en millones de cordobas
14.Nigeria	C-GIDD (Canback Global Income Distribution Database). Archived from the original on 11 March 2012. Retrieved 20 August 2008. GDP and GRP 2007 in millions USD (PPP); GDP per capita and GRP per capita 2007 in USD (PPP)
15.Pakistan	The only source that contains comparable data on Pakistan and its ERAs is “REGIONAL

	ACCOUNTS OF PAKISTAN: METHODOLOGY AND ESTIMATES (1973 – 2000)” http://www.spdc.org.pk/Data/Publication/PDF/WP5.pdf . GDP, GRP, GDP pc, GRP pc in Pakistani rupees.
16.Papua New Guinea	National statistical service http://www.nso.gov.pg/ GDP in local currency (K, in millions) and GDP per capita in local currency (K)
17.Philippines	National statistical service http://countrystat.psa.gov.ph/selection.asp GDP and GRP 1987-2008 at 1985 prices; 2009-2014 at 2000 prices, millions pesos http://psa.gov.ph/regional-accounts/grdp/data-and-charts GDP per capita (1998-2014) and GRP per capita (2010-2014), in prices constant (2000), pesos
18.Sudan	AFRICAN DEVELOPMENT BANK GROUP http://www.afdb.org/fileadmin/uploads/afdb/Documents/Project-and-Operations/SUDAN%20-%20Country%20Brief%202010-2011.pdf GDP and GDP per capita, at Current Prices in dollars South Sudan statistical service http://static1.1.sqspcdn.com/static/f/750842/25421027/1410241500537/The+GDP+Final+2013.pdf?token=JCSP%2FdOxbjcbpdcS2eVfR6UuYX8%3D GRP and GRP per capita at Current Prices in dollars
19.Tajikistan	National statistical service http://www.stat.tj/ru/analytical-tables/real-sector/ GDP and GRP in millions Somonies
20.Tanzania	National statistical service http://www.nbs.go.tz/ GDP at current market prices in millions of Shillings; GDP per capita at current market prices in Thousands of Shillings. National statistical service of Zanzibar http://www.ocgs.go.tz/ GRP at current market prices in millions of Shillings; GRP per capita at current market prices in Thousands of Shillings.
21.Ukraine	National statistical service http://www.ukrstat.gov.ua/ GDP and GRP current prices in millions of hryvnia; GDP per capita and GRP per capita current prices in millions of hryvnia
22.Uzbekistan	National statistical service http://www.stat.uz/ru/dinamicheskie-ryady GDP (2012-2015) in billion so‘m Karakalpakstan http://www.stat.uz/ru/141-otkrytye-dannye/1103-respubliki-karakalpakstan GRP (2012-2015) in billion so‘m

For some autonomies, data are absent due to the lack of statistics (the abolished autonomous regions of the Russian Federation were not taken into account in national statistics separately from the "mother" subjects of the Russian Federation) - coded as "NA".

For some autonomies, economic statistics have not be found; the respective cells were left blank.

- All Ethiopian ERAs
- Rotuma
- Ajara (GRP per capita)
- Gagauzia (GRP per capita)
- Rodrigues
- All the ERAs of Myanmar
- Both ERAs of Nikaragua (GRP per capita)
- All the ERAs of Panama
- Bougainville
- Gorno-Badakhshan (GRP per capita)
- Karakalpakstan (GRP per capita)

Demographic indicators

Due to the fact that there are no systematic databases on the population of the sub-national units, especially in terms of ethnic structure, all the data are taken from national censuses (except some cases, see below). To calculate the share of ERA in the population of the country ("share_pop_aut"), comparable data on the population of the country and the ERA are needed. Therefore the data on the population of the country are taken from the same censuses, not from World Development Indicators as it was for the variable "pop_country" in the section "1.2. General characteristics of the country".

Coding rules in special cases:

- 1) If there are some dominant and / or titular groups, their shares in the population of the country / autonomy are summarized
- 2) If a dominant group / group in the country is simultaneously a titular in the ERA, the rules for these cases are as follows:

Belgium	Two dominant ethnic groups: Flemings and Walloons. Each of them is a titular group in own entity of the federation. Therefore, when calculating the share of the dominant group in the population of the country, Flemings and Walloons are summarized, while in calculating the share of the dominant group in the ERA's population, only the opposite group is taken: Walloons for Flanders, Flemings for Wallonia.
Bosnia and Herzegovina	Three dominant ethnic groups: Bosniaks, Croats, and Serbs. Serbs is a titular group in one entity of the federation (Republika Srpska) while Bosniaks and Croats are titular ethnic groups in the other entity of the federation (Federation of Bosnia and Herzegovina). Therefore, when calculating the share of the dominant group in the population of the country, all three groups (Serbs, Croats and Bosniaks) are summarized, while in calculating the share of the dominant group in the ERA's population, only the opposite group (group) is taken: for FBG - Serbs, for RS - the sum of Croats and Bosniaks.
Ethiopia	Two dominant ethnic groups: Amhara and Tigrayans, and both are titular ethnic groups in their "own" entity of the federation. However, there are 7 other ethnic regions (ERAs). Therefore, when calculating the share of the dominant group in the population of the country, as well as calculating the share of the dominant group in the population of all ERAs (including Amhara and Tigray), Amhara and Tigrayans are summarized. That is why in two ERAs (Amhara and Tigray), the ratio between the dominant and titular groups is not "exclusive", and the titular group is part of the dominant one.
Pakistan	Punjabis are a dominant group in the county and a titular group in Punjab. Therefore, with respect to Punjab, the shares of the dominant and titular groups in the population of the country coincide. Similarly, the share of the dominant group in the Punjab and the titular group in the Punjab are the same.

As noted above, in 5 cases the title group cannot be precisely defined. Therefore, in Arunachal Pradesh (India), in the calculation of the share of titular groups, the data for the two largest language groups (nissi / dafla = 18.97% + adi = 16.91%) are summarized. In the remaining cases (four Nigerian states), no assumptions were required as there are no statistics on the ethnic structure of the population in Nigeria.

Since the censuses are not conducted every year, the values of all demographic variables are projected prospectively and retrospectively for the entire period of 2001 - 2015. If in the period of 2001 - 2015 there were several censuses, the data were

projected only prospectively. In a few cases (China, India, the Netherlands, the Philippines), one census contains questions on ethnicity, and the other does not, in such cases, only the census that contains the necessary data is used in order to make the data comparable and calculate the values of the variables correctly.

It has to be taken into account that censuses' questionnaires are very different in various countries: firstly, ethnic markers are different (language, religion, etc.), and secondly, depending on the rules / traditions prevailing in the country, respondents are asked different questions.

- In many countries, censuses contain a "direct question" about "ethnicity", sometimes about "identity." In such cases, this information is used;
- In most countries there are no "direct questions" about ethnicity in censuses, but there are questions about the language (mother language, first language, family language, etc.). If the language in this case is a significant marker of ethnicity, these questions were used for coding;
- In several cases, not a language but a religion is a main marker of ethnicity (Adzharia, Rodriguez, the Indian states of Punjab and Jammu and Kashmir), so the variables were coded on the basis of the questions about confessional affiliation;
- In such cases, as the Netherlands, the issue of "origin" was taken into account.
- In number of countries there are no questions, which concern ethnicity, in censuses. Therefore, the share of autonomy in the population of the country was calculated on the basis of the censuses but the other variables were coded on the basis of the other sources, first of all expert evaluations (listed below).

For some autonomies, there are no information on ethnic composition of the population of both country and ERAs even in expert evaluations; the respective cells were left blank.

- All Nigerian ERAs
- Kurdistan

1.Belgium	Census 2011 http://census2011.fgov.be/ However, there are no questions, which concern ethnicity, in censuses. The latest census with the question about language was in 1947. This census is the basis for expert evaluations of ethnic composition of the country (http://www.populstat.info/populhome.html) as well as of separate regions (Hooghe, L. (2003), 'Belgium: From Regionalism to Federalism', in Coakley, J., ed., The Territorial Management of Ethnic
-----------	--

	Conflict (London: Frank Cass). P.73).
2.Bosnia_Herzegovina	<p>Census 2013 http://www.popis2013.ba/index.php/en/</p> <p>However, the results of the census on ethnic questions have not been published because of political sensitivity of the issue.</p> <p>Expert evaluations on both the country and entities of Bosnia and Herzegovina: http://www.worldwizzy.com/library/Federation_of_Bosnia_and_Herzegovina http://www.worldwizzy.com/library/Republic_of_Srpska#Demographics http://www.worldwizzy.com/library/Bosnia_and_Herzegovina#Demographics</p>
3.Canada	<p>Censuses 2001, 2006, 2011 http://www12.statcan.gc.ca</p> <p>Ethnicity is defined on the basis of the question about “mother language” (in Nunavut in 2001 and 2016 – “unofficial language”)</p>
4.China	<p>Censuses 2000 and 2010, but only 2000 is taken, since there are no questions about ethnicity in 2010. http://www.stats.gov.cn/english/statisticaldata/yearlydata/YB2002e/htm/d0408e.htm http://www.stats.gov.cn/english/statisticaldata/yearlydata/YB2002e/htm/d0406e.htm</p> <p>For ERAs, the share of titular groups is taken from: “Population Structure and Changes in the Tibet Autonomous Region” (https://www.case.edu/affil/tibet/tibetanSociety/documents/TAR-Census2.pdf).</p>
5.Denmark	<p>In Denmark, after 1970, registers of the population are used instead censuses⁵. Therefore we have annual data: http://bank.stat.gl/pxweb/en/Greenland/Greenland_BE_BE01/BEXSAT1.px/table/tableViewLayout1/?rxid=62937d53-4c61-4a88-9696-0842213e43ae</p> <p>Faroe Islands: census 2011 http://statbank.hagstova.fo/pxweb/en/H2/H2_IB_IB01/fo_vital_md.px/table/tableViewLayout1/?rxid=6837cd67-4893-4f1a-82fe-1e28acffef9e</p> <p>Ethnicity is defined on the basis of the question about first language http://statbank.hagstova.fo/pxweb/en/H2/H2_MT_MT01_MT0103/MT17.px/table/tableViewLayout1/?rxid=6837cd67-4893-4f1a-82fe-1e28acffef9e</p>

⁵ The Nordic countries, such as Denmark or Finland, in the last few years have developed a mechanism for using data from administrative registers in the census, which makes it possible to rewrite the population without going home. Thus, the population does not participate directly in the census procedure. Since register data are always available, censuses can be conducted annually (<http://www.demoscope.ru/weekly/2012/0531/analit03.php>).

	d=c7d4994d-0926-43bd-ac95-b8a90b9f4440 Greenland – expert evaluations 2010 (CIA The World Factbook)
6.Ethiopia	Census 2007 (there is a direct question about ethnicity): http://www.csa.gov.et/index.php?option=com_phocadownload&view=category&id=104&Itemid=112
7.Fiji	Census 2007 (there is a direct question about ethnicity): http://www.statsfiji.gov.fj/statistics/2007-census-of-population-and-housing
8.Finland	In Finland, registers of the population are used instead censuses . Therefore we have annual data: http://pxnet2.stat.fi/PXWeb/pxweb/en/?rxid=570178fc-78ea-4371-bff9-3c9490c3de40 Ethnicity is defined on the basis of the question about language.
9.France	Censuses 1999, 2007, 2012 http://www.insee.fr/en/ However, there are no questions, which concern ethnicity, in censuses, therefore ethnic composition is taken from GROWup project (groupsize). New Caledonia conducts its own censuses (1996, 2004, 2009, 2014), which include the question on ethnicity http://www.isee.nc/population/recensement/structure-de-la-population-et-evolutions French Polynesia: expert evaluations (CIA The World Factbook). Corsica: expert evaluations https://en.wikipedia.org/wiki/Corsican_language
10.Georgia	Census 2014 http://census.ge/en/results/census Ajarans are not recognized officially as an ethnic group. However, the census contains the information on religion. Therefore the values of variables are calculated as proxies on the assumption that Georgians as a dominant group would be Christians while Ajarans as a titular group would be Muslims.
11.India	Censuses 2001 and 2011, but only 2001 is taken, since there are no questions about ethnicity in 2011. http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/data_on_language.aspx For the country and most of ERAs, ethnicity is defined on the basis of the question about language. In some ERAs, where titular groups are so small peoples that they are not in all-Indian statistics, it is proposed that their number in the state would be the same as their number in the country. In 2 ERAs – Punjab, Jammu and Kashmir, where ethnicity is based on rather religion then language, the questions about religion, not language were taken.
12.Indonesia	Census 2010 r. http://sp2010.bps.go.id/ The results of the census on ethnic questions, which are not available in the web-site. Were taken from:

	<p>“Changing Ethnic Composition: Indonesia, 2000-2010” http://iussp.org/sites/default/files/event_call_for_papers/IUSSP%20Ethnicity%20Indonesia%20Poster%20Section%20G%202708%202013%20revised.pdf For Aceh, the information is available in: Suryadinata, Leo, Evi Nurvidya Arifin, and Aris Ananta. Indonesia's population: Ethnicity and religion in a changing political landscape. No. 1. Institute of Southeast Asian Studies, 2003. Papua: there are no Papuans as an ethnic group in the census; instead there are Javanese (dominant group) and multitude tribes. All these tribes are accounted as Papuans so that Papuans would be all the population of ERA except Javanese. Also, it was proposed that there would be no Papuans outside the region, which allows calculating the values of the variables.</p>
13.Iraq	<p>The latest census was in 1987, for the period of 2000s there are only expert evaluations. Data on the population of the country and regions are taken from the web-site of the Central Statistical Organization of Iraq http://cosit.gov.iq/ar/ (Kurdistan is the sum of three provinces) Ethnic composition of Iraq is based on expert evaluations of CIA https://www.cia.gov/library/publications/the-world-factbook/geos/iz.html The data on ethnic composition of Kurdistan have not been found.</p>
14.Italy	<p>Censuses 2001 and 2011 http://dati-censimentopopolazione.istat.it/Index.aspx?lang=en However, there are no questions, which concern ethnicity, in censuses, therefore the share of Italians as a dominant group in the country is taken from GROWup project. South Tyrol conducts its own censuses, which include the question on language (Statistisches Jahrbuch für Südtirol 2014 / Annuario statistico della Provincia di Bolzano 2014). It was proposed that there would be no South Tyrolians outside the region, which allows calculating the values of the variables. For all other ERAs, the results of the surveys on the usage of language in the family were taken (Lingue e dialetti e lingue straniere in Italia 2006). Aosta Valley: 54% used Italian language in family communication whereas 45% used other languages (they were assumed as a titular ethnic group. Also, it was proposed that there would be no French speakers outside the region, which allows calculating the values of the variables). Friuli Venezia Giulia: 35.8% used Italian language in family communication whereas 62.5% used other languages. However, there are some other non-Italian ethnic groups in the region. When subtracting their</p>

	<p>number (based on estimates) from the total ERA's population, the share of Friulians is approximately 49.2% of the population of the autonomy. Also, it was proposed that there would be no Friulians outside the region, which allows calculating the values of the variables.</p> <p>Sardinia: 52% used Italian language in family communication whereas 48% used other languages (they were assumed as a titular ethnic group. Also, it was proposed that there would be no Sardinians outside the region, which allows calculating the values of the variables).</p> <p>Sicily: 26% used Italian language in family communication whereas 71% used dialect (this is Sicilian language). They were assumed as a titular ethnic group. Also, it was proposed that there would be no Sardinians outside the region, which allows calculating the values of the variables.</p>
15.Malaysia	<p>Census 2010 (there is a direct question about ethnicity):</p> <p>https://web.archive.org/web/20150301154300/http://www.statistics.gov.my/portal/download_Population/files/census2010/Taburan_Penduduk_dan_Ciri-ciri_Asas_Demografi.pdf</p>
16.Mauritius	<p>Censuses 2000 and 2011 http://statsmauritius.govmu.org</p> <p>There are questions on both mother language and religion. But since the differences between the dominant and titular groups are based on rather religion (Indo-Mauritians speak in both Indian and Creole), the questions about religion are used (Indo-Mauritians are mostly Hindus while the population of Rodriguez are Catholics⁶).</p>
17.Moldova	<p>Census 2004 (there is a direct question about ethnicity)</p> <p>http://www.statistica.md/pageview.php?l=ru&idc=295</p>
18.Myanmar	<p>Census 2014 г. http://myanmar.unfpa.org/en/node/15104</p> <p>However, the results of the census on ethnic questions have not been published therefore the data on ethnic composition of the country are taken from GROWup project (groupsize).</p> <p>For the calculation of the ethnic composition of ERAs, the results of the previous census of 1983 were used (Ethnic Conflicts in Southeast Asia / eds. Kusuma Snitwongse, Willard Scott Thompson. Institute of Southeast Asian Studies, 2005. P. 68).</p>
19.Netherlands	<p>The censuses are conducted separately in different part of the Kingdom of the Netherlands, therefore their</p>

⁶ Experts suppose that in Mauritius «religion... is the best local proxy for ethnicity» (<http://www.brandeis.edu/globalbrandeis/documents/Nigel%20-%20Mauritius%20FINAL.pdf>)

	<p>results have been summed up.</p> <p>Netherlands – censuses 2001 and 2011 http://statline.cbs.nl/Statweb/dome/?TH=5410&LA=en. Ethnicity is defined on the basis of the question about origin (“Dutch background” are interpreted as belonging to the dominant group while “Antiles background” – to the titular group that is common for both Aruba and the Netherlands Antilles.</p> <p>Aruba – censuses 2000 and 2010 http://cbs.aw/wp/index.php/category/population/census/</p> <p>Netherlands Antilles (Curaçao) – censuses 2001 and 2010 http://www.cbs.cw/website/statistics-by-subject_3406/</p> <p>In both ERAs ethnicity is defined on the basis of the question about language (Dutch and Papiamentu) though in Curaçao this question is only in 2010, therefore for comparability only the data on 2010-2011 were taken.</p>
20.Nicaragua	<p>Census 2005 (there is a direct question about ethnicity)</p> <p>http://www.inide.gob.ni/censos2005/resumencensal/resumen2.pdf</p>
21.Nigeria	<p>Census 2006 r. http://nigeria.opendataforafrica.org/ifpbxbd/state-population-2006</p> <p>However, there are no questions concerning ethnicity</p>
22.Pakistan	<p>Census 1998⁷ http://www.pbscensus.gov.pk/content/population-and-housing-indicators</p> <p>Ethnicity is defined on the basis of the question about “mother tongue”</p>
23.Panama	<p>Census 2010 (there is a direct question about ethnicity) https://www.contraloria.gob.pa/inec/Default.aspx</p>
24.Papua_New_Guinea	<p>Censuses 2000 and 2011 http://www.nso.gov.pg/</p> <p>There are no questions concerning ethnicity in the censuses, however the specificity of the definition of a titular ethnic group in Bougainville (see above) allows us to assume that all the residents of the ERA belong to a titular ethnic group («Bougainvilleans»). Then we take all the population of the ERA as a titular group and all the population of the country except Bougainville as a dominant group. It allows calculating the values of the variables.</p>
25.Philippines	<p>Censuses 2000 and 2010, but only 2000 is taken, since there are no questions about ethnicity in 2010.</p> <p>http://www.census.gov.ph/content/statistical-tables-sample-variables-results-2010-census-population-and-housing-philippines</p> <p>Ethnicity is defined on the basis of the question about language. Since 13 ethnic groups are related to the</p>

⁷ The latest census was in 2017 but the data have not been published.

	Moro language, they are summed up.
26.Russia	<p>Censuses 2002 and 2010 (there is a direct question about ethnicity)</p> <p>http://www.perepis2002.ru/index.html?id=11</p> <p>http://www.gks.ru/free_doc/new_site/perepis2010/perepis_itogi1612.htm</p>
27.Spain	<p>Censuses 2001 and 2011</p> <p>http://www.ine.es/dynt3/inebase/en/index.htm?type=pcaxis&path=/t20/e242/p03/a2001/&file=pcaxis</p> <p>http://www.ine.es/dynt3/inebase/index.htm?type=pcaxis&path=/t20/e244/avance/p01/&file=pcaxis&L=1</p> <p>However, there are no questions, which concern ethnicity, in censuses, therefore the share of Spaniards as a dominant group in the country is taken from GROWup project.</p> <p>For the separate ERAs the information about the usage of languages from statistical web-sites of these ERAs is used</p> <p>http://www.idescat.cat/pub/?id=eulp&n=7198&lang=en</p> <p>http://www.ige.eu/estatico/estat.jsp?ruta=html/gl/ecv/ECV_ResumoResultados_galego.html#02</p> <p>http://www.navarra.es/AppsExt/GN.InstitutoEstadistica.Web/InformacionEstadistica.aspx?R=1</p> <p>http://en.eustat.eus/estadisticas/tema_458/opt_0/ti_Linguistic_competence/temas.html</p>
28.Sudan	<p>Census 2008</p> <p>http://static1.1.sqspcdn.com/static/f/750842/10069944/1294146347077/Sudan+Census+Priority+Results+English+08+April+2009.pdf?token=JCSP%2FdOxbjcbpdcS2eVfR6UuYX8%3D</p> <p>For South Sudan, the data on 10 Southern states are summed up.</p> <p>Ethnicity is defined on the basis of that «Northern Sudanese» (considered as a dominant group) and «Southern Sudanese» (considered as a titular group) are distinguished in the census.</p>
29.Switzerland	<p>The data on the population in the statistics are dynamic, because annual monitoring is carried out. However, there are no questions, which concern ethnicity; therefore the share of German Swiss as a dominant group (as well as French Swiss and Italian Swiss) is taken from official statistical web-site</p> <p>http://www.bfs.admin.ch/bfs/portal/en/index/themen/01/05/blank/key/sprachen.html</p> <p>Since the titular ethnic group of Jura is French speaking Jurassians, not French Swiss, it is possible to assume that there would be all the residents of the ERA belong to a titular ethnic group and there would be no French speaking Jurassians outside the region, which allows calculating the values of the variables.</p>
30.Tajikistan	<p>Censuses 2000 and 2010 http://www.stat.tj/ru/img/526b8592e834fcaaccec26a22965ea2b_1355501132.pdf</p>

	http://www.stat.tj/en/img/b417f44e3113e555ffff3cd143d5b3fe_1404817165.pdf There is a direct question about ethnicity in the censuses, however Pamiris (titular group of Gorno-Badakhshan) are not officially recognized in the country as a specific ethnic group and included in the category of Tajiks. Therefore, their number in the country and in the autonomy is calculated on the basis of expert evaluations: http://www.demoscope.ru/weekly/2005/0191/analit05.php
31.Tanzania	Censuses 2002 and 2012 http://www.nbs.go.tz/ There are no questions concerning ethnicity in the censuses, however the specificity of the definition of a titular ethnic group in Zanzibar (see above) allows us to assume that all the residents of the ERA belong to a titular ethnic group («Zanzibaris»). Then we take all the population of the ERA as a titular group and all the population of the Mainland as a dominant group. It allows calculating the values of the variables.
32.Ukraine	Census 2001 http://2001.ukrcensus.gov.ua/rus/ There is a direct question about “nationality” (in the sense of “ethnicity”) in the census.
33.United_Kingdom	Census 2011 https://www.ons.gov.uk/census/2011census There are no questions concerning ethnicity in the census, therefore the share of a dominant group in the country is taken from CREG project. The data on ethnic composition in ERAs are taken of the surveys on the issues of identity.
34.Uzbekistan	The latest census was conducted in 1989. Therefore the official estimations of the population size is used. http://stat.uz/ru/demograficheskie-dannye For ethnic composition, expert evaluations are used: for the country - CIA World Factbook; for Karakalpakstan – Mayhew, Bradley (2007). Central Asia: Kazakhstan, Tajikistan, Uzbekistan, Kyrgyzstan, Turkmenistan. Lonely Planet. p. 258.

4. Political characteristics of the host country

Description of the variable	Label	Type	Sources	Comments
4. Political characteristics of the host country				
4.1. General political and institutional indicators				
Political regime (Polity)	polity	Ordinal	Polity IV http://www.systemicpeace.org/inscrdata.html	
Political regime (Freedom House)	freedom_house	Ordinal	Freedom House https://freedomhouse.org/report-types/freedom-world	
Government system	gov_arrang	Ordinal	Database of Political Institutions 2015 http://www.iadb.org/en/research-and-data/publication-details,3169.html?pub_id=IDB-DB-121	Parliamentary (2), Assembly-elected President (1), Presidential (0)
Federation	fed_arrang	Binominal	Lists of federations: Anderson L. Federal Solutions to Ethnic Problems: Accommodating Diversity. New York, 2013; Bednar J. The Robust Federation. New York, 2008; Elazar D. Federal Systems of the World: a Handbook of Federal, Confederal and Autonomy Arrangements. Harlow, 1994; Watts R. Comparing Federal Systems in the 1990s. Kingston, 1996	(0) no (1) yes
Federalism and its types	ethnofed_type	1 string and 3 binominal	Anderson (2014) Ethnofederalism: The Worst Form of Institutional Arrangement	
Number of sub-national units of the first level	number_units	Interval	http://www.statoids.com/	

Number of ethnic sub-national units of the first level	number_ethnic_units	Interval	On the basis of original list of ERA	
4.2. Lower house of national parliament and representation of ERA				
Electoral system	elect_syst	1 string and 5 binominal	Electoral statistics*	FPTP - First Past The Post System Two-Round – Two-Round System AlterVote – Alternative Vote System Parallel – Parallel System LPR – List Proportional Representation
Number of seats	total_seats	Interval	Electoral statistics*	
The coincidence of boundaries of electoral districts and boundaries of administrative units	district_reg	Ordinal	Electoral statistics*	(0) no (1) partially (2) yes (FPTP or LPR if mmds coincide with administrative units)
Number of deputies in the lower house from the autonomy	aut_seats	Interval	Electoral statistics*	It is applicable only if the variable «district_reg» takes the values of (1) and (2)
Share of deputies from the autonomy in the total number of seats	aut_seats_share	Interval	aut_seats / total_seats	In the cases, when «district_reg» = (1), the share is calculated on the basis of applicable deputies
Index of representation of a region in the lower house	aut_seats_index	Interval	aut_seats_share / share_pop_aut	
The presence of ethno-regional parties of the autonomy in national politics*	erp_nat	Binominal	Electoral statistics*	(0) no (1) yes
Share of ethno-regional parties of the autonomy in the number of deputies elected in the lower house from the autonomy *	erp_share_natleg	Interval	Electoral statistics*	
4.3. Upper house of national parliament and representation of ERA				
Upper house of national	upper_house	Binominal	Electoral statistics*	(0) no

parliament				(1) yes
Number of seats	up_seats	Interval	Electoral statistics*	
Regional basis of the formation of upper house	up_reg	Binominal	Electoral statistics*	(0) no (1) yes, including the cases of partially regional basis of the formation of upper house
Number of deputies in the upper house from the autonomy	aut_seats_uh	Interval	Electoral statistics*	
Share of deputies from the autonomy in the total number of seats of upper house	aut_seats_uh_share	Interval	aut_seats_uh / uh_seats	
Index of representation of a region in the upper house	aut_seats_uh_index	Interval	aut_seats_uh_share / share_pop_aut	
4.4. Central government				
Governmental power-sharing in national government*	power_sharing_natgov	Ordinal	Electoral statistics*	(2) formal or informal power-sharing arrangements with participation of ethno-regional parties of the autonomy (1) – situational post-electoral coalitions with participation of ethno-regional parties of the autonomy (0) – no power-sharing

Notes and comments

Missing data

The data are coded as "not applicable" and respective cells are marked with "NA" in the following cases:

- China: “gov_arrang” cannot be coded on the scale of DPI (political system of the Soviet type)
- Myanmar (until 2010): “gov_arrang” cannot be coded on the scale of DPI (military rule)

- South Tyrol: “aut_seats” and other variables associated with it cannot be coded as deputies of the national parliament are elected from the region Trentino-Alto Adige, not from South Tyrol
- Aruba and Netherlands Antilles: “aut_seats” and other variables associated with it cannot be coded as these ERAs don’t participate in the elections of the Nitherland parliament
- Rotuma (2014-15), Gagauzia, Russian ERAs (2007-15), Crimea (2006-11): “aut_seats” and other variables associated with it cannot be coded as “district_reg” takes the value “0”
- In some cases, there is no upper house of the national parliament

For some autonomies, the information has not be found; the respective cells were left blank.

- “aut_seats” in Karakalpakstan (2001-09)
- “aut_seats_uh” in Ethiopian ERAs (2005-09 and 2015)

Ethno-regional political parties

Unlike all- national parties (ANP), ethno-regional parties (ERP) are political parties which manifest themselves as representatives of particular ethnic groups, and their activities are primarily limited to the regions where respective groups are localized.

The ERAD takes into account only ERPs of ethnic groups, which are titular ones for a given ERA or ERPs of ethnic minorities localized in a given ERA:

- Regional parties, which represent a dominant ethnic group in the country (unionist parties in Northern Ireland), are not considered as ERPs
- Ethnic political parties, which represent a dominant ethnic group in the country (Indo-Fijian and native-Fijian parties), are not considered as ERPs
- Ethnic political parties, which represent an ethnic minority, if the latter is not localized in ERA, are not considered as ERPs (for example, Chinese and Indian parties in Malaysia are not considered as ERPs, because these ethnic communities are not localized, unlike the Kadazan and Daiak parties, which are considered as ERPs as these ethnic groups are titular ones for Sarawak and Sabah. Similarly, Mohajirs in Pakistan: although to some extent they are localized in Sind, it is not sufficient to consider them as an ethnic minority of Sind)

In most cases, ERPs are not hard to be identified on the basis of electoral statistics and the criteria presented above. In some fairly complicated cases (India, Italia, Spain, Pakistan) special literature was used (Brancati, 2007; Dandoy, 2010; Schakel, 2013; Schakel and Swenden, 2016; Oziashvili, 2012).

Coding rules

- The binominal variable "erp_nat" concerns the very fact of the presence of ERPs of the respective ERA in national politics whereas the variable "erp_share_natleg" measures the “weight” of ERPs through the share of the ERP of the given ERA among the deputies elected from the ERA to the national parliament (if there are several ERPs, their deputies were summed up).
- If the value of the variable "erp_nat" = 0, the variable "erp_share_natleg" also takes the value "0". If the variable "erp_nat" = 1, the variable "erp_share_natleg" takes a value in the interval [0 ... 1].
- The variable "erp_nat" was assigned a value also in those cases when other variables from this section were coded as "NA". If its value is "0", then the variable "erp_share_natleg" in all cases was assigned the value "0". If its value is "1", in those cases (see above), when it is impossible to determine the composition of deputies elected from the ERA to the national parliament, it was coded as "NA".

Power-sharing in national government

Power-sharing is defined, in a broad sense, as a set of institutional arrangements that ensure participation of the representatives of politically salient ethnic groups in political decision-making, including the allocation of governmental positions on the basis of ethnicity (Lijphart, 2002, p. 39).

In the ERAD power-sharing is understood in a narrow sense - as a governmental power-sharing. In the case of national governments, power-sharing is recorded when any one or several ERPs of this ERA participate in a government coalition at the national level. It may be either a coalition of ANP and ERP of a given ERA or a coalition of ERPs of different ERAs (there are several cases - Belgium, BiH, Ethiopia - when ANP are absent, therefore the government is formed entirely from the ERPs)

Coding rules

- In the full sense of the word, governmental power-sharing presuppose that the participation of the ERP in the government is guaranteed irrespective of the results of the elections. In these cases the variable “power_sharing_natgov” takes the value of "2".

- Frequently, ERPs participate in a coalition government and as a result of certain election results. In these cases the variable “power_sharing_natgov” takes the value of "1".
- The variable “power_sharing_natgov” is coded as "1" or "2" only if ERPs of a given ERA participate in the national government. Personal membership of the ERA’s officials in the national government (Aruba, Netherlands Antilles, Nigerian states, Zanzibar, Gagauzia, etc.) is not considered as a power-sharing
- The variable “power_sharing_natgov” is coded as "0", if a coalition includes EPR, but these are ERP from the other ERA
- The variable “power_sharing_natgov” is coded as "0", if a coalition government is based on power-sharing arrangements, but it involves only ANPs as, for instance in Switzerland
- The variable "power-sharing_natgov" was assigned a value also in those cases when other variables from this section were coded as "NA".

Electoral statistics

Some systematized datasets are the main sources of information

Foundation for Electoral Systems (IFES) <http://www.electionguide.org/>

ACE Electoral Knowledge Network <http://aceproject.org/>

International Institute for Democracy and Electoral Assistance (IDEA) <http://www.idea.int>

PARLINE Database on National Parliaments <http://www.ipu.org/parline>

European Election Database http://www.nsd.uib.no/european_election_database

African Elections Database <http://africanelections.tripod.com/index.html>

Psephos Adam Carr's Election Archive <http://psephos.adam-carr.net/about/about.shtml>

Election Passport <http://www.electionpassport.com/>

Global Elections Database <http://www.globalelectionsdatabase.com/index.php/index>

Constituency-Level Elections Archive (CLEA) <http://www.electiondataarchive.org/>

Democratic Electoral Systems (DES) dataset <http://mattgolder.com/elections>

In some cases additional sources (national electoral statistics) are used:

Bosnia and Herzegovina –Statistical Yearbooks <http://www.rzs.rs.ba> <http://www.fzs.ba>

Canada <http://www.elections.ca/home.aspx>

France <https://www.interieur.gouv.fr/Elections>

Georgia <http://cesko.ge/eng>

India <http://elections.traceall.in/> <http://www.elections.in/>

Myanmar Burma's 2010 Elections: a comprehensive report. The Burma Fund-UN Office, New York, 2011.

Pakistan <https://ecp.gov.pk/>

Russia <http://www.cikrf.ru/>

Tanzania <http://nec.go.tz/>

United Kingdom <https://www.electoralcommission.org.uk> <http://www.parliament.uk/about/how/elections-and-voting/>

5. Political characteristics of an autonomy

Description of the variable	Label	Type	Sources	Comments
5. Political characteristics of an autonomy				
Government system in the autonomy	aut_gov_arrang	Ordinal	Coded on the basis of the scale of the Database of Political Institutions 2015	Parliamentary (2), Assembly-elected President (1), Presidential (0)
Electoral system in regional legislature's elections	regleg_electsystem	String	Electoral statistics (see notes to Section 4)	
Number of seats in regional legislature	regleg_seats	Interval	Electoral statistics (see notes to Section 4)	
The presence of ethno-regional parties of the autonomy in regional politics (on the ERA)*	erp_reg	Binominal	Electoral statistics (see notes to Section 4)	(0) no (1) yes
Share of ethno-regional parties of the autonomy in the number of deputies of a regional legislature*	erp_share_regleg	Interval	Electoral statistics (see notes to Section 4)	
Upper house of regional legislature	aut_upperhouse	Binominal	Electoral statistics (see notes to Section 4)	(0) no (1) yes
Number of seats in the upper house of regional legislature	aut_up_seats	Interval	Electoral statistics (see notes to Section 4)	
The presence of ethno-regional parties of the autonomy in regional government*	erp_reggov	Binominal	Electoral statistics (see notes to Section 4)	(0) no (1) yes, in the coalition with national parties (2) yes, only ethno-regional parties
Governmental power-sharing in regional government*	power_sharing_reggov	Ordinal	Electoral statistics (see notes to Section 4)	(2) formal or informal power-sharing arrangements with participation of ethno-regional parties of the autonomy (1) – situational post-electoral coalitions with participation of ethno-regional parties of the autonomy

				(0) – no power-sharing
Ethnic sub-regional autonomies	ethnic_subregional_unites	Binominal	http://www.iso.org/iso/country_codes http://www.statoids.com/	(0) no (1) yes

Notes and comments

Missing data

The data are coded as "not applicable" and respective cells are marked with "NA" in the following cases:

- China: “aut_gov_arrang” and “regleg_electsystem” cannot be coded (political system of the Soviet type)
- Myanmar (until 2010): “aut_gov_arrang” and “regleg_electsystem” cannot be coded (military rule)
- Rotuma and ERAs of Panama: “aut_gov_arrang” and “regleg_electsystem” cannot be coded as regional parliament is formed on the traditional ground
- Rotuma, Ajara (2004-15), ERAs of Myanma, Gorno-Badakhshan, ERAs of Russia (2005-11): “aut_gov_arrang” cannot be coded as the leader of ERA is appointed by the central authorities
- Karakalpakstan, Dagestan (2001-04): “aut_gov_arrang” cannot be coded on the basis of DPI scale because of very specific institutional arrangements
- In some cases, there is no upper house of the national parliament
- Aceh: the first elections of regional parliament were held only in 2009 while the Governor was elected in 2006.
- Bougainville: the first elections of regional parliament were held only in 2005
- Chechnya: the first elections of regional parliament were held only in 2005
- South Tyrol: until 2003, regional elections were held on the level of the region Trentino-Alto Adige

For some autonomies, the information has not been found; the respective cells were left blank.

- Ajara until 2004 (personalist regime of Abashidze)
- Papua – composition of the regional parliament
- Gorno-Badakhshan – electoral system and composition of the regional parliament

- Karakalpakstan – electoral system and composition of the regional parliament
- ERAs of Myanmar – composition of the regional governments

Ethno-regional political parties

ERPs are interpreted in the same way as in the section 4

Coding rules are the same as in the section 4

- The binominal variable "erp_reg" concerns the very fact of the presence of ERPs of the respective ERA in regional politics whereas the variable "erp_share_regleg" measures the “weight” of ERPs through the share of the ERPs among the deputies of regional parliament (if there are several ERPs, their deputies were summed up).
- If the value of the variable "erp_reg" = 0, the variable "erp_share_regleg" also takes the value "0". If the variable "erp_reg" = 1, the variable "erp_share_regleg" takes a value in the interval [0 ... 1].
- The variable "erp_reg" was assigned a value also in those cases when other variables from this section were coded as "NA". If its value is "0", then the variable "erp_share_regleg" in all cases was assigned the value "0". If its value is "1", in those cases (see above), when it is impossible to determine the composition of regional legislature, it was coded as "NA".

Power-sharing in regional government

Power-sharing is interpreted in the same way as in the section 4

Coding rules (the variable “erp_reggov”)

- The variable “erp_reggov” concerns the very fact of the presence of ERPs in the regional government.
- It is coded through the ordinal scale. The value of "2" is assigned if in the regional government consists of only the ERPs of this autonomy; "1" - if ERPs participate in the regional government in the coalition with ANPs; "0" - if there no EPRs of this autonomy in the regional government.
- The variable "erp_reggov" was assigned a value also in those cases when other variables from this section were coded as "NA".

Coding rules (the variable “power_sharing_reggov”)

- The variable “power_sharing_reggov” concerns power-sharing arrangements in the regional government in the sense that both ANPs and ERPs participate in governmental coalition
- It is coded as “2” if participation of both ANPs and ERPs in the government is guaranteed irrespective of the results of the elections.
- It is coded as “1”, if a coalition of ANPs and ERPs is a result of certain election results.
- It is coded as “0”, if there is no a coalitional government of ANPs and ERPs, including the cases when there is a coalition of some ERPs (Kurdistan in Iraq) or some ANPs (Jura in Switzerland)
- In the cases when there are no ANPs in the country, a coalition between ERPs of the ERA and ERPs of the other ERA is coded as a power-sharing (Bosnia and Herzegovina)
- Informal arrangements such as distribution of governmental positions between ethnic groups on personal ground does not considered as power-sharing in the ERAD; and the value of the variable is “0” (informal power-sharing arrangements in some ERAs in Russia, China, Nigeria)
- The variable "power_sharing_reggov" was assigned a value also in those cases when other variables from this section were coded as "NA".

6. Preferential policies

6. Preferential policies				
6.1. RAI indices (absolute values)*				
The extent to which a regional government is autonomous rather than deconcentrated	tier_instdepth	Ordinal	RAI	Self-rule
The range of policies for which a regional government is responsible	tier_policy	Ordinal	RAI	
The extent to which a regional government can independently tax its population	tier_fiscauto	Ordinal	RAI	
The extent to which a regional government can borrow	tier_borrowauto	Ordinal	RAI	
he extent to which a region has an independent legislature and executive, which is the sum of assembly and executive	tier_rep	Ordinal	RAI	
The extent to which regional representatives co–determine national legislation	tier_lawmaking	Ordinal	RAI	Shared-rule
The extent to which a regional government co–determines national policy in intergovernmental meetings	tier_execon	Ordinal	RAI	
The extent to which regional representatives co–determine the distribution of national tax revenues	tier_fiscon	Ordinal	RAI	
The extent to which a regional government co–determines subnational and national borrowing constraints	tier_borrowcon	Ordinal	RAI	
The extent to which regional representatives co–determine constitutional change	tier_constit	Ordinal	RAI	
Total index of autonomy (self-rule)	tier_selfrule	Ordinal	RAI	Total
Total index of participation (shared-rule)	tier_sharedrule	Ordinal	RAI	
Total index RAI	tier_RAI	Ordinal	RAI	
6.2. RAI indices (deviation from the standard regions)*				
The extent to which a regional government is autonomous rather than deconcentrated	tier_instdepth_dist	Ordinal	Calculated on the basis of RAI	Self-rule
The range of policies for which a regional government is responsible	tier_policy_dist	Ordinal	Calculated on the basis of RAI	
The extent to which a regional government can independently tax its population	tier_fiscauto_dist	Ordinal	Calculated on the basis of RAI	

The extent to which a regional government can borrow		tier_borrowauto_dist	Ordinal	Calculated on the basis of RAI	Shared-rule
he extent to which a region has an independent legislature and executive, which is the sum of assembly and executive		tier_rep_dist	Ordinal	Calculated on the basis of RAI	
The extent to which regional representatives co–determine national legislation		tier_lawmaking_dist	Ordinal	Calculated on the basis of RAI	
The extent to which a regional government co–determines national policy in intergovernmental meetings		tier_execon_dist	Ordinal	Calculated on the basis of RAI	
The extent to which regional representatives co–determine the distribution of national tax revenues		tier_fiscecon_dist	Ordinal	Calculated on the basis of RAI	
The extent to which a regional government co–determines subnational and national borrowing constraints		tier_borrowcon_dist	Ordinal	Calculated on the basis of RAI	
The extent to which regional representatives co–determine constitutional change		tier_constit_dist	Ordinal	Calculated on the basis of RAI	Total
Total index of autonomy (self-rule)		tier_selfrule_dist	Ordinal	Calculated on the basis of RAI	
Total index of participation (shared-rule)		tier_sharedrule_dist	Ordinal	Calculated on the basis of RAI	
Total index RAI		tier_RAI_dist	Ordinal	Calculated on the basis of RAI	
6.3. Linguistic policy					
Official status of the language of the titular group	titul_lang_status	Ordinal	official acts*	(0) no; (1) yes, on the level of ERA; (2) yes, on the national level; (3) yes, on both regional and national levels ⁸	
Learning the language of the titular group in secondary schools	titul_lang_educ	Ordinal	official acts, expert evaluations, academic publications, and other sources*	(0) no; (1) optional; (2) compulsory for the members of the titular group (3) compulsory for all the population of ERA	
The presence of schools, where the language of the titular group or other ethnic minorities is the language of instruction	titul_lang_of_instr	Ordinal	official acts, expert evaluations, academic publications, and other sources*	(0) no; (1) yes	

⁸ There are some cases when the language of the titular group has official status on national level, but has not on regional level (Jura in Switzerland, Jammu and Kashmir in India). Therefore, it needs to distinguish the values “2” and “3”.

Notes and comments

Index RAI: content and missing data

The Regional Authority Index (RAI) is the most extensive and systematic database in contemporary political science, which compiles and codes the information on a wide range of issues related to the degree of autonomy (self-rule block) of sub-national units, as well as their relationship with central authorities (shared-rule block), in 81 countries around the world for the period 1950 -2010 years. (Hooghe at al, 2016). Each of two blocks consists of 5 variables, which are included in the ERAD

- For South Tyrol, the provincial indicators (as an exception), rather than the Trentino-Alto Adige region, are taken.
- For Russia, Tatarstan and Bashkortostan in RAI are coded separately, all the other 19 republics are in the category of "republics", and all other ERAs (autonomous okrugs and autonomous regions) are in the category of "subjects of the federation".

Despite a fairly large number of countries in the RAI (81 countries), only 16 of the 34 countries, which have ERAs, are represented in it (not counting the Netherlands, since Aruba and Netherland Antilles are not represented in the RAI). In addition, RAI lacks French Polynesia and New Caledonia. Thus, the RAI data could be used to measure the relevant variables for only 65 ERAs out of 139. Units of observation for all missing in the RAI ERA are coded as "NA". Similarly, the units of observation for 2011-2015 are coded as "NA".

The RAI contains only absolute values of the indices; however special preferences of ERAs in comparison to “standard regions” are of particular interest for the ERAD, therefore the variables, which measure deviation of the ERAs from the standard regions, have been created for all indices. The values of these variables were calculated by subtracting the value of the index of the standard region from the value of the ERA index.

- In the ERAs of Belgium and Bosnia and Herzegovina, there are no referents for comparison, since there are no "standard regions".
- For the same reason, it is impossible to correlate Mindanao (Philippines) and "standard region".

Consequently, the number of ERAs that can be measured in this part of the ERAD has been reduced to 60. Units for all other ERAs are coded as "NA".

Linguistic policy

In the 5 ERAs (Ajara in Georgia, Zanzibar in Tanzania, Ningxia Hui in China, Rodrigues in Mauritius and Sicily in Italy), the titular groups differ from the dominant group not by language. Therefore, all the variables in the linguistic policy in these cases are coded as "NA". From the linguistic point of view, there are also no differences between the main ethnic groups in Bosnia and Herzegovina - Serbs, Croats and Bosniaks (in the ERAD, the variable "dist_lang" is coded as "0"), however, all three ethnic groups carry out strong politicization of the language, therefore all the variables in the linguistic policy in these cases are coded.

Sources of information for the coding of the variables in the linguistic policy

Systematized sources

Ethnologue site. Languages of the world. Belgium. [Electronic document]. URL: <http://www.ethnologue.com/country/BE/languages> (circulation date June 19, 2016)

World data on education VII Ed. 2010/11. [Electronic document] URL: http://www.ibe.unesco.org/fileadmin/user_upload/Publications/WDE/2010/pdf-versions/China.pdf (circulation date June 27, 2016)

Website of the Council of Europe. European Charter for Regional or Minority Languages with an explanatory report (translated into Russian). [Electronic resource]. URL: http://www.coe.int/t/dg4/education/minlang/textcharter/Charter/Charterexpl_eng.pdf. (date of circulation: 3/10/2016.)

Site of EuroEducation.Net. The European Education Directory. Bosnia & Herzegovina. [Electronic document]. URL: <http://www.euroeducation.net/prof/boherco.htm> (circulation date June 19, 2016)

Council of Europe. Language Education Policy Profiles. [Electronic document]. URL: <http://www.coe.int/t/dg4/linguistic>

Council of Europe. National or Regional Reports. [Electronic document]. URL: <http://www.coe.int/t/dg4/linguistic>

Additional sources on the individual countries and ERAs

Belgium

Décret relatif à l'organisation pédagogique du 1er degré de l'enseignement secondaire, 2006. http://www.gallilex.cfwb.be/document/pdf/30998_008.pdf

Steven Delarue & Johan De Caluwe. Eliminating social inequality by reinforcing standard language ideology? Language policy for Dutch in Flemish schools. Journal: Current Issues in Language Planning Volume 16, Issue 1-2, April 2015, pages 8-25. DOI:10.1080/14664208.2014.947012

Bosnia and Herzegovina

BiH Educational Authorities: Shared modernization strategy for primary and general secondary education in BiH, October 2003. http://www.esiweb.org/pdf/bridges/bosnia/BiH_Edu_Education_modernisation.pdf

Wolff S., van Houten P., Anghelea A.M., Djuric I. Minority Rights in the Western Balkans, 2008.

<http://www.stefanwolff.com/publications/minority-rights-in-the-western-balkans>

Primary and secondary education in Bosnia and Herzegovina (current state and recommendations for reforms). <http://eu-monitoring.ba/site/wp-content/uploads/2015/06/Primary-and-secondary-education-in-Bosnia-and-Herzegovina.pdf>.

Canada

Charter of the French language Cet hyperlien s'ouvrira dans une nouvelle fenêtre. – Chapter VIII – The language of instruction. <http://legisquebec.gouv.qc.ca/en/ShowDoc/cs/C-11>.

CONSOLIDATION OF OFFICIAL LANGUAGES ACT S.Nu. 2008. http://www.gov.nu.ca/sites/default/files/2015-07-28-official_language_act-conssnu2008c10.pdf.

Inuktitut and Inuinnaqtun resources // Ntip.gov.nu.ca. <http://ntip.gov.nu.ca/new-hires/essential-information/inuktitut-and-inuinnaqtun-resources>.

China

Language Policy in the Republic of china. Annual Review of Applied Linguistics / Volume 2 / Mapr 1981, pp 33-47 DOI: <http://dx.doi.org/10.1017/S0267190500000234>.

John Kwock-Ping Tse. Language Policy in the Republic of china. Annual Review of Applied Linguistics / Volume 2 / March 1981, pp 33-47 DOI: <http://dx.doi.org/10.1017/S0267190500000234>

Anwei Feng & Mamtimyn Sunuodula. Analysing language education policy for China's minority groups in its entirety. Journal: International Journal of Bilingual Education and Bilingualism Volume 12, Issue 6, November 2009, pages 685-704. DOI:10.1080/13670050802684396

Minglang Zhou. Language Policy and Illiteracy in Ethnic Minority Communities in China. Journal: Journal of Multilingual and Multicultural Development Volume 21, Issue 2, March 2000, pages 129-148. DOI:10.1080/01434630008666398

Wuyuncang Ojijed. Language Competition in an Ethnic Autonomous Region. Journal: Chinese Education & Society Volume 43, Issue 1, January 2010, pages 58-69. DOI:10.2753/CED1061-1932430105

Agnes S. L. Lam. The multi-agent model of language choice: national planning and individual volition in China. Journal: Cambridge Journal of Education Volume 37, Issue 1, March 2007, pages 67-87. DOI:10.1080/03057640601179186

Wang Tiezhi. Preferential policies for ethnic minority students in China's college/university admission. Journal: Asian Ethnicity Volume 8, Issue 2, June 2007, pages 149-163. DOI:10.1080/14631360701406288

Denmark

Education in Greenland // Alaska Native knowledge network. URL: <http://www.ankn.uaf.edu/IEW/edgreen.html> (accessed: 17.04.2017).

Education in the Faroe Islands // En.wikipedia.org. URL:

https://en.wikipedia.org/wiki/Education_in_the_Faroe_Islands#cite_note-Skulabladid.fo-14 (accessed: 17.04.2017).

GREENLAND'S LANGUAGE SUCCESS // Office of the Languages Commissioner of Nunavut. URL:

<http://langcom.nu.ca/blog-entry/dshadbolt/greenlands-language-success> (accessed: 17.04.2017).

Greenland Inuits and problems in language teaching // Sorosoro.org. URL: <http://blog.sorosoro.org/en/greenland-inuits-and-problems-in-language-teaching> (accessed: 17.04.2017).

Ethiopia

Midega, M. (2017). Ethiopian Federalism and the Ethnic Politics of Divided Cities: Consociationalism without Competitive Multiparty Politics in Dire Dawa. *Ethnopolitics*, 16(3), 279-294. doi:10.1080/17449057.2016.1254412

Seidel, K., & Moritz, J. (2009). Changes in Ethiopia's Language and Education Policy – Pioneering Reforms? Changes in Ethiopia's Language and Education Policy – Pioneering Reforms?, (pp. 1125-1138). Trondheim. URL:

<http://portal.svt.ntnu.no/sites/ices16/Proceedings/Volume%204/K.%20Seidel%20and%20J.%20Moritz%20-%20Changes%20in%20Ethiopia%E2%80%99s%20Language%20and%20Education%20Policy.pdf> (accessed: 15.06.2017).

Fiji

On the Use of Rotuman Language // University of Hawai'i. URL:

<http://www.hawaii.edu/oceanic/rotuma/os/Forum/Forum28.html> (accessed: 15.06.2017).

Ministry Looks at Language Policy // Fiji Government Online Portal. URL: <http://www.fiji.gov.fj/Media-Center/Press-Releases/Ministry-looks-at-language-policy.aspx> (accessed: 15.06.2017).

Dutcher N. The Use of First and Second Languages in Education: a Review of International Experience // Pacific Islands discussion paper series; no. 1. Washington, D.C.: The World Bank. URL:

<http://documents.worldbank.org/curated/en/131161468770987263/pdf/multi-page.pdf> (accessed: 15.06.2017).

Finland

The example of Åland: autonomy as a minority protector // the Ministry for Foreign Affairs, Department for Communications. URL: <http://finland.fi/public/default.aspx?contentid=160122> (accessed: 15.06.2017).

Scarpulla C. The Constitutional Framework for the Autonomy of Åland: A Survey of the Status of an Autonomous Region // MEDDELANDEN FRÅN ÅLANDS HÖGSKOLA. 2002. URL: <http://www.ha.ax/files/scarpulla.pdf> (accessed: 15.06.2017).

France

Schiffman HF. Linguistic culture and language policy, 1998. URL: <http://www.english.illinois.edu/-people-/faculty/debaron/584/584reading/frenchlangpol.pdf> (accessed: 15.06.2017).

Robert J. Blackwood. L'Exception française? Post-war Language Policy on Corsica. Journal: Journal of Multilingual and Multicultural Development Volume 28, Issue 1, January 2007, pages 18-33. DOI:10.2167/jmmd463.1

Kanak languages gaining popularity again in New Caledonia // Radio Australia. URL: <http://www.radioaustralia.net.au/international/radio/onairhighlights/kanak-languages-gaining-popularity-again-in-new-caledonia> (accessed: 15.06.2017).

Christine Hélot, Andrea Young. The Notion of Diversity in Language Education: Policy and Practice at Primary Level in France. Journal: Language, Culture and Curriculum Volume 18, Issue 3, November 2005, pages 242-257. DOI:10.1080/07908310508668745

Philip Riley. The state, the activists and the islanders. Language policy on Corsica. Journal: Journal of Multilingual and Multicultural Development Volume 31, Issue 4, July 2010, pages 435-437. DOI:10.1080/01434630903497996

Tom Moring. Discourse and struggle in minority language policy formation: Corsican language policy in the EU context of governance. Journal: Journal of Multilingual and Multicultural Development Volume 33, Issue 3, May 2012, pages 321-323. DOI:10.1080/01434632.2012.656972

India

Benedikter, T. (2013). Language Policy and Linguistic Minorities in India: An appraisal of the linguistic rights of minorities in India.

Indonesia

Astri Y. Does the Spoken Language help or Hinder the Spelling in the Second Language? A Case of Indonesian and Acehese Language // The Office of Graduate Studies of Texas A&M University. 2009. URL: <http://oaktrust.library.tamu.edu/bitstream/handle/1969.1/ETD-TAMU-2009-05-530/YULIA-THESIS.pdf>

Iraq

Three more languages recognized officially in Iraq // World bulletin, 09 January 2014. URL:

<http://www.worldbulletin.net/?aType=haber&ArticleID=126556> (accessed: 27.06.2016)

Italy

SPECIAL STATUTE FOR TRENTINO-ALTO ADIGE. Modified text of the Constitution of the “Trentino – Alto Adige” Region and the Provinces of Trento and Bolzano. Art.19

Caŕ Ufficio Scolastico Regionale per il Friuli Venezia Giulia. Tutela della lingua e della cultura friulana. URL:

http://www.scuola.fvg.it/usr/fvg/USRFVG/Progetti_scuola/lingue/lingue_minoritarie/Friulano.html (accessed 27.06.2016)

Institut de Sociolingüística Catalana. Sardinian in Italy. URL:

<https://www.uoc.edu/euromosaic/web/document/sard/an/i1/i1.html#2.4> (accessed: 27.06.2016).

Gianmario Demuro, Francesco Mola, Ilenia Ruggiu. Identità e Autonomia in Sardegna e Scozia. Maggioli S.p.A. 2013.

Sicilie. Regione Autonoma Siciliana // L'aménagement linguistique dans le monde. URL:

<http://www.axl.cefan.ulaval.ca/europe/italiesicile.htm> (accessed: 27.06.2016).

Sicilian language // Revolv.com. URL:

https://www.revolv.com/main/index.php?s=Sicilian%20language&item_type=topic (accessed: 25.08.2017)

Malaysia

Malaysia Education Act 1996 // Portal of Education Plans and Policies. URL:

http://planipolis.iiep.unesco.org/upload/Malaysia/Malaysia_Education_Act_1996.pdf (accessed: 15.06.2017).

Asmah Haji Omar. Languages in the Malaysian Education System: Monolingual Strands in Multilingual Settings. Routledge, 2015. URL: https://books.google.ru/books?id=iaQ0CwAAQBAJ&hl=ru&source=gbp_navlinks_s (accessed: 15.06.2017).

Saran Kaur Gill. Change in Language Policy in Malaysia: The Reality of Implementation in Public Universities. Journal: Current Issues in Language Planning Volume 7, Issue 1, February 2006, pages 82-94. DOI:10.2167/cilp083.0

Carolina López C.. Language is the Soul of the Nation: Language, Education, Identity, and National Unity in Malaysia. Journal: Journal of Language, Identity & Education Volume 13, Issue 3, May 2014, pages 217-223.

DOI:10.1080/15348458.2014.919812

Santhiram R. Raman & Tan Yao Sua. Ethnic segregation in Malaysia's education system: enrolment choices, preferential policies and desegregation. Journal: Paedagogica Historica Volume 46, Issue 1-2, February 2010, pages 117-131. DOI:10.1080/00309230903528496

Malasie // L'aménagement linguistique dans le monde. URL: <http://www.axl.cefan.ulaval.ca/asie/malaysia.htm> (accessed: 25.08.2017)

Moldova

ЗАКОН No. 3465 от 01.09.1989 О функционировании языков на территории Молдавской ССР.

<http://lex.justice.md/ru/312813/> ЗАКОН Nr. 344 от 23.12.1994 об особом правовом статусе Гагаузии (Гагауз Ери).

<http://lex.justice.md/viewdoc.php?action=view&view=doc&id=311656&lang=2>

Matthew H. Ciscel. Uneasy Compromise: Language and Education in Moldova. Journal: International Journal of Bilingual Education and Bilingualism. Volume 11, Issue 3-4, July 2008, pages 373-395. DOI:10.1080/13670050802148756

Steven D Roper. Regionalism in Moldova: The Case of Transnistria and Gagauzia. Journal: Regional & Federal Studies Volume 11, Issue 3, September 2001, pages 101-122. DOI:10.1080/714004699

Myanmar

Schein J. Ethnofederalism and the accommodation of ethnic minorities in Burma: United They Stand // Calhoun. Institutional Archive of the Naval Postgraduate School. URL: <https://calhoun.nps.edu/handle/10945/34739> (accessed: 19.06.2016).

Netherlands

Migge, Bettina; Léglise, Isabelle; Bartens, Angela (2010). Creoles in Education: An Appraisal of Current Programs and Projects. Amsterdam: John Benjamins Publishing Company. p. 268.

LANDSVERORDENING van de 28ste maart 2007 houdende vaststelling van de officiële talen (Landsverordening officiële talen). Government of the Netherlands. URL:

<http://decentrale.regelgeving.overheid.nl/cvdr/XHTMLoutput/Actueel/Curaçao/144328.html> (accessed: 27.06.2016)

Over Curaçao. Government of Curaçao. URL:

<http://www.gobiernu.cw/web/site.nsf/web/overcura%C3%A7ao?opendocument&language=nederlands> (accessed: 27.06.2016)

Nicaragua

ROBERT F. ARNOVE. Adult Education and State Policy in Latin America: The contrasting cases of Mexico and Nicaragua. Journal: Comparative Education Volume 31, Issue 3, November 1995, pages 311-326. DOI:10.1080/03050069529001

Amanda Minks. Interculturalidad en el discurso de los niños miskitos en Corn Island.

Costa Atlántica con poco acceso a la educación // La prensa. 28/04/2002. URL:

<http://www.laprensa.com.ni/2002/04/28/nacionales/838301-costa-atlantica-con-poco-acceso-a-la-educacin> (accessed: 27.06.2016)

Programa de educación bilingüe intercultural (PEBI). URL:

http://www.montraykreyol.org/sites/default/files/breve_resena_de_la_experiencia_del_pebi_en_la_costa_caribe_nicaraguens

[e.pdf](#) (accessed: 27.06.2016)

Nigeria

Suberu R. Nigeria // Diversity and unity in federal countries / L. Moreno, C. Colino (eds.) – L., Ithaca: VcGrill – Queen's univ. press, 2010.

Olagbaju O. Multilingual Education in Nigeria: Policy, Practice, Challenges and Solutions // Journal of Education and Practice. 2014. №6.

Omoniyi A. The Languages in Nigerian Socio-political Domains: Features and Functions // English Language Teaching. 2012. № 10

Pakistan

Syed Abdul Manan, Maya Khemlani David, Francisco Perlas Dumanig. Disjunction between language policy and children's sociocultural ecology – an analysis of English-medium education policy in Pakistan. Journal: Language and Education Volume 29, Issue 5, September 2015, pages 453-473. DOI:10.1080/09500782.2015.1046882

Syed Abdul Manan, Francisco Perlas Dumanig, Maya Khemlani David. The English-medium fever in Pakistan: analyzing policy, perceptions and practices through additive bi/multilingual education lens. Journal: International Journal of Bilingual Education and Bilingualism pages 1-17. DOI:10.1080/13670050.2015.1080659

Sabiha Mansoor. The Status and Role of Regional Languages in Higher Education in Pakistan. Journal: Journal of Multilingual and Multicultural Development Volume 25, Issue 4, August 2004, pages 333-353. DOI:10.1080/01434630408666536

Syed Abdul Manan, Maya Khemlani David. Mapping ecology of literacies in educational setting: the case of local mother tongues vis-à-vis Urdu and English languages in Pakistan. Journal: Language and Education Volume 28, Issue 3, May 2014, pages 203-222. DOI:10.1080/09500782.2013.800550

Siddiqi Farhan Hanif. The Politics of Ethnicity in Pakistan: The Baloch, Sindhi and Mohajir Ethnic Movements. Routledge, 2012.

Panama

Educación Intercultural Bilingüe Ngobe y Kuna. URL: <http://consentimiento.culturalsurvival.org/educacion-intercultural-bilingue-ngobe-y-kuna> (accessed: 27.06.2016)

Papua New Guinea

Bougainville before the Conflict. Edited by Anthony J. Regan and Helga M. Griffin. Canberra, Pandanus Books, 2005.

Hutjena Secondary School // Papua New Guinea Education News. 6/28/2016. URL: <http://edu.pngfacts.com/secondary->

[schools-in-png/category/secondary-schools-in-bougainville](#) (accessed: 27.06.2016).

Philippines

Language Policy and Language-in-Education Policy in the Philippines. Annual Review of Applied Linguistics / Volume 2 / Март 1981, pp 48-59 DOI: <http://dx.doi.org/10.1017/S0267190500000246> (accessed: 27.06.2016).

Russia

Образование в России. 2003: Стат.сб./ Госкомстат России – М., 2003.

Алпатов В. Языковая ситуация в регионах современной России. [Электронный ресурс]. URL: <http://www.strana-oz.ru/2005/2/yazykovaya-situaciya-v-regionah-sovremennoy-rossii> (дата обращения: 10.05.2017).

Обзор: Изучение национальных языков в республиках и округах России. [Электронный ресурс]. URL: <http://nazaccent.ru/content/11552-89-rodnyh-yazykov.html> (дата обращения: 10.05.2017).

Как изучаются государственные и родные языки в России. [Электронный ресурс]. URL: <http://echomskufa.ru/2017/08/07/21639> (дата обращения: 10.05.2017).

ГОСУДАРСТВЕННЫЙ ЯЗЫК И КУЛЬТУРНОЕ МНОГООБРАЗИЕ РОССИЙСКОЙ ФЕДЕРАЦИИ.

АНАЛИТИЧЕСКИЙ ВЕСТНИК № 2. [Электронный ресурс]. URL: <http://iam.duma.gov.ru/node/8/4911#> (дата обращения: 10.05.2017).

Языки меньшинств: юридический статус и повседневные практики. Российско-французский диалог. Отв. ред. и пер. с фр. Елена Филиппова. – М.: ФГНУ "Росинформротех", 2013. 202 с.

Spain

Kathryn A. Woolard and Tae-Joong Gahng. Changing language policies and attitudes in autonomous Catalonia. Language in Society / Volume 19 / Issue 03 / Сентябрь 1990, pp 311 - 330 DOI: <http://dx.doi.org/10.1017/S0047404500014536>

Carla Amorós-Negre. Different paradigms in the history of Spanish language policy and planning. Journal: Journal of Multilingual and Multicultural Development pages 1-14. DOI:10.1080/01434632.2016.1165232

Catrin Wyn Edwards. Language-in-education policies, immigration and social cohesion in Catalonia: the case of Vic. Journal: International Journal of Bilingual Education and Bilingualism pages 1-16. DOI:10.1080/13670050.2015.1023253

Melinda Dooly, Virginia Unamuno. Multiple languages in one society: categorisations of language and social cohesion in policy and practice. Journal: Journal of Education Policy Volume 24, Issue 3, May 2009, pages 217-236.

DOI:10.1080/02680930902823039

Sudan

Language of Instruction for Increased Access to Relevant Education for Conflict-Affected Children in South Sudan. URL:

<http://www.iar-gwu.org/sites/default/files/articlepdfs/Language%20of%20Instruction%20in%20South%20Sudan-%20Hollyn%20Hammond.pdf> (accessed: 02.03.2016).

Switzerland

Gunther M Hega. Regional Identity, Language and Education Policy in Switzerland. Journal: Compare: A Journal of Comparative and International Education Volume 31, Issue 2, June 2001, pages 205-227. DOI:10.1080/03057920120053210

Tajikistan

Худоеров М.М. ПРОБЛЕМЫ РАЗВИТИЯ И СОХРАНЕНИЯ ПАМИРСКИХ ЯЗЫКОВ В СОВРЕМЕННОМ ТАДЖИКИСТАНЕ. - www.gramota.net/materials/3/2011/7-1/50.html (дата обращения: 06.07.2017).

Додыхудоева Л. Р. ПАМИРСКИЕ ЯЗЫКИ В УСЛОВИЯХ ЯЗЫКОВОГО СДВИГА // Языковая политика и языковые конфликты в современном мире. - М., 2014. С. 430-437

Tanzania

Eustard Tibategeza, Theodorus du Plessis. Language-in-education policy development in Tanzania: An overview. Journal: Language Matters Volume 43, Issue 2, November 2012, pages 184-201 DOI:10.1080/10228195.2011.573801

Ukraine

Эмирова А.М. Русско-крымскотатарская интерференция и проблемы возрождения миноритарных языков Крыма. // ВІСНИК Луганського національного педагогічного університету імені Тараса Шевченка. Філологічні науки. – 2007, № 18 (134). – С. 183–192. [Электронный ресурс] URL:

http://www.emirova.com/Emirova_Izbrannye_Nauchnye_Raboty.pdf (дата обращения: 16.06.2016).

Kulyk V. Language Policy in Ukraine What People Want the State to Do. - East European Politics and Societies and Cultures. 2013. Volume 27. Number 2, pp. 280-307.

United_Kingdom

Welsh Language Act 1993. [Электронный документ]. URL: <http://www.legislation.gov.uk/ukpga/1993/38/contents> (accessed: 8.10.2015).

Welsh Language (Wales) Measure 2011. [Электронный документ]. URL: <http://www.legislation.gov.uk/mwa/2011/1/enacted> (дата обращения 22.06.2016).

Gaelic Language Plan. Overview of the Gaelic Language (Scotland) Act 2005 and the preparation of Gaelic Language Plans [Электронный документ] – URL: <http://www.gov.scot/Publications/2010/07/06161418/3> (дата обращения 10.08.2016).

Pádraig Ó Riagáin, Glyn Williams, F. Xavier Vila i Moreno. Young People and Minority Languages: Language use outside the classroom (in Irish speaking schools). URL: <http://www.comhairle.org/english/wp->

[content/uploads/YoungPeopleAndMinorityLanguages_POR_2007.pdf](#) (accesses: 26.06.2016)

The role of Gaelic Education 3-18 in 'Language Learning in Scotland: A 1+2 Approach 2010. URL:

http://www.educationscotland.gov.uk/Images/RoleofGaelicEducationina12Approach_tcm4-826256.pdf (accesses: 22.06.2016)

The Scots language in education in Scotland // European Research Centre on Multilingualism and Language Learning. URL:

https://www.mercator-research.eu/fileadmin/mercator/documents/regional_dossiers/scots_in_scotland_2nd.pdf (accesses: 26.06.2016).

Terry Lamb. Language policy in multilingual UK. Journal: The Language Learning Journal Volume 23, Issue 1, June 2001, pages 4-12. DOI:10.1080/09571730185200031

Uzbekistan

Pearl of East: Your travel guide in Uzbekistan. URL: http://pearlofeast.com/post.php?id=443&news_page=34 (accessed: 26.06.2016).

7. Conflicts

Description of the variable	Label	Type	Sources	Comments
7. Conflicts				
7.1. Variables from ACD2EPR*				
Ethnic conflicts in the year of observation	incidence_flag	Binominal	ACD2EPR	(0) no (1) yes
Onset of the ethnic conflict in the year of observation	onset_ko_flag	Binominal	ACD2EPR	(0) no (1) yes
7.2. Variables on the basis of Conflict Barometer Annual Reports*				
Ethnic conflicts of the Type I in the year of observation	conflict_1	Binominal	Conflict Barometer Annual Reports	
Intensity of an ethnic conflict of the Type I in the year of observation	conflict_1_int	Ordinal	Conflict Barometer Annual Reports	
The issue of an ethnic conflict of the Type I in the year of observation <ul style="list-style-type: none"> • autonomy • secession 	conflict_1_aut conflict_1_sec	Two binominal	Conflict Barometer Annual Reports	
Start of the conflict	conflict_1_start	String	Conflict Barometer Annual Reports	
Conflict parties	conflict_1_parties	String	Conflict Barometer Annual Reports	
Ethnic conflicts of the Type II in the year of observation	conflict_2	Binominal	Conflict Barometer Annual Reports	
Intensity of an ethnic conflict of the Type II in the year of observation	conflict_2_int	Ordinal	Conflict Barometer Annual Reports	
The issue of an ethnic conflict of the Type II in the year of observation <ul style="list-style-type: none"> • autonomy • secession 	conflict_2_aut conflict_2_sec conflict_2_dom	Three binominal	Conflict Barometer Annual Reports	

• regional domination				
Start of the conflict	conflict_2_start	String	Conflict Barometer Annual Reports	
Conflict parties	conflict_2_parties	String	Conflict Barometer Annual Reports	
Ethnic conflicts in the year of observation	conflict	Binominal	Conflict Barometer Annual Reports	
Intensity of an ethnic conflict in the year of observation	conflict_int	Ordinal	Conflict Barometer Annual Reports	

Notes and comments

Variables from ACD2EPR

ACD2EPR 2014 is a docking dataset links conflicts inventoried in UCDP/PRIO Armed Conflict Dataset to politically relevant ethnic groups of EPR project (Wucherpfennig et al., 2012)⁹.

Operationally, the "armed conflict" in ACD2EPR is defined as: "a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in deaths." "battle-related deaths."

at least 25 battle-related deaths" (UCDP/PRIO Armed Conflict Dataset Codebook. - <http://ucdp.uu.se/downloads/ucdpprio/ucdp-prio-acd-172.pdf>)

The ERAD takes from ACD2EPR all the units of observations which refer to ERAs. Two variables are used:

- "incidence_flag" - presence / absence of ethnic conflicts in the autonomy in a given year
- "onset_ko_flag" - the emergence of new ethnic conflicts in the autonomy in a given year

Since the data for 2014-2015 are missing in ACD2EPR, these units of observations are coded as "NA".

⁹ <https://icr.ethz.ch/data/epr/acd2epr/>

Variables on the basis of Conflict Barometer Annual Reports

Because ethnopolitical conflicts are not necessarily armed, the ERAD uses materials from another project, the Conflict Barometer (Heidelberg Institute for International Conflict Research). In this project, a conflict is defined as follows: “Conflict measures are actions and communications carried out by a conflict actor in the context of a political conflict. They are constitutive for an identifiable conflict if they lie outside established procedures of conflict regulations and – possibly in conjunction with other measures – if they threaten the international order or a core function of the state” (Conflict Barometer Dataset 2015: 6).

Since information on conflicts in the Conflict Barometer is presented in the form of annual monitoring, all the reports for 2001-2015 were examined for the identification of inter-ethnic conflicts directly related to the ERAs included in the ERAD.

Selection of the conflicts related to the ERAs

- 1) Conflict Barometer distinguishes 4 types of the conflicts (interstate, intrastate, sub-state, and trans-state) and 10 items of the conflicts. The ERAB takes the conflicts, which a) types: intrastate and sub-state; b) items: autonomy, secession, sub-national domination.
- 2) Conflict Barometer defines the actors involved in the conflict and (in some cases) the territory of the conflict. The ERAD takes the conflicts, which: a) actors: ERA or ethnic group (titular for the ERA or ethnic minority of the ERA); b) take place in territory of the ERA.
- 3) All other conflicts are not taken in the ERAD, for instance
 - religious conflicts are not included if they are cross-regional (Nigeria: Christians versus Muslims, Northerners vs. Southerners, Boko Haram, Lord's Resistance Army, India: Muslims are Hindus, Christians are Hindus, Islamists are against the government, Pakistan: TTP, LeJ, SSP - various religious groups, Taliban - various tribes; Bosnia and Herzegovina: wahhabi militants);
 - inter-ethnic communal conflicts are not included, if they are cross-regional (Nigeria: farmers versus shepherds);
 - conflicts between the ERA's titular group and other actors are not included if they occur in another region where the representatives of this group live (China: Tibetans vs. Hui vs. Han);
 - autonomist / secessionist conflicts are not included if they are cross-regional and have no direct relationship to ERA (Nigeria: struggle for the creation of Biafra, Italy: Lega Nord / northern Italy);
 - subnational domination conflicts between different parties of the titular group are not included (Kurdistan: conflict between KDP and PUK)

- 4) On the basis of the actors involved in the conflict, it is defined whether it is a new or ongoing conflict.
- 5) If the conflict occurs in some ERAs, it is taken into account in each ERA separately.

Two types of the conflicts in ERAs

All the selected conflicts have been divided into 2 types (identified on the basis of items and conflict parties)

- 1) The struggle of titular groups for greater autonomy / secession: "titular ethnic group versus the central government" or "ERA (regional government) versus the central government". The issues of such type of conflict may be a) autonomy; b) secession;
- 2) The struggle of groups that have not yet autonomy, for autonomy or secession: "ethnic minority versus central / regional government" or "ethnic minority versus titular / dominant ethnic group". The issues of such type of conflict may be a) autonomy; b) secession; c) subnational domination¹⁰.

Coding rules

- 1) 5 variables are taken from the reports of Conflict Barometer:
 - “conflict” - presence - absence of the conflict in the unit of observation (autonomy – year) (binominal);
 - “conflict_int” - intensity of the conflict (ordinal):
 - 1 – dispute: ни одна из сторон не угрожает насилием
 - 2 – non-violent crisis: одна из сторон угрожает насилием
 - 3 – violent crisis
 - 4 – limited war
 - 5 – war
 - issue of the conflict (autonomy, secession, subnational domination) - some binominal variables;
 - “conflict_start” - the year of the start of the conflict (string);
 - “conflict_parties” - the actors involved in the conflict (string)

¹⁰ As a rule, subnational domination is defined as an additional item.

- 2) Both types of conflicts are coded separately - as "conflict_1" and "conflict_2" with the corresponding variables. If there is no conflict in the unit of observation, the variables "beginning of the conflict" and "parties to the conflict" are coded as "NA".
- 3) In addition, two aggregated variables have been developed (summarizing the conflicts of both types):
 - “conflict” - presence - absence of the conflict in the unit of observation (autonomy – year);
 - “conflict_int” - intensity of the conflict

References

- Anderson L. (2016) Ethnofederalism and the Management of Ethnic Conflict: Assessing the Alternatives // *Publius: The Journal of Federalism*. № 1.
- Benedikter T. (2009) *Solving Ethnic Conflict through Self-government: a Short Guide to Autonomy in South Asia and Europe*. Bolzano.
- Brancati D. (2008) The origins and strengths of regional parties // *British Journal of Political Science*. Vol. 38, № 1, pp. 135-159.
- Conflict Barometer (2015) Heidelberg: The Heidelberg Institute for International Conflict Research.
- Dandoy R. (2010) Ethno-regionalist parties in Europe: a typology // *Perspectives on Federalism*. Vol. № 2, p. 194-220.
- Dandoy R., Schakel A. (eds) (2013) *Regional and National Elections in Western Europe. Territoriality of the Vote in Thirteen Countries*. Houndmills.
- Elazar D. (1994) *Federal Systems of the World: a Handbook of Federal, Confederal and Autonomy Arrangements*. Harlow.
- Ganguly R., MacDuff I. (2003) *Ethnic Conflict and Secessionism in South and Southeast Asia: Causes, Dynamics, Solutions*. Thousand Oaks.
- Ghai Y. (2000) *Autonomy and Ethnicity: Negotiating Competing Claims in Multi-Ethnic States*. Cambridge.
- Ghai Y., Woodman S. (2013) *Practising Self-Government: a Comparative Study of Autonomous regions*. Cambridge.
- Hale H. (2004) Divided We Stand: Institutional Sources of Ethnofederal State Survival and Collapse // *World Politics*. № 1.
- Henders S. (2010) *Territoriality, Asymmetry, and Autonomy: Catalonia, Corsica, Hong Kong, and Tibet*. New York.
- Hooghe, L., Marks G., Schakel A., Chapman-Osterkatz S., Niedzwiecki S. Shair-Rosenfield S. (2016) *Measuring regional authority. Volume I: A postfunctionalist theory of governance*. Oxford.
- Horowitz D. (1985) *Ethnic Groups in Conflict*. Berkeley.
- Lapidoth R. (1997) *Autonomy: Flexible Solutions to Ethnic Conflicts*. Washington.
- Lijphart A. (1977) *Democracy in Plural Societies: A Comparative Exploration*. New Haven.
- O'Leary B. (2001) An Iron Law of Nationalism and Federation? A (Neo-Diceyan) Theory of the Necessity of a Federal Staatsvolk, and of Consociational Rescue // *Nations and nationalism*. № 3.
- Oziashvili Y. (2012) *Regional Party Systems in Ethnofederal States*. PhD dissertation. The City University of New York, 2012.
- Roeder P. (2014) Secessionism, Institutions, and Change // *Ethnopolitics*. № 1.
- Schakel A. (2013) Congruence Between Regional and National Elections // *Comparative Political Studies*. № 5.

- Schakel A., Swenden W. (2016) Rethinking Party System Nationalization in India (1952–2014) // Government and Opposition (Published online: 08 March 2016 - doi:10.1017/gov.2015.42)
- Wimmer A. (2013) Ethnic Boundary Making: Institutions, Power, Networks. New York.
- Wolff S. (2010) Approaches to Conflict Resolution in Divided Societies // Ethnopolitics Papers. № 5.
- Wucherpfennig, Julian, Nils W. Metternich, Lars-Erik Cederman, and Kristian Skrede Gleditsch (2012) Ethnicity, the State, and the Duration of Civil War // World Politics. № 1.

Appendix. The list of ethnic regional autonomies

Core list

Country	ERA	Titular ethnic group ¹¹	Dominant in the country ethnic group (Staatsvolk) ¹²
Belgium****	Flemish Region	Flemings	Flemings / Walloons
	Walloon Region	Walloons	Flemings / Walloons
Bosnia and Herzegovina****	Federation of Bosnia and Herzegovina	Bosniaks / Croats	Bosniaks / Croats / Serbs
	Republika Srpska	Serbs	Bosniaks / Croats / Serbs
Canada	Nunavut	Inuit people	English Canadians
	Quebec	French Canadians	English Canadians
China	Guangxi Zhuang	Zhuang people	Han
	Inner Mongolia	Mongols	Han
	Ningxia Hui	Hui people	Han
	Tibet	Tibetans	Han
	Xinjiang Uyghur	Uyghurs	Han
Denmark	Faroe Islands	Faroese people	Danes
	Greenland	Inuit people	Danes
Ethiopia****	Afar	Afars	Amhara / Tigrayans
	Amhara	Amhara people	Amhara / Tigrayans
	Benishangul-Gumuz	Berta / Gumuz	Amhara / Tigrayans
	Gambela	Nilotic peoples	Amhara / Tigrayans
	Harari	Harari people	Amhara / Tigrayans
	Oromia	Oromo people	Amhara / Tigrayans
	Somali	Somalis	Amhara / Tigrayans
	Southern Nations, Nationalities and Peoples Region	Southern Peoples	Amhara / Tigrayans
	Tigray	Tigrayans	Amhara / Tigrayans
Fiji	Rotuma	Rotumans	Fijians indigenous / Indo-Fijians
Finland	Aland Islands	Swedes	Finns
France*****	Corsica	Corsicans	French

¹¹ In 8 ERAs there are two or even three titular ethnic groups. Furthermore, in some cases, “the group of peoples” is interpreted as the titular ethnic group: in Ethiopia - Southern Nations, Nationalities and Peoples Region and Gambela (Nilotic peoples), Southern Peoples in South Sudan, Dagestan in Russia. As a result of internal heterogeneity of ethnic group, in some cases it is very hard to define whether the group should be considered as a “single ethnic group” or a “group of peoples” (Moro people in Philippines, Bouganvilleans in Papua New Guinea, Naga people in India, etc.).

¹² In 5 countries there are some dominant ethnic groups (the cases are described in Notes).

India *****	Andhra Pradesh	Telugu speakers	Hindi speakers
	Assam	Assamese	Hindi speakers
	Goa	Konkani speakers	Hindi speakers
	Gujarat	Gujaratis	Hindi speakers
	Jammu and Kashmir	Kashmiri Muslims	Hindi speakers
	Karnataka	Kannada speakers	Hindi speakers
	Kerala	Malyalam speakers	Hindi speakers
	Maharashtra	Marathi speakers	Hindi speakers
	Manipur	Manipuri (Meitei) people	Hindi speakers
	Meghalaya	Khasi / Garo / Jaintia	Hindi speakers
	Mizoram	Mizo people	Hindi speakers
	Nagaland	Naga people	Hindi speakers
	Orissa	Oriya people	Hindi speakers
	Punjab	Punjabi Sikhs	Hindi speakers
	Tamil Nadu	Tamils	Hindi speakers
	Telangana (since 2014) *	Urdu speaking Muslims	Hindi speakers
	Tripura	Bengalis	Hindi speakers
	West Bengal	Bengalis	Hindi speakers
Indonesia	Aceh (since 2005) *	Acehnese people	Javanese
	Papua *	Papuans	Javanese
Iraq	Kurdistan ** (since 2005) *	Kurds	Arabs
Italy	Aosta Valley	French speakers	Italians
	Friuli Venezia Giulia	Friulians	Italians
	Sardinia	Sardinians	Italians
	South Tyrol ***	South Tyrolians	Italians
Mauritius	Rodrigues (since 2002) *	Creoles of African descent	Indo-Mauritians
Moldova	Gagauzia	Gagauz people	Moldavians
Myanmar	Chin	Chin people	Burmese
	Kachin	Kachin people	Burmese
	Kayah	Karenni people (Red Karens)	Burmese
	Kayin	Karens	Burmese
	Mon	Mon people	Burmese
	Rakhine	Rakhine people	Burmese
	Shan	Shan people	Burmese
Nicaragua	North Caribbean Coast Autonomous Region	Indigenous peoples	Mestizos
Nicaragua	South Caribbean Coast Autonomous Region	Indigenous peoples	Mestizos
Nigeria *****	Adamawa	Fulbe	Hausa (Fulani) / Yoruba / Igbo

	Akwa Ibom	Ibibio people	Hausa (Fulani) / Yoruba / Igbo
	Bayelsa	Ijaw (Izon) people	Hausa (Fulani) / Yoruba / Igbo
	Benue	Tiv / Idoma	Hausa (Fulani) / Yoruba / Igbo
	Borno	Kanuri people	Hausa (Fulani) / Yoruba / Igbo
	Delta	Urhobo people	Hausa (Fulani) / Yoruba / Igbo
	Edo	Edo (Bini) people	Hausa (Fulani) / Yoruba / Igbo
	Kogi	Igala people	Hausa (Fulani) / Yoruba / Igbo
	Niger	Nupe people	Hausa (Fulani) / Yoruba / Igbo
	Rivers	Ogoni people	Hausa (Fulani) / Yoruba / Igbo
	Yobe	Kanuri people	Hausa (Fulani) / Yoruba / Igbo
Pakistan ****	Baluchistan	Baluchi	Punjabis
	Khyber Pakhtunkhwa	Pashtuns	Punjabis
	Punjab	Punjabis	Punjabis
	Sindh	Sindhis	Punjabis
Panama	Embera Wounaan	Indigenous Embera	Mestizos
	Guna Yala	Indigenous Kuna	Mestizos
	Ngabe Bugle	Indigenous Ngöbe	Mestizos
Philippines	Mindanao	Muslims Moro people	Tagalog people
Russia	Adygeya	Adyghe	Russians
	Agin-Buryat AD (until 2008) *	Buryats	Russians
	Altay	Altaians	Russians
	Bashkortostan	Bashkirs	Russians
	Buryatiya	Buryats	Russians
	Chechnya	Chechens	Russians
	Chukotka	Chukchi people	Russians
	Chuvashiya	Chuvash people	Russians
	Dagestan	Dagestani peoples	Russians
	Evenk AD (until 2007) *	Evenks	Russians
	Ingushetiya	Ingush people	Russians
	Jewish AO	Jews	Russians
	Kabardino-Balkariya	Kabardians / Balkars	Russians
	Kalmykiya	Kalmyks	Russians
	Karachaevo-Cherkessiya	Karachays / Circassians	Russians

	Kareliya	Karelians	Russians
	Khakasiya	Khakas people	Russians
	Khanty-Mansi AD	Khanty / Mansi	Russians
	Komi	Komi people	Russians
	Komi-Permyak AD (until 2005) *	Komi-Permyaks	Russians
	Koryak AD (until 2008) *	Koryaks	Russians
	Marii El	Mari people	Russians
	Mordoviya	Mordvins	Russians
	Nenets AD	Nenets	Russians
	North Ossetiya	Ossetians	Russians
	Tatarstan	Tatars	Russians
	Taymyr AD (until 2007) *	Dolgans / Nenets	Russians
	Tuva	Tuvinians	Russians
	Udmurtiya	Udmurt people	Russians
	Ust-Orda Buryat AD (until 2008) *	Buryats	Russians
	Yakutiya	Yakuts	Russians
	Yamalo-Nenets AD	Nenets	Russians
Spain ****	Basque Country	Basques	Spaniards
	Catalonia	Catalans	Spaniards
	Galicia	Galicians	Spaniards
	Navarre	Basques	Spaniards
Sudan	South Sudan ** (2005-2011) *	Southern Peoples	Arabs
Switzerland****	Jura	French speaking Jurassians	German Swiss
Tajikistan	Gorno-Badakhshan	Pamiris	Tajiks
Ukraine	Crimea (until 2014) *	Russians	Ukrainians
United Kingdom	Northern Ireland	Irish people	English
	Scotland	Scots	English
	Wales	Welsh	English
Uzbekistan	Karakalpakstan	Karakalpaks	Uzbeks

*Border-line list*¹³

France ****	French Polynesia	Polynesians	French
	New Caledonia	Kanakas	French
Georgia	Ajara	Ajarans	Georgians
India ****	Arunachal Pradesh	No definite titular group, the largest groups are Nyishi, Adi	Hindi speakers
	Sikkim	Nepali speakers	Hindi speakers
Italy	Sicily	Sicilians	Italians
Malaysia ****	Sabah	Kadazans	Malays
	Sarawak	Dayaks	Malays
Netherlands ****	Aruba	Creoles	Dutch
	Netherlands Antilles (Curacao) *	Creoles	Dutch
Nigeria ****	Cross River	No definite titular group, the largest groups are Efik, Ejagham	Hausa (Fulani) / Yoruba / Igbo
	Nasarawa	No definite titular group, the largest group is Eggon people	Hausa (Fulani) / Yoruba / Igbo
	Plateau	No definite titular group, the largest group is Berom people	Hausa (Fulani) / Yoruba / Igbo
	Taraba	No definite titular group, the largest groups are Mumuyes, Jukun, Tiv	Hausa (Fulani) / Yoruba / Igbo
Papua New Guinea ****	Bougainville	Bouganvilleans	Papuans
Tanzania ****	Zanzibar	Zanzibaris (Shirazi and arabs)	Mainland Africans

¹³ 16 autonomies are included in the “border-line list” because of two sources of their ambiguity. 1) Adjarians in Georgia, Sicilians in Italy, Zanzibaris in Tanzania and Bouganvilleans in Papua New Guinea might be considered as distinct ethnic groups, however, the strength of the ethnic grounds for the self-identification there remains unclear. In other words, all these autonomies are undoubtedly connected to regional identity, but it is hard to tell the salience of its ethnic component.; 2) all other cases and Zanzibar again - it is hard to define clearly to what extent ethnic character was a significant reason for the establishment of autonomous region.

Notes and comments

The list includes all sub-national units of the first level with special status in unitary states and entities of the federations, which correspond to criteria of ERA and exist at present or ceased to exist in the beginning of XXI century. Some countries (Nepal, South Sudan, Somali), where transition to the federation has not yet finished, are not examined. Because of the vagueness in constitutional foundations in some countries (Spain, Tanzania, South Africa, Iraq, etc.), there is no full agreement among scholars on the issue of how many countries are in fact fully federal states. Some disputed cases, which are considered as federations by some authors but not the others, are interpreted on the basis of their legal (constitutional) definitions. Therefore, Tanzania, Iraq, Sudan (2005-2011), and Myanmar (since 2008, when new Constitution was passed) are defined as federations, namely partial ethnofederations with seven ethnic regions in Myanmar and one per each remaining case¹⁴. On the other hand, for the decentralized states (regionalized states, quasi-federations) as Spain, Italy, and South Africa are defined as unitary states. Therefore only ethnic regions with special status are included in the list of ERAs in Spain and Italy, while none of regions of South Africa is included in the list because all of them have equal status.

* **The cases, which exist as ERA only part of the period 2001-2015.** Rodrigues was granted autonomy in 2002, Aceh – in 2005, autonomy of Kurdistan was reestablished in 2005. South Sudan existed as autonomous region in of Sudan in 2005-2011 and then became an independent state. A newest Indian state Telangana was established just in 2015. On the other hand, 6 autonomous okrugs of the Russian Federation were eliminated during administrative reforms in 2005-2008. In Crimea case, there was an irredentist secession in 2014.

Netherlands Antilles: Until 2010, Curaçao, Bonaire, Sint Maarten, Sint Eustatius and Saba constituted The Netherlands Antilles as one of the three self-governing entities of Kingdom of the Netherlands, so that they are presented in the list as a single ERA in spite of that in 2010 reorganization took place, and only two islands (Curaçao and Sint Maarten) get the status of self-governing entities.

Papua in Indonesia: in 2003, it was decided to divide it into two provinces with a special status, but this process is still incomplete, and Papua is presented in the list as a single region.

** **The first sub-national level:** In most cases administrative division of country is pretty clear, so it is not hard to distinguish sub-national unites of the first level. However, in some countries formally first sub-national division has rather administrative than political significance ('statistical regions' in Philippines, federal districts in Russia, etc.). It is expressed in lack of own government structure. In all these cases "actual" first sub-national level with own government structure was examined. Furthermore, there some special cases, where ERAs were established between the national and the first sub-national level - Iraqi Kurdistan, Zanzibar (Tanzania), and South Sudan (Sudan, 2005-2011). All of them are included in ERAD in contrast to territorial autonomous units, which were founded between the first and the second sub-national levels (for instance, Chittagong Hills in Bangladesh).

*** **South Tyrol:** it is the only second-level sub-national unit that for special reasons is included in ERAD. South Tyrol is a province (the second sub-national level), which is part of Trentino-Alto Adige, one of the five autonomous special regions of Italy. Trentino-Alto Adige is divided into two ethnic territories: the German-speaking South Tyrol and the Italian Trento (Trentino). Formation of such an artificial region was a consequence of complex historical constellations. After the Second World War, South Tyrol remained in Italy in exchange for guarantees of its special rights. The special status of South Tyrol and Trento as autonomous provinces is fixed constitutionally, and

¹⁴ From the view of distinguishing ERA, in all these cases it is not important whether they unitary states or federations, since all ethnic regions have a special status.

unlike all other regions of Italy, most of the powers of Trentino-Alto Adige are transferred to the provincial level in such a way that it is actually the provinces and not the region that have real self-government. Even official socio-economic statistics are not conducted for Trentino-Alto Adige, but for South Tyrol and Trento separately.

****** Titular and dominant ethnic groups in full ethno-federations:** In full ethno-federations (Belgium, Bosnia and Herzegovina, Pakistan, and Ethiopia), all the first tier sub-national entities are defined in ethnic terms. It follows that within the full-federations there is no clearly identified “dominant” ethnic group. Thus, in Belgium both Flemings and Walloons should be considered as “dominant” while each of them is concurrently a “titular group” in own region. Similarly, in Bosnia and Herzegovina three ethnic groups – Bosniaks, Croats, and Serbs are interpreted as “dominant”; Serbs is also a titular group in Republika Srpska; Bosniaks and Croats – in the Federation of Bosnia and Herzegovina. Unlike, in Ethiopia, for historical reasons only two ethnic groups – Amhara and Tigrayans – are usually identified as Staatsvolk; both of them also have their own regions, where they are titular ethnic groups. In Pakistan, Punjabis are usually considered as Staatsvolk though it is a disputable issue. At the same time, Punjabis have their own region.

******* Notes on some specific cases:**

India: It is common to consider Hindi-speaking and professing Hindu people as Staatsvolk. Accordingly, ERAs may be not Hindi-speaking states - 20 states out of 29¹⁵. Creation of most of them was the result of the reform of the administrative-political division of the country according to the linguistic principle (The States Reorganisation Act, 1956). In several cases, along with the language, religious specificity (Jammu and Kashmir, Punjab, Telangana) is important for ethnic boundaries. Several states were British colonial provinces, but some of them also had the character of ethnic peripheries in the British Empire, and after granting India independence, they were transformed into ERAs (West Bengal, Orissa, Jammu and Kashmir). Some others states – former colonial provinces were reorganized in accord with linguistic lines so that they became ethnic states (Assam, Tamil Nadu, Punjab, Karnataka). Two of the 20 states (Arunachal Pradesh, Sikkim) are included in the border-line list, as they appeared as a result of the annexation of some territories to India, and despite their ethnic specifics, there is no certainty that ethnicity was important factor for their creation.

Spain: of 19 autonomous regions (17 autonomous communities and 2 autonomous cities), only 4 - Catalonia, Galicia, Basque Country and Navarre - are interpreted as ERAs because they have a special status, which in turn is clearly related to their ethnic character. The process of regionalization in Spain began after the adoption of the democratic Constitution of 1978, and special track (fast track) of autonomization was used only for three “historical regions” (Catalonia, Basque Country, Galicia). This, in particular, was due to the fact that they had experience of obtaining self-government during the Second Republic. In addition, the Spanish legislation recognized the special historical status of Navarre - *fueros* (*fueros* - right, privilege).

Malaysia: The Malays are certainly a dominant ethnic group in the country; and only 2 of the 13 entities of the Malaysian Federation - Sarawak and Sabah have a special ethnic character. Nevertheless, the creation of Sarawak and Sabah as entities of the Malaysian Federation had a variety of reasons, and it is difficult to definitely say to what extent the ethnic specificity of the population (the presence of indigenous peoples, which have received a special status of *bumiputera* in exchange for their agreement with entering these English colonies into the Malaysian Federation) was essential

¹⁵ In addition, there are 7 union territories, but although some of them are related to ethnic specifics (the Andaman and Nicobar Islands, Lakshadweep), only 2 of the 7 territories have self-government (the rest are managed by the central government) - Delhi and Pondicherry, and both are not ethnic.

in the process of post-imperial transformation. Therefore, Sarawak and Sabah are included in the border-line list.

Nigeria: here three ethnic groups (Hausa, Igbo, and Yoruba) are commonly recognized as “the main ethnic groups” and can be considered as dominant groups in the country. In the onset of the Nigerian state three regions corresponding with these groups were founded so that initially, Nigeria was a full ethnofederations. However, after some administrative reforms, three entities were divided into 36 states so that each of three main ethnic groups dominates in some states while the rest of the regions are linked to different minority groups. This moves contemporary Nigeria to the partial ethnodederation; and due to this logic, all 15 states, where neither of three main ethnic groups dominates, can be interpreted as ERAs. While in some cases the creation of “minority states” was the result of respective ethnic movement, in other cases there is no clear link between the creation of the state and the ethnic group. Sometimes a state is divided so as to create an “ethnic state”, and it is not entirely obvious that the other state that emerged after the separation of the first state is ethnic. All these cases are included in the supplementary list.

Netherlands: The Kingdom of the Netherlands is a very specific state form, invented in 1954 in the process of post-imperial transformation. Until 2010, it consisted of three territories (countries): the Netherlands itself - a state in Western Europe, consisting of 12 provinces, and 2 countries in the Caribbean: the Netherlands Antilles and Aruba (until 1985, Aruba was also included in the Netherlands Antilles). Accordingly, they can be considered as ERAs, although since in both cases it is not entirely clear to what extent ethnic specificity was the basis for granting them special statuses, both these territories are included in the border-line list¹⁶. In 2010, the Netherlands Antilles were liquidated (see note *), and only two of the islands (Curaçao and Sint Maarten) received the status of self-governing countries similar to Aruba (Bonaire, Sint Eustatius and Saba had the status of special municipalities in the Netherlands, however, do not belong to any of the provinces).

Papua New Guinea: Bougainville became a part of the PNG as a result of the post-imperial transformation, but from the very beginning it was striving for self-determination. Many years of struggle led to a compromise and the granting of autonomy to the island. As in the country as a whole, the processes of ethnogenesis in Bougainville are far from over, about two dozen language groups stand out among the island's population. Nevertheless, in the struggle for self-determination, common Bougainville identity has been constructed. Sometimes, it is interpreted as ethnic, and Bougainvilleans are slightly different from the Papuans (residents of the main part of the country) in a racial sense (darker skin). If you follow this logic, Bougainvilleans can be qualified as a titular ethnic group of the ERA, while Papuans – as a dominant ethnic group of the country. However, since such an interpretation is not conventional, this case is included in the border-line list.

Tanzania: Zanzibar became a part of Tanzania as a result of the post-imperial transformation, when a federation was formed from the British colonies of Tanganyika and Zanzibar. Unlike the mainland, the history of Zanzibar is closely related to the Arabs, who in the Middle Ages made the island the center of the slave trade. The centuries-old contacts between the Arabs and the indigenous population led to the Islamization of the island's population, as well as the emergence of Shirazi as a special group of people: having African appearance, they do not have “their own language” but speak in Swahili (the official language of Tanzania) and profess Islam¹⁷. Thus, Zanzibaris is often perceived

¹⁶ In addition, the specificity of these ERAs is that they are not represented in the Netherlands parliament. The latter is the parliament of the Netherlands proper, and not of the Kingdom of the Netherlands, which does not have a parliament at all. However, both ERAs have special ministers (Minister Plenipotentiary), who represent them in the Dutch government in The Hague.

¹⁷ Benedikter T. World's Modern Autonomy Systems: Concepts and Experiences of Regional Territorial Autonomy. – Bozen/Bolzano, 2009. – P. – 159-160.

as an ethnic group, and if you follow this logic, they can be qualified as a titular ethnic group of the ERA, while “Africans of the mainland” (“Bantu people”) – as a dominant ethnic group of the country. However, since such an interpretation is not conventional, this case is included in the border-line list. Moreover, it is not obvious that it was ethnic specific that was the basis for granting autonomy to Zanzibar

France: It is a unitary state with a very complex administrative division: there are 18 regions in the country (until 2016 there were 27 regions) and 101 departments. Of all the metropolitan regions, only Corsica has a special status and corresponds with all criteria of ERA. In addition, there are 5 overseas regions, which, however, do not have a special status (Guadeloupe, Martinique, French Guiana, Reunion, Mayotte). Furthermore, there are 3 overseas units with a special status, but 2 of them (Clipperton and the French Southern and Antarctic territories) are managed directly from Paris, and only New Caledonia has autonomy. Finally, there is such a special type of administrative units as the “overseas communities” (Saint-Pierre and Miquelon; Saint Barthelemy; Saint-Martin; Wallis; Futuna; French Polynesia), which have expanded rights in comparison with the usual departments. However, of all of them only French Polynesia has a special status of “overseas country”. Thus, of all the overseas territories of France, New Caledonia and French Polynesia can be interpreted as ERAs, but since in both cases it is not entirely clear to what extent ethnic specificity was the basis for granting them special status, both these territories are included in the border-line list.

Switzerland represents a special case due to the specific way of the generation of the state. The confederation of the cantons arose in the nineteenth century, evidently not on the basis of ethnicity. For this reason it is very hard to interpret Switzerland as an ethnic federation. Despite the fact that the 26 cantons are fairly strictly divided into three zones according to their languages (German, French and Italian, and only three cantons are “mixed”, Franconian), none of these groups can be regarded as having its own “own ERA” in the composition of Switzerland. There are different interpretations of which group should be considered as Staatsvolk in Switzerland - only German-speaking Swiss, or all 3 ethno-linguistic groups, but in any logic the Swiss cantons do not fit the concept of ERA. The only exception is the canton of Jura, which was created in 1979 as a result of a social movement that had ethnic features (language / religion) and aimed at the exit of the French-speaking and predominantly Catholic population from the canton of Bern (mostly German-speaking and Protestant). Therefore, Jura is qualified as ERA, and French speaking Jurassians - as a titular ethnic group of the ERA.