Pro-Government Militias Database Version 2.0 Codebook

Sabine C. Carey and Neil J. Mitchell

October 29, 2021

Project website and email:

http://www.sabinecarey.com/militias

militias@uni-mannheim.de

When using the data, please cite:

Carey, Sabine C., and Neil J. Mitchell. 2021. Pro-Government Militias Database (PGMD) Version 2.0 Codebook, University of Mannheim.

Carey, Sabine C., Neil J. Mitchell and Katrin Paula. 2022. "The Life, Death and Diversity of Pro-Government Militias: The Fully Revised Pro-Government Militias Database Version 2.0". *Research & Politics* 9(1), https://doi.org/10.1177/20531680211062772.

Contents

1 Acknowledgements **8**

2 Introduction **8**

3 Sources **9**

3.1 News sources . . . 9

3.2 Web of Science . . . 9

3.3 Amnesty International and Human Rights Watch reports . . . 10

3.4 Wikipedia . . . 10

4 Structure of the database **10**

4.1 Online PGMD 2.0 . . . 10

4.2 PGMD 2.0 Group Data file `pgmdv2_group.dta` . . . 10

4.3 PGMD 2.0 Country-Year Data file `pgmdv2_countryyear.dta` . . . 10

4.4 Militias Guidebook . . . 11

5 Coding Procedure **12**

6 Variables **12**

6.1 Basic information . . . 12
- 6.1.1 Country name (`country`) . . . 12
- 6.1.2 Country code (`gwno`) . . . 12
- 6.1.3 PGM name (`name`) . . . 12
- 6.1.4 PGM id (`id`) . . . 12
- 6.1.5 Name assigned by coder (`name_assigned`) . . . 12
- 6.1.6 Private Military Company (`pmc`) . . . 12

6.2 Start and end dates of the PGM . . . 13
- 6.2.1 Date formed (`date_formed`) . . . 13
- 6.2.2 Day formed (`day_formed`) . . . 13
- 6.2.3 Month formed (`month_formed`) . . . 13
- 6.2.4 Year formed (`year_formed`) . . . 13
- 6.2.5 Accuracy of formation date (`accuracy_formed`) . . . 13
- 6.2.6 Date terminated (`date_terminated`) . . . 14
- 6.2.7 Day terminated (`day_terminated`) . . . 14
- 6.2.8 Month terminated (`month_terminated`) . . . 14
- 6.2.9 Year terminated (`year_terminated`) . . . 14
- 6.2.10 Accuracy of termination date (`accuracy_terminated`) . . . 15
- 6.2.11 Year terminated, alternative measure . . . 15
- 6.2.12 Last recorded activity (`last_activity`) . . . 15

6.3 Life cycle of the group pre- and post-PGM . . . 15
- 6.3.1 Predecessor's PGM name (`predecessor01, predecessor02`) . . . 15
- 6.3.2 Predecessor's PGM id (`predecessor_id01, predecessor_id02`) . . . 15
- 6.3.3 Former group (`prior_group`) . . . 15

6.3.4 Former armed group (prior_armed) . . . 16
6.3.5 Former rebel group (prior_rebel) . . . 16
6.3.6 UCDP actor ID of prior rebel group (prior_ucdp) . . . 16
6.3.7 Predecessor's PGM name (predecessor) . . . 16
6.3.8 Predecessor's PGM id (predecessor_id) . . . 16
6.3.9 Successor's PGM name (successor) . . . 17
6.3.10 Successor's PGM id (successor_id) . . . 17
6.3.11 Subsequent rebel group (post_rebel) . . . 17
6.3.12 UCDP actor ID of subsequent rebel group (post_ucdp) . . . 17

6.4 Origin of the PGM . . . 17
6.4.1 Created by government (gov_formed) . . . 17
6.4.2 Created by a state institution (govcreator_state_institution) . . . 17
6.4.3 Created by an individual (govcreator_person) . . . 18
6.4.4 Created by a governing political party (govcreator_party) . . . 18
6.4.5 Created by the military (govcreator_military) . . . 18
6.4.6 Created by a sub-national government (govcreator_subnational) . . . 18
6.4.7 Created by other (govcreator_other) . . . 18
6.4.8 Creating body unclear (govcreator_unclear) . . . 18
6.4.9 Origin (origin) . . . 18

6.5 Termination . . . 18
6.5.1 Disarmed (termtype_disarmed) . . . 18
6.5.2 Government change (termtype_gov_change) . . . 18
6.5.3 Border change (termtype_border_change) . . . 19
6.5.4 PGM defects (termtype_pgm_defect) . . . 19
6.5.5 Government defects (termtype_gov_defect) . . . 19
6.5.6 PGM integrated into regular forces (termtype_integrate) . . . 19
6.5.7 Termination type unclear (termtype_unclear) . . . 19
6.5.8 Termination (termination) . . . 19

6.6 Connection to the government . . . 19
6.6.1 Government relation (government_relation) . . . 19
6.6.2 Link to a state institution (link_state_institution) . . . 20
6.6.3 Link to an individual (link_person) . . . 20
6.6.4 Link to a governing political party (link_party) . . . 20
6.6.5 Name of party link (party_id) . . . 20
6.6.6 Link to the military (link_military) . . . 21
6.6.7 Link to a sub-national government (link_subnational_gov) . . . 21
6.6.8 Link is unclear (link_unclear) . . . 21
6.6.9 Weapons and training (weapon_training) . . . 21
6.6.10 Government trained or equipped (gov_trained_equipped) . . . 21
6.6.11 Shared information or shared operations (shared_info_ops) . . . 21
6.6.12 Shared personnel (shared_personnel) . . . 22

6.7 PGM organizational characteristics . . . 22
6.7.1 Location (location) . . . 22
6.7.2 Highest maximum headcount (headcount_high) . . . 22
6.7.3 Lowest maximum headcount (headcount_low) . . . 22
6.7.4 Size (size) . . . 22
6.7.5 Organization (organization) . . . 22

6.8 Membership . . . 22
6.8.1 Primary membership (primary_mem) . . . 23

6.8.2 Alternative primary membership (alt_primary_mem) . . . 24
6.8.3 Members identifiable by ethnicity (member_ethnic) . . . 24
6.8.4 Ethnicity of membership (ethnicmember01–ethnicmember04) . . . 24
6.8.5 EPR id of ethnic membership (cowgroupid_m01–cowgroupid_m04) . . . 24
6.8.6 Quality of information for ethnic membership (info_ethnic_mem) . . . 25
6.8.7 Members identifiable by ideology (member_ideology) . . . 25
6.8.8 Members identifiable by religion (member_religious) . . . 25
6.8.9 Members identifiable as political party activists (member_party_activist) . . . 25
6.8.10 Members identifiable as villagers/from rural areas (member_village_rural) . . . 25
6.8.11 Members from urban areas (member_urban) . . . 25
6.8.12 Members belonging to the security forces (member_security_forces) . . . 25
6.8.13 Members include former members of security forces (member_ex_soldier) . . . 25
6.8.14 Members include former rebels (member_former_rebel) . . . 26
6.8.15 Members include mercenaries (member_mercenary) . . . 26
6.8.16 Members identifiable as workers (member_worker) . . . 26
6.8.17 Members identifiable as peasants (member_peasant) . . . 26
6.8.18 Members include students (member_student) . . . 26
6.8.19 Members include adolescents (member_adolescent) . . . 26
6.8.20 Members include children (member_children) . . . 26
6.8.21 Members with foreign nationalities (member_foreign_national) . . . 26
6.8.22 Members include unemployed (member_unemployed) . . . 26
6.8.23 Members include criminals (member_criminals) . . . 26
6.8.24 Members coerced (members_coerced) . . . 26
6.8.25 Members paid (members_paid) . . . 27
6.8.26 Members killed (members_killed) . . . 27
6.8.27 Reasons for membership (reason_membership . . . 27

6.9 Purpose . . . 27
6.9.1 Protect institution (purp_protect_institutions) . . . 27
6.9.2 Protect leader (purp_protect_leader) . . . 27
6.9.3 Protect border (purp_protect_border) . . . 28
6.9.4 Protect property (purp_protect_property) . . . 28
6.9.5 Self-defense and security (purp_protect_security) . . . 28
6.9.6 Intelligence (purp_intelligence) . . . 28
6.9.7 Intimidate civilians (purp_intimidate_civilian) . . . 28
6.9.8 Intimidate ethnic or religious groups (purp_intimidate_ethnic_religious) . . . 28
6.9.9 Name of intimidated ethnic group (ethnicpurpose01–ethnicpurpose03) . . . 28
6.9.10 EPR id of intimidated ethnic group (cowgroupid_p01–cowgroupid_p03) . . . 28
6.9.11 Quality of information for intimidated ethnic group (info_ethnic_purp) . . . 28
6.9.12 Intimidate political opposition (purp_intimidate_opposition) . . . 29
6.9.13 Fight insurgents (purp_fight_insurgents) . . . 29
6.9.14 Fight crime (purp_anti_crime) . . . 29
6.9.15 Destabilise neighbouring countries (purp_destabilise_neighbour) . . . 29
6.9.16 Purpose (purpose_text) . . . 29

6.10 Relative benefit . . . 29
6.10.1 Knowledge (rel_benefit_knowledge) . . . 29
6.10.2 Local presence (rel_benefit_local_presence) . . . 29
6.10.3 Local support (rel_benefit_local_support) . . . 29
6.10.4 Faster mobilization (rel_benefit_faster_mobil) . . . 30
6.10.5 Cheaper mobilization (rel_benefit_cheaper_mobil) . . . 30
6.10.6 Deniability (rel_benefit_deniability) . . . 30
6.10.7 Coup-proofing (rel_benefit_coup_proof) . . . 30

6.10.8 Violence (rel_benefit_use_violence) . . . 30
6.10.9 Loyalty (rel_benefit_loyalty) . . . 30

6.11 Target . . . 30
6.11.1 Target unarmed government critics (target_govcritic) . . . 30
6.11.2 Target armed group (target_rebel) . . . 30
6.11.3 Target religious group (target_religious) . . . 31
6.11.4 Target ethnic group (target_ethnic) . . . 31
6.11.5 Ethnic target name (ethnictarget01–ethnictarget03) . . . 31
6.11.6 EPR id of targeted ethnic group (cowgroupid_t01–cowgroupid_t03) . . . 31
6.11.7 Target students (target_student) . . . 31
6.11.8 Quality of information for targeted ethnic group (info_ethnic_targ) . . . 31
6.11.9 Target peasants (target_peasant) . . . 31
6.11.10 Target journalists (target_journalist) . . . 31
6.11.11 Target foreigners (target_foreigner) . . . 32
6.11.12 Target foreigners (target_aidworker) . . . 32
6.11.13 Target immigrants(target_immigrant) . . . 32
6.11.14 Target criminals (target_criminal) . . . 32
6.11.15 Target PGMs (target_pgm) . . . 32
6.11.16 Target military (target_military) . . . 32
6.11.17 Target foreign military (target_foreign_military) . . . 32
6.11.18 Target peacekeepers (target_pko) . . . 32
6.11.19 Target civilians (target_civilian) . . . 32
6.11.20 Target unclear (target_unclear) . . . 32
6.11.21 Treatment of civilians (treatment_civilians) . . . 32
6.11.22 Type of violence: beating (violencetype_beat) . . . 33
6.11.23 Type of violence: kidnapping, abductions (violencetype_kidnap) . . . 33
6.11.24 Type of violence: torture (violencetype_torture) . . . 33
6.11.25 Type of violence: sexual violence (violencetype_sexual) . . . 33
6.11.26 Type of violence: killings (violencetype_kill) . . . 33

6.12 Types of support . . . 33
6.12.1 Support from the domestic government (support_gov) . . . 34
6.12.2 Support from the military (support_military) . . . 34
6.12.3 Support from business corporations (support_corporation) . . . 34
6.12.4 Support from landowners (support_land) . . . 34
6.12.5 Support from villagers (support_village) . . . 34
6.12.6 Support from a foreign government (support_foreign_gov) . . . 34
6.12.7 Foreign government sponsor (sponsor_NAMEOFCOUNTRY) . . . 34
6.12.8 Support from a foreign actor (support_foreign) . . . 34
6.12.9 Support through drugs (support_drugs) . . . 34
6.12.10 Support through plunder and loot (support_loot) . . . 34
6.12.11 Support through crime (support_crime) . . . 34
6.12.12 PGM being self-maintained (support_self) . . . 35
6.12.13 Support unclear (support_unclear) . . . 35
6.12.14 Other connections (other_connections) . . . 35

6.13 PGM activities (activities) . . . 35

6.14 Other (other) . . . 35

7 References for meta-analysis **35**

8 Pieces of evidence **35**

9 Interrupted PGMs **35**

9.1 Afghanistan . . . 36
9.2 Bangladesh . . . 36
9.3 Congo, Democratic Republic of . . . 36
9.4 Georgia . . . 36
9.5 Haiti . . . 37
9.6 Iraq . . . 37
9.7 Indonesia . . . 37
9.8 Libya . . . 37
9.9 Malawi . . . 37
9.10 Nepal . . . 38
9.11 Pakistan . . . 38
9.12 Philippines . . . 38
9.13 Somalia . . . 38
9.14 Sudan . . . 38
9.15 Zimbabwe . . . 38

10 Country notes **39**

10.1 Afghanistan . . . 39
10.2 Burundi . . . 39
10.3 Brazil . . . 39
10.4 Chile . . . 39
10.5 Colombia . . . 39
10.6 Comoros Islands . . . 39
10.7 Cote d'Ivoire . . . 40
10.8 Croatia . . . 40
10.9 Egypt . . . 40
10.10 Ethiopia . . . 41
10.11 Indonesia . . . 41
10.12 Iraq . . . 41
10.13 Kenya . . . 41

10.14 Mexico . . . 42
10.15 Mozambique . . . 42
10.16 Myanmar . . . 42
10.17 Nepal . . . 42
10.18 Nigeria . . . 42
10.19 Philippines . . . 42
10.20 Romania . . . 44
10.21 Serbia . . . 44
10.22 Kosovo . . . 44
10.23 Somalia . . . 44
10.24 Sudan . . . 45
10.25 South Africa . . . 45
10.26 Syria . . . 45
10.27 Uganda . . . 45
10.28 Ukraine . . . 45
10.29 United Kingdom . . . 46
10.30 United States of America . . . 46
10.31 Venezuela . . . 46
10.32 Yugoslavia . . . 46
10.33 Zambia . . . 46
10.34 Zimbabwe . . . 46

References **47**

1 Acknowledgements

The PGMD 2.0 was funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement 336019. The first version of the PGMD was funded by the Economic and Social Research Council (ESRC), UK, RES-062-23-0363, with additional support from the Working Group "Human Rights, Governance and Conflict" at the Centre for the Study of Civil War (CSCW) at the Peace Research Institute Oslo (PRIO). We are particularly grateful to our excellent research assistants Patricia Blank, Anna-Lena Hadbawnik, Paula Jöst, Stella Riedel

2 Introduction

The Pro-Government Militias Database (PGMD) v2.0 contains information about over 500 organised armed groups that are identified by documentary, media and academic sources to be pro-government but are not part of regular state security forces. Consistency across sources is checked, but no further claims are made about the accuracy or completeness of this information from these sources. For some groups there is substantial reporting and information, for others less information is available and there are more missing data.

The information covers all countries during the time period from 1981 to 2014. Is not restricted to on-going conflicts. The main unit of observation is the pro-government militia (PGM). Other labels used for such groups are "militias", "paramilitary groups", "irregular forces", "death squads" or "vigilantes." The dataset includes private military companies, which are also labelled ""military contractor", "private security company" or 'private military firm," if they satisfy the definition of pro-government militias outlined below and were identified in the search process.

Operational definition of a pro-government militia

We define a pro-government militia as a group that

1. is identified by the source as pro-government or sponsored by the government (national or sub-national),
2. is identified as not part of the regular security forces,
3. is armed and
4. has some level of organisation.

Criterion 1

The group is identified by the source as pro-government or sponsored by the government, either national or sub-national. Evidence that a group does not oppose the government or that the group shares an enemy with the government is not sufficient for it to be considered pro-government.

Our focus is on national governments and the individual ministerial or institutional links between the government and the militia, but if a group is identified as allied with or backed by a sub-national government and the national (e.g. federal) government is neutral in its position towards this group, it is included as a PGM. However, if sources describe the national (e.g. federal) government as opposed to this group, then it is not classified as a PGM.

We focus on domestic groups, meaning pro-government groups that are active within their own borders. Groups can be sponsored or trained by other countries, but they need to be linked to the government of the territory in which they are active. For private military companies, this means that they are recorded

as being present in the country they are operating in, i.e. the country for which they contracted to operate on the government's side.

Criterion 2

The group is not the regular state security force as described in government information or other sources on armed forces and police, although the group may operate with the regular state security forces, or even be composed of members of the security forces organized clandestinely as an unofficial or informal group (death squads). This relationship with the regular forces might include, in addition to sharing of personnel as in the above example, information sharing, joint operations or training. Not identified as PGM are therefore the police, regular military forces or special units that are part of the military. Paramilitaries do not enter the PGM dataset when the term paramilitary is used to refer to regular forces, i.e. police units with military status, as in the case of the French gendarmes or Italian caribinieri.

Criterion 3

The group is armed and equipped for violence, but does not have to commit violence to be included. This criterion is not limited to firearms; some groups are equipped with machetes or clubs.

Criterion 4

The group has some evidence of organisation (for example an identifiable leader, or name, or a geographical, ethnic, religious or political basis) to exclude a "flash" or spontaneous mob.

3 Sources

To gather information on pro-government armed groups, we use news sources, academic publications and NGO reports to cover a wide range of information and to triangulate and cross-check details provided by the individual sources. With this approach, we aim to provide comprehensive, reliable and valid coverage of these groups and their characteristics. While some types of sources include English translations of reports originally published in other languages, we rely predominantly on English language reports and publications.

Using different types of sources sources reduces the problem of missing data and of wrongly classifying the groups, but we expect that the PGMD will likely under-represent the presence of pro-government militias. Sources might also under-report the features of the group that we aim to capture with this dataset. Therefore, we do not implement a threshold in the form of number of reports to code particular characteristics. See specific variables for more details.

3.1 News sources

The PGMD contains open source information about pro-government armed groups obtained from LexisNexis searches of news sources from around the world. The sources include transcripts and reports that were initially published in their local language and translated into English to be included in BBC World Summaries of local news, Agence France Presse, Xinhua General News Service, and major international newspapers. Search terms included "government militia", "paramilitary", "government death squads", "government irregular forces" and "vigilante". While this search strategy is heavily labour-intensive, it reduced the risk of missing information.

3.2 Web of Science

We use the Web of Science to locate relevant academic sources on specific pro-government militias and then apply our coding procedures in a meta-analysis to the documents. We apply a PGM-level search by using the name of the PGM as search term, and a country-level search approach with the search terms

"militia*" together with the country name. We use the publications abstract to decide whether the article likely contains relevant information for our meta-analysis.

3.3 Amnesty International and Human Rights Watch reports

We search the online reports from Amnesty International and Human Rights Watch, using only the names of the PGMs in combination with the country name.

3.4 Wikipedia

We use a cautious approach to search information provided on Wikipedia on pro-government militias. We only use the names of the PGMs as search terms, in combination with the country name. If information about the PGMs is not backed up by a source, we ignore it. If the information about the PGM is referenced with a source, we go directly to the source and check whether the information can indeed by verified by the source. If we cannot find in the original source the information the Wikipedia site listed, we ignore the information. If we can verify it, we use it to inform our coding and reference the source and the link to the permanent Wikipedia site. If we are unable to access the source, for example due to language difficulties, we only use the information if the source seems to be of good quality. As this is difficult to assess, we apply a cautionary and conservative approach.

4 Structure of the database

The PGMD 2.0 provides extensive information on pro-government militias from around the globe between 1981 and 2014. The information is presented in various formats, which can be accessed from the website http://www.sabinecarey.com/militias.

4.1 Online PGMD 2.0

The most comprehensive version of the PGMD is the online relational database. It allows access to all tables that code various aspects of the pro-government militias. It also links each group to the main documentary evidence researchers used to create the dataset.

The main unit of analysis of the PGMD 2.0 is the pro-government militia. Therefore, all variables (with the exception of `activities`) are time-invariant. Because information on PGMs is not available in regular and narrow time intervals, it is not feasible to code changes in the characteristics over time in a consistent and reliable manner for all groups. If we observe significant temporal changes in the behaviour or characteristics of a group, they are described in the variable `other`.

4.2 PGMD 2.0 Group Data file `pgmdv2_group.dta`

This static and downloadable data file of the PGMD contains all time-invariant variables on each group. Like in the online version, its unit of analysis is the pro-government militia. Unlike the online version, it does not include the documentary evidence used for coding the groups.

4.3 PGMD 2.0 Country-Year Data file `pgmdv2_countryyear.dta`

This static and downloadable datafile uses the country-year as the unit of analysis. It transforms selected variables from the `pgmdv2_group.dta` into a country-year format. For the transformation we use the variable `year_terminated_alt` to identify the year the PGM is coded as terminated.

Note that it accounts for interrupted PGMs, i.e. groups that cease to exist as PGMs for a certain time period and then become a PGM again at a later time point.

The variables in this dataset are either binary measures that indicate whether at least one PGM existed in that country-year with the respective characteristic (`variablename`), or count measures (`variablename_count`), that reflect how many PGMs were present in that country-year with the respective characteristics.

The following variables are included in the `pgmdv2_countryyear.dta` dataset. For simplicity, we only list the binary PGM variables. The count variables have the same name but with the suffix `_count`:

- `gwno`: country code
- `year`
- `country name`
- `presence`: presence of PGMs
- `presence_informal`: informal PGMs
- `presence_semiofficial`: semi-official PGMs
- `presence_gov_formed`: PGMs formed by government
- `presence_notgov_formed`: PGMs not formed by government
- `presence_gov_formed_on`: onset of PGMs formed by government
- `link_person`: PGMs with a link to a person
- `link_party`: PGMs with a link to a political party
- `link_state_institution`: PGMs with a link to an institution
- `link_subnational_gov`: PGMs with a link to a sub-national institution
- `link_military`: PGMs with a link to the military
- `primary_mem_ethnic`: PGMs with ethnicity as primary or alternative primary membership
- `primary_mem_ideol`: PGMs with ideology as primary or alternative primary membership
- `primary_mem_local`: PGMs with local primary or alternative primary membership
- `primary_mem_nat`: PGMs with nationalist primary or alternative primary membership
- `primary_mem_nonciv`: PGMs with non-civilian primary or alternative primary membership
- `primary_mem_pol`: PGMs with political primary or alternative primary membership
- `primary_mem_rel`: PGMs with religious primary or alternative primary membership

Important note: Those who want to create their own country-year dataset from the `pgmdv2_group.dta` dataset they need to make sure that they adequately control for the interrupted PGMs and for the exclusion of Lebanon prior to 2008 and of Somalia prior to 2012. The file `2_createdatasets_extended.do` can be used as a template to transform the PGM group dataset `pgmdv2_group.dta` into the country-year format.

4.4 Militias Guidebook

The online Militias Guidebook at https://militias-guidebook.com describes the basic characteristics of each group in text form. It includes details on the formation of the group, its purpose, organization, size, weapons and training, as well as reasons for membership and treatment of civilians. It also lists the names of the sources used for the coding. If it is only based on news sources, it refers to the PGMD. It is used additional sources, such as academic studies or reports from NGOs, then this are referenced separately.

5 Coding Procedure

Each country is coded by one coder. The coding of the variables is then checked against the collected evidence by a second coder. The PI reviews the recommendations made by the second and makes the final decision in case of divergent coding. Therefore, the coding for each group is reviewed by three different individuals.

6 Variables

6.1 Basic information

6.1.1 Country name (`country`)

The country name identifies the country where the PGM is active and to whose government the PGM is linked. If a PGM operates in a disputed border region that is subject to competing claims by more than one country, then this region is treated as being part of the country that the PGM is from and claims sovereignty over.

6.1.2 Country code (`gwno`)

The country code for the country in which the PGM is (primarily) active and linked to the government of that country is identified with the country code based on Gleditsch and Ward (1999).

6.1.3 PGM name (`name`)

This variable lists the name(s) used in the source to identify the PGM or the name of the private military company (PMC) where a group can be identified as such. Repeated non-specific mentions of militias in a country are included under one PGM if there are no indications that the source is referring to a new or separate militia.

6.1.4 PGM id (`id`)

This variable provides a unique identifier for each PGM.

6.1.5 Name assigned by coder (`name_assigned`)

If the PGM is unnamed, the coders have assigned a descriptive name, coded `1` and `0` otherwise.

6.1.6 Private Military Company (`pmc`)

This variable identifies whether the group is a private military company or a private security company. If the group is identified as a PMC, then all recorded information about this group refers to the PMC's activities in that particular country. For example, if a report lists year 1990 as the year in which the PMC was formed, but lists year 1995 as the year in which the PMC became active in the country of interest, then we record year 1995 as date of formation. The variable takes on the value `1` for not being a PGM and `1` if it is a PGM. For PMCs we code the link to the government with the variable `government_relation` as semi-official and `members_paid` as paid, unless the available information suggests otherwise.

6.2 Start and end dates of the PGM

We provide an exact date for when the PGM was formed or when the group was first mentioned or when its contract started. The group might have been newly created on this date or an already existing group might have become a militia as defined in our criteria outlined above. For example, if a rebel or opposition armed group becomes pro-government because of a change in government, then this variable codes the date of becoming pro-government, not the earlier date in which the group formed as a rebel organisation.

In some cases it might not be clear who the government is for a particular time period. To identify the government, which is necessary to be able to identify an armed group as being pro-government, whenever possible we rely on the coding decisions and case description provided by Goemans, Gleditsch and Chiozza (2009).

The information on the formation of the PGM is recorded as precisely as possible. If a source refers to a range of time instead of a specific time point (for example referring to “the 1980s” instead of a specific year), then the earliest time point of the given range is coded (e.g. “the 1990s” is coded as 1990, etc).

Some groups become reactivated after having been terminated. This could be the result of a temporary change in government, side-switching of the group, or simply an interruption in the link to the government, for example. As each group has only one observation to avoid artificially counting the same group twice, even if the group existed at two distinct time periods, we indicate such interrupted PGMs in the relevant text fields, e.g. "Details of Formation", Details of Termination", or "Other". We list all such interrupted PGMs in the Section Interrupted PGMs.

6.2.1 Date formed (`date_formed`)

`date_formed` presents the formation date of the PGM in the format YYYY-MM-DD.

6.2.2 Day formed (`day_formed`)

`day_formed` lists the day the group was formed or first mentioned, listing the first of the month if the sources do not provide a specific day.

6.2.3 Month formed (`month_formed`)

`month_formed` lists the month the group was formed or first mentioned, listing the month of January (`1`) if the sources do not provide a specific month.

6.2.4 Year formed (`year_formed`)

`year_formed` lists the year the group was formed or first mentioned.

6.2.5 Accuracy of formation date (`accuracy_formed`)

As we code an exact date for formation of the group, this variable indicates the level of detail this date is based upon. `accuracy_formed` codes whether the sources reported the exact day (`dy`), the month (`mo`), the year (`yr`) or the decade (`dec`). If the date formed refers to the date of the first mention of the group rather than the date it was actually formed, then degree of accuracy is coded as the year. Year is the most common level of accuracy in the data.

6.2.6 Date terminated (`date_terminated`)

`date_terminated` presents the date the PGM is coded as terminated in the format YYYY-MM-DD.

A PGM is coded as being terminated when it no longer fulfils any of the above identified criteria. This includes the disarmament or banning of the group by the government or its integration into the regular security forces.

If a president or party ceases to be in government, then the PGMs affiliated with them also cease to be PGMs. If regular forces, such as the police, take action against or are in conflict with the group and if members are imprisoned for their group activities, we no longer classify the group as being pro-government, unless there is evidence of continued support by the executive (e.g. coup prevention).

Groups can also cease to be classified as PGMs as a result of a border change. For example, armed groups linked to the Indonesian government were active in East Timor, fighting the independence movement. These pro-Indonesia groups end in our dataset with the establishment of the UN Transitional Administration in East Timor created by the UN Security Council in October 1999. Even though some of these groups were still active within East Timor after this date, they were then linked to the Indonesian government. Therefore, they were not active in their own border anymore and no longer fit our definition of a domestic pro-government militia.

If the leader of a PGM changes and the group still operates under the same name, we do not code this PGM as terminated. The same is true, when the leader of a PGM creates a new splinter group, while the original group continues to operate under the same name but with a new leader, as in the case of the Sudan People's Liberation Army – United. The group splintered in 1994 and its leader Riek Machar renamed his small group of supporters. Lam Akol broke with Riek Machar and became the new chairman of SPLM/A-United.

In some cases it might not be clear who or which group forms the government for a particular time period. To identify the government, which is necessary to be able to identify an armed group as being pro-government, we rely whenever possible on the coding decisions and case description provided by Goemans, Gleditsch and Chiozza (2009).

The information on the termination of the PGM is recorded as precisely as possible. If a source refers to a range of time instead of a specific time point, then the last time point of the given range is coded, unless the sources suggest another more likely date. For example, if a group is said to have disarmed in January, without any further specification, we code it terminated on January 31. If a group is said to have terminated in 1992, we code it as terminated on Dec 31, 1992.

6.2.7 Day terminated (`day_terminated`)

`day_terminated` lists the day the group ceased to exist as PGM, listing the last day of the month if the sources do not provide a specific day.

6.2.8 Month terminated (`month_terminated`)

`month_terminated` lists the month the PGM is coded as terminated, listing the month of December `(12)` if the sources do not provide a specific month.

6.2.9 Year terminated (`year_terminated`)

`year_terminated` lists the year the PGM is coded as terminated.

6.2.10 Accuracy of termination date (accuracy_terminated)

As we code an exact date for termination of a PGM, this variable indicates the level of detail this date is based upon. This variable codes whether the sources reported the exact day (dy), the month (mo), the year (yr) or the decade (dec). Day is the most common level of accuracy in the data.

6.2.11 Year terminated, alternative measure

(year_terminated_alt) The sources do not always provide information about the PGM that indicates that the group no longer fulfills our criteria for being coded as pro-government militia, such as disarmament or change in government, and can therefore be coded as terminated. Sometimes the reports about a group simply end without any further details about the group. This variable provides an alternative measure for PGM termination by using information from the variable last_activity. If no clear indication of a PGM's termination is available, this variables uses the last year of reported activity as year of termination, if this year is prior to 2010. If the last year of activity falls within the last five years covered by the dataset (2010-2014), then it is not used as alternative termination date. In this case, the termination date remains missing.

6.2.12 Last recorded activity (last_activity)

This variable lists the year of the last recorded activity of the PGM.

6.3 Life cycle of the group pre- and post-PGM

Several variables identify whether, and if so in what form, the PGM existed as a group before and after it met all criteria to be coded as a pro-government militia. For example, The could have existed as unarmed group before becoming an armed PGM. Examples are the Malawi Pioneers, which were initially set up to work on agricultural development, and armed grass roots self-defence group, like the Rondas Campesinas in Peru. Some PGMs were previously rebel groups and might have become pro-government when the rebel groups joined the government. Or PGMs might turn rebel groups if their government is ousted from power. In some cases, PGMs merge, change from informal to semi-official, or simply change names.

6.3.1 Predecessor's PGM name (predecessor01, predecessor02)

If a PGM has a predecessor PGM, then this variable provides the name of the PGM that preceded the coded PGM. The PGM might have changed its name or the PGM could have transformed into a new PGM. PGMs have at most two PGM predecessors, their names are captured with the variables predecessor01 and predecessor02 respectively.

6.3.2 Predecessor's PGM id (predecessor_id01, predecessor_id02)

If a PGM has a predecessor PGM, then this variable provides the unique id of the PGM that preceded the coded PGM. predecessor_id01 is the PGM id for the variable predecessor01 and predecessor_id02 is the PGM id for the PGM listed in predecessor02.

6.3.3 Former group (prior_group)

This variable assesses whether the PGM existed as an organized group prior to becoming a PGM according to the PGM definition laid out above. It takes on the following values:

1 no

2 yes

3 unclear

6.3.4 Former armed group (prior_armed)

This variable assesses whether the PGM existed as an organized armed group prior to becoming a PGM according to the PGM definition laid out above. It can take on the value 2 (yes) only, if prior_group is coded 2 (yes). The variable takes on the following values:

. missing if prior_group == no

1 no

2 yes

3 unclear

6.3.5 Former rebel group (prior_rebel)

This variable assesses whether the PGM existed as an armed rebel group group prior to becoming a PGM according to the PGM definition laid out above. If the sources do not mention a rebel group that preceded the PGM, we code the absence of such information as indication that the PGM did not exist as a rebel group before becoming a PGM, unless we have reason to assume the contrary. The variable takes on the following values:

. missing if prior_group & prior_armed == no

1 no

2 yes

3 unclear

6.3.6 UCDP actor ID of prior rebel group (prior_ucdp)

If the PGM previously existed as a rebel group and was coded as a UCDP actor, this field records the ID from the UCDP Actor Dataset, version 20.1 (Pettersson and Öberg 2020). No PGM is associated with more than two preceding UCDP rebel groups. The variables prior_ucdp01 and prior_ucdp02 list the UCDP actor code for the first and second UCDP rebel group respectively that preceded the PGM.

6.3.7 Predecessor's PGM name (predecessor)

In some cases a PGM undergoes a name change or the PGM itself is transformed into a new PGM. If a PGM has a predecessor PGM, then this variable provides the name of the PGM that preceded the coded PGM. PGMs have at most two PGM predecessor, their names are captured with the variables predecessor01 and predecessor02 respectively.

6.3.8 Predecessor's PGM id (predecessor_id)

If a PGM has a predecessor PGM, then this variable provides the unique id of the PGM that preceded the coded PGM. predecessor_id01 is the PGM id for the variable predecessor01 and predecessor_id02 is the PGM id for the PGM listed in spredecessor02.

6.3.9 Successor's PGM name (successor)

In some cases a PGM undergoes a name change or the PGM itself is transformed into a new PGM. If a PGM has a successor PGM, then this variable provides the name of the PGM that succeeded the coded PGM. PGMs have at most two PGM successors, their names are captured with the variables successor01 and successor02 respectively.

6.3.10 Successor's PGM id (successor_id)

If a PGM has a successor PGM, then this variable provides the unique id of the PGM that succeeded the coded PGM. successor_id01 is the PGM id for the variable successor01 and successor_id02 is the PGM id for the PGM listed in successor02.

6.3.11 Subsequent rebel group (post_rebel)

This variable assesses whether the PGM turned into a rebel group. If the sources do not mention a rebel group that forms from the PGM, we code the absence of such information as indication that the PGM did not turn into a rebel group, unless we have reason to assume the contrary. The variable takes on the following values:

1 no

2 yes

3 unclear

6.3.12 UCDP actor ID of subsequent rebel group (post_ucdp)

If the PGM turned into rebel group and was coded as a UCDP actor, this field records the ID from the UCDP Actor Dataset, version 20.1 (Pettersson and Öberg 2020). No PGM is associated with more than two subsequent UCDP rebel groups. The variables post_ucdp01 and post_ucdp02 list the UCDP actor code for the first and second UCDP rebel group that the PGM turned into.

6.4 Origin of the PGM

6.4.1 Created by government (gov_formed)

This variable indicates whether the PGM was created by the government. The absence of any information that suggests that the group was created by the government is coded as not being created by government. The variable takes on the following values:

1 no

2 yes

3 unclear

If the PGM is coded as having been created by the government, the next set of variables identify which element of the government formed the PGM.

6.4.2 Created by a state institution (govcreator_state_institution)

This binary variable is coded 1 if a state institution that is part of or controlled by the government created the PGM.

6.4.3 Created by an individual (govcreator_person)

This binary variable is coded 1 if a particular individual who belongs to the government created the PGM.

6.4.4 Created by a governing political party (govcreator_party)

This binary variable is coded 1 if a political party who is part of the government created the PGM.

6.4.5 Created by the military (govcreator_military)

This binary variable is coded 1 if the military created the PGM.

6.4.6 Created by a sub-national government (govcreator_subnational)

This binary variable is coded 1 if a sub-national government created the PGM.

6.4.7 Created by other (govcreator_other)

This binary variable is coded 1 if another government body created the PGM.

6.4.8 Creating body unclear (govcreator_unclear)

This binary variable is coded 1 if it is unclear which government body created the PGM.

6.4.9 Origin (origin)

This text field describes the details of the group's formation as a PGM.

6.5 Termination

We identify different, non-exclusive reasons for why the PGM was coded as terminated, if such information is provided by the sources, with several binary variables.

6.5.1 Disarmed (termtype_disarmed)

This binary variable indicates whether a PGM is coded as terminated because it was disarmed. If only a sub-section of the PGM is said to have been disarmed, then the group is not coded as terminated.

6.5.2 Government change (termtype_gov_change)

If the government changes and there is no information that suggests that the PGM has established a link to the new government, then the PGM counts as terminated because of a change in government, captured with this binary variable.

6.5.3 Border change (termtype_border_change)

The binary variable border_change indicates whether due to a border change a PGM is no longer pro-government. By ending up in a different country, the group might still be linked to the previous government, which is now located in a different country. We code the PGM as terminated because of a border change unless information suggests that the PGM is linked to the government in the country it is located after the border change has occurred.

6.5.4 PGM defects (termtype_pgm_defect)

This binary variable codes whether the PGM defects from the government by cutting ties to the government.

6.5.5 Government defects (termtype_gov_defect)

This binary variable codes whether the government defects from the PGM by cutting the ties to the PGM, for example by declaring the group to be illegal. This is also the case if a higher-order government body (e.g. national executive) defects from the group while a lower-order government body (e.g. state executive) continues to support the PGM.

6.5.6 PGM integrated into regular forces (termtype_integrate)

This binary variable codes whether a PGM has been terminated by having been integrated into the regular security forces, such as military or police. If only part of the PGM has been integrated into the regular forces, while other parts remain active as PGM, the PGM is not coded as terminated.

6.5.7 Termination type unclear (termtype_unclear)

This binary variable indicates that although the group has been terminated, the information provided by the sources on the termination is not sufficiently clear to code any of the variables on the termination type.

6.5.8 Termination (termination)

This text field that describes the details of the group's termination as a PGM.

6.6 Connection to the government

Several variables capture the nature of the connection between the government and the pro-government militia. We capture whether the relationship is of informal or semi-official character and identify with four non-exclusive categories to which element in the government or to which institution the PGM is linked to. Additional variables describe the nature of the collaboration between the government and the PGM.

6.6.1 Government relation (government_relation)

This variable describes the relationship that the PGM has with the government. While all groups in the database are pro-government, they vary with respect to the degree of formality and visibility of the link. In some instances the government tries to keep the group at arm's length, while in other cases

governments openly include these groups in legislation, reports and official documents. As an effort to capture this type of proximity to the government we use two categories, *informal* and *semi-official PGMs*. The key difference between informal and semi-official PGMs is that the link to informal PGMs is not officially or formally acknowledged.

`government_relation` takes on three values:

- Value `1` identifies an *informal* link between the government and the militia. In these cases, the sources describe the group as pro-government, government militia, linked to the government, government-backed, or government-allied. The PGM may be armed or trained by the government. Death squads, even when closely linked to the government, are normally informal and clandestine, and are therefore categorized as informal PGMs. Examples include the Janjaweed in Sudan, the Young Patriots in Cote d'Ivoire, the Ansar-e Hezbollah in Iran and the Interahamwe Militia in Rwanda during the early 1990s.
- Value `2` identifies a *semi-official* link between the government and the militia. In this case, the PGM has a formally and/or legally acknowledged status, in contrast to the looser affiliation of informal PGMs. A semi-official PGM might be sub-ordinate to the regular security forces, but is separate from the regular police and security forces. The link between the PGM and the government is more formal and institutionalized, for example when the PGM is recognized and acknowledged in official and legal documents of the state. The group is still identified as a distinct organisation from the regular security forces as with Village Defence Committees in India.
- Value `0` identifies a group for which the link to the government is *unclear*. This coding is used for groups that do not fit the differentiation between informal and semi-official as outlined above. For example, the Gladio groups found in most European countries during the Cold War were formally organized and trained within NATO, but these groups were kept secret, including from most parliaments in Europe. Therefore, they cut across the categories *informal* and *semi-official*. This coding is also used for groups for which the available information does not allow for a classification on this dimensions, as is the case for several groups in Syria during the Syrian civil war.

6.6.2 Link to a state institution (`link_state_institution`)

This binary variable is coded `1` if the PGM is linked to an institution of the state, which is part of or controlled by the government, such as the executive or the defense ministry. In this case, the group's primary connection is not to the person holding the office, but to the office or institution itself.

6.6.3 Link to an individual (`link_person`)

This binary variable is coded `1` if the PGM is linked to a specific person that is part of the government, such as the president or a minister. For example, this might be the case if a PGM was created directly by the president.

6.6.4 Link to a governing political party (`link_party`)

This binary variable is coded `1` if the PGM is linked to a political party who is part of the government.

6.6.5 Name of party link (`party_id`)

If the PGM is linked to a political party, then `party_id` lists the name of that party.

6.6.6 Link to the military (`link_military`)

This binary variable is coded `1` if the PGM is linked to the military. If a PGM is linked to the institution and the state is a military regime, then we code the link as being both to state (institution) and to military (institution). If the military is openly in opposition to the government, then a connection between the military and an armed group does not make this group a pro-government militia.

6.6.7 Link to a sub-national government (`link_subnational_gov`)

This binary variable is coded `1` if the PGM is linked to a sub-national government, such as a state in a federal system. If the national government opposes a PGM that is linked to a sub-national government, the group does not enter the database as it is not considered to be pro-government.

6.6.8 Link is unclear (`link_unclear`)

This binary variable is coded `1` if the information on how the PGM is linked to the government does not match the link-variables listed above or where the sources are too vague to allow a more precise coding.

6.6.9 Weapons and training (`weapon_training`)

This text field describes the kind of weapons the PGM received and who trained them.

6.6.10 Government trained or equipped (`gov_trained_equipped`)

We record whether the PGM has been trained by government forces **or** has received equipment, such as arms or provisions, from the government. **NOTE** that this variable does, therefore, not only code training by the government. This variable does **not** distinguish whether the PGM received training **or** equipment from the government. This variable **can not** be used to measure one of these two elements separately. The absence of any information that the PGM was trained/or equipped by the government is interpreted as no training or equipment by the government. The variable takes on the following values:

`1 no` (i.e. the government provided neither equipment nor training)

`2 yes` (i.e. the government provided either only equipment, only training, or both equipment and training)

`3 unclear`

6.6.11 Shared information or shared operations (`shared_info_ops`)

We record whether the PGM has received information from the government related to the PGM's mission or whether the PGM has carried out operations jointly with the government, for example with regular police or army units. Information sharing captures the government sharing information with and passing information on to the group. If a group passes information on to the government (which the government might not have requested nor potentially have wanted to have), without the government sharing information, then this does not classify as information sharing. The absence of any information that the PGM shared information or operations with the government is interpreted as not sharing information or operations. The variable takes on the following values:

`1 no`

`2 yes`

`3 unclear`

6.6.12 Shared personnel (shared_personnel)

We record whether members of the security forces are also active members of the PGM. Examples of such personnel overlap are instances where soldiers out of uniform participate in death squads. he absence of any information that members of regular security forces were also active members in the PGM is interpreted as not having shared personnel. This variable takes on the following values:

1 no

2 yes

3 unclear

6.7 PGM organizational characteristics

6.7.1 Location (location)

This text field records the main location of PGM activities as precisely as the sources allow.

6.7.2 Highest maximum headcount (headcount_high)

We record the estimated size, in terms of membership, of the group as reported in the sources. This variable provides the largest estimated size of the group. If only one figure is given by the source, then we record this figure as the maximum. If the PGM is a PMC, force strength refers to the staff of the PMC that were active in the particular country, not the strength of the PMC as a whole.

6.7.3 Lowest maximum headcount (headcount_low)

If different sources provide different estimates of the PGM's size in terms of membership, then we record the lowest maximum in this variable. If the PGM is a PMC, force strength refers to the staff of the PMC that were active in the particular country, not the strength of the PMC as a whole.

Note: Since the data are time-invariant, we capture estimates of force strength when the group was at its largest, not variations over time, such as estimates of smaller sizes that only refer to an early time period, for example. Both variables, headcount_high and headcount_low aim to capture the size of the group at the height of the PGM. The variables reflect variation in what different sources report as the size of the group, not variation across time. Changes over time are captured in the text field size.

6.7.4 Size (size)

This text field describes the force strength of the group, including how it might have changed over time or reports discrepancies in reporting.

6.7.5 Organization (organization)

This text field records information about the organization of the group, including information about leadership and chain of command, as far as provided by the sources.

6.8 Membership

Several variables describe characteristics of the individuals that participate in PGMs. The two variables primary_mem and alt_primary_mem are nominal variables and identify the *main characteristic* that

distinguishes the majority of PGM members. In contrast, the set of binary variables `members_**` record whether sources identify *some members* as have this particular characteristics, *irrespective of the importance* of a particular characteristic relative to the whole group.

6.8.1 Primary membership (`primary_mem`)

This nominal variable captures the main feature that defines and characterizes the group's members. Information on the purpose or the behaviour of the group is not used to infer membership characteristic. For example, if a PGM is created by a political party or incumbent to intimidate opposition voters during elections, then the primary membership of this PGM is not automatically political.

If a PGM belongs to an umbrella PGM, for which primary membership can be clearly identified, we code the same primary membership for this individual PGM if there is no further information about primary membership of this PGM or no details that would suggest otherwise. The same counts the other way around. If we have information on primary membership for individual PGMs, which are grouped together in an umbrella PGM, for which we have no information about primary membership being different, we use the same primary membership for the umbrella PGM as well (see the Pro-Integration Fighters PPI from Indonesia as an example).

If members of one (previous) PGM also belong to or move to another PGM, we do not automatically code the same membership of this new group as for the old one. For example, the first PGM might be characterized by ethnicity but consists also of non-civilians. Some of the non-civilian members might be mobilized for a second PGM. If they are recruited due to their experience in violence rather than due to their ethnicity, the primary membership of the second PGM would be non-civilian and not ethnic like in the first PGM.

The variable `primary_mem` has the following values:

> `no information (0)` The sources do not provide evidence on what the main and defining characteristic is of the PGM's membership.
>
> `ethnic (1)`: The defining feature of the PGM is that their members are recruited along ethnic lines. This does not imply that they are recruited from only one ethnic group. An example is the Interahamwe Militia in Rwanda.
>
> `ideological (2)`: The defining feature of the PGM members is that they share a common ideological identity or a common ideological goal, such as the Revolutionary Committees in Libya.
>
> `local (3)`: The defining feature of the PGM members is that they are civilians that are recruited, stationed and active within their local communities. PGMs with primary local membership are deployed in the same areas they are recruited from. If the evidence suggests that they are not active within their own communities, then primary membership is not coded as local. If there is no information on deployment but only information on PGM members being locally recruited, we assume that they are also primarily locally deployed and code primary membership as local. Similarly, if the sources refer to local deployment within a specific town or region, without further specifying where the members come from, we assume that they are locally recruited, unless there are indicators that would suggest otherwise. Examples include the Edo State Vigilante from Nigeria.
>
> `nationalist (4)`: The defining feature of the PGM members is that they share a common nationalistic goal. An example of nationalist membership is the Patriotic League in Bosnia-Herzegovina or several militias in Indonesia fighting against the independence of East Timor.
>
> `non-civilian (5)`: The defining feature of non-civilian PGM members is that they are off-duty militias or are drawn from parallel forces. They might be regular police or military if they act outside

their formal organization, or former members of the regular security apparatus and mercenaries. PMCs are coded as having non-civilian members, as they consisted of paid mercenaries. Examples include the Ninjas in Algeria and Indonesia or the Commando Rodrigo Franco in Peru.

`political (6)`: The defining feature of the PGM members is that they share a common political goal or political identity. Examples include PGMs that are formed by political parties and where their members share a common political goal. Examples are the Young Pioneers or the Young Democrats in Malawi.

`religious (7)`: The defining feature of the PGM is that their members are recruited along religious lines. An example is the Ismaili Militia in Afghanistan.

6.8.2 Alternative primary membership (`alt_primary_mem`)

Some PGMs have two equally important defining features of their membership, where one cannot clearly rank one above the other. In these cases, we identify an alternative primary membership, using the same categories as for the variable `primary_mem`. This variable is **only** filled in, if it seems to distort the presentation of the PGM membership if only one primary category was selected. It does **not** provide a ranking. If for one PGM both primary membership and alternative primary membership are coded, then both are judged to be of equal value. Examples are the Awakening groups in Iraq, with primary membership being coded as ethnic and religious, and the Anti-Maoist vigilantes in Nepal, for which primary membership is identified as local and ideological. For most PGMs, this variable is left blank.

6.8.3 Members identifiable by ethnicity (`member_ethnic`)

If the sources identify some members of the PGM by a particular ethnicity, this binary variable takes on the value `1`.

6.8.4 Ethnicity of membership (`ethnicmember01-ethnicmember04`)

If `member_ethnic` is coded `1`, we identify the name of the ethnic group as given by the Ethnic Power Relations (EPR) core dataset 2019 (Vogt et al. 2015). PGM members are identified with a maximum of four different ethnicities in the PGMD Version 2.0. Therefore, the ethnicity of the PGM membership is captured with four variables, `ethnicmember01, ethnicmember02, ethnicmember03, ethnicmember01`. The variables record the name of the ethnic group if such ethnic group membership can be identified and is coded `0` otherwise.

If the group given by our sources is mentioned as a sub-group of an EPR-identified group, we list the EPR group as the main group and provide the name of the sub-group in the variable `other`. If we cannot identify a group, we use the Ethnolog website to try to match the name of the ethnic group. If our sources do not provide a specific name for the ethnic group, we try to establish a link via other connections, such as the leader of the group.

6.8.5 EPR id of ethnic membership (`cowgroupid_m01-cowgroupid_m04`)

Whenever we can connect the ethnicity of PGM members to an ethnic group listed in the Ethnic Power Relations (EPR) core dataset 2019 (Vogt et al. 2015), these four variables contain the cowgroupid from the EPR dataset. `Cowgroupid_0*` is the code for the ethnic group name in `ethnicmember0*`.

6.8.6 Quality of information for ethnic membership (info_ethnic_mem)

If the PGM is recorded as having ethnically based membership (member_ethnic = 1 and the variable cowgroupid_m01–cowgroupid_m04 identifies the ethnic group, we code whether the information about the ethnic group was given by the PGMD sources or inferred through a link, such as the leader of the group. This variable takes on the following values:

given (1)

inferred (2)

not applicable (3)

6.8.7 Members identifiable by ideology (member_ideology)

If the sources identify some members of the PGM by a particular ideology, this binary variable takes on the value 1.

6.8.8 Members identifiable by religion (member_religious)

If the sources identify some members of the PGM by a particular religion, this binary variable takes on the value 1.

6.8.9 Members identifiable as political party activists (member_party_activist)

If sources identify some members of the PGM as activists of a political party that forms part of the government, this binary variable takes on the value 1.

6.8.10 Members identifiable as villagers/from rural areas (member_village_rural)

If sources identify some members of the PGM as villagers or coming from rural areas, this binary variable takes on the value 1.

6.8.11 Members from urban areas (member_urban)

If sources identify some members of the PGM as coming from urban areas, this binary variable takes on the value 1.

6.8.12 Members belonging to the security forces (member_security_forces)

If sources identify some members of the PGM as belonging to the regular security forces, such as military or police, this binary variable takes on the value 1.

6.8.13 Members include former members of security forces (member_ex_soldier)

If sources identify some members of the PGM as having previously belonged to the formal security apparatus, this binary variable takes on the value 1. Rarely it might include former members of the security forces of another country. In those instances the variable member_foreign_national is also coded 1.

6.8.14 Members include former rebels (member_former_rebel)

If sources identify some members of the PGM as former rebels, this binary variable takes on the value 1.

6.8.15 Members include mercenaries (member_mercenary)

If sources identify some members of the PGM as mercenaries, this binary variable takes on the value 1.

6.8.16 Members identifiable as workers (member_worker)

If sources identify some members of the PGM as workers, this binary variable takes on the value 1.

6.8.17 Members identifiable as peasants (member_peasant)

If sources identify some members of the PGM as peasants, this binary variable takes on the value 1.

6.8.18 Members include students (member_student)

If sources identify some members of the PGM as students, this binary variable takes on the value 1.

6.8.19 Members include adolescents (member_adolescent)

If sources identify some members of the PGM as adolescents, this binary variable takes on the value 1. We do not attach a specific age range to this variable but rely on the description provided by the sources.

6.8.20 Members include children (member_children)

If sources identify some members of the PGM as children, this binary variable takes on the value 1. We do not attach a specific age range to this variable but rely on the description provided by the sources.

6.8.21 Members with foreign nationalities (member_foreign_national)

If sources identify some members of the PGM as having a foreign nationality, this binary variable takes on the value 1.

6.8.22 Members include unemployed (member_unemployed)

If sources identify some members of the PGM as unemployed, this binary variable takes on the value 1.

6.8.23 Members include criminals (member_criminals)

If sources identify some members of the PGM as criminals, this binary variable takes on the value 1.

6.8.24 Members coerced (members_coerced)

This nominal variable indicates whether (some) members of the PGM were forced to join the group. If children are recruited, this variable is always coded yes. Absence of relevant information is coded as no coercion. The variable takes on the following values:

1 no

2 yes

3 unclear

6.8.25 Members paid (`members_paid`)

This variable codes whether members of the PGM received payment as PGM members. It is coded as yes also if only some members receive monetary payment. Absence of relevant information is coded as no payment. The variable takes on the following values:

1 no

2 yes

3 unclear

6.8.26 Members killed (`members_killed`)

This ordered variable captures the frequency with which PGM members were killed, irrespective of the perpetrator of the killings. It broadly measures the frequency of the casualties, not the extent or intensity of the killings of PGM members. If a large number of PGM members were killed in one or two events, the coding would be "rarely". If a very small number of PGM members are killed frequently, then the coding would be "often". This variable takes on the following values:

. unclear

0 no information

1 never

2 rarely

3 sometimes

4 often

6.8.27 Reasons for membership (`reason_membership`

This text field summarises the reasons for why individuals are reported to have joined the PGM.

6.9 Purpose

We record the intended purpose of the PGM as provided by the sources. This might not necessarily be identical with the main activities of the group. We identify the stated purpose of the group with twelve binary variables and an open text field.

6.9.1 Protect institution (`purp_protect_institutions`)

If the purpose of the PGM is reported as protecting national or sub-national political institutions, this variable is coded `1` and `0` otherwise.

6.9.2 Protect leader (`purp_protect_leader`)

If the purpose of the PGM is reported as protecting the political national or sub-national leader (the individual, not the office), this variable is coded `1` and `0` otherwise.

6.9.3 Protect border (purp_protect_border)

If the purpose of the PGM is reported as protecting the borders of the country, this variable is coded 1 and 0 otherwise.

6.9.4 Protect property (purp_protect_property)

If the purpose of the PGM is reported as protecting property, this variable is coded 1 and 0 otherwise.

6.9.5 Self-defense and security (purp_protect_security)

If the purpose of the PGM is reported as self-defense and providing security, this variable is coded 1 and 0 otherwise.

6.9.6 Intelligence (purp_intelligence)

If the purpose of the PGM is reported as gathering intelligence and providing information, this variable is coded 1 and 0 otherwise.

6.9.7 Intimidate civilians (purp_intimidate_civilian)

If the purpose of the PGM is reported to intimidate unarmed civilians, this variable is coded 1 and 0 otherwise.

6.9.8 Intimidate ethnic or religious groups (purp_intimidate_ethnic_religious)

If the purpose of the PGM is reported to intimidate ethnic or religious groups, this variable is coded 1 and 0 otherwise.

6.9.9 Name of intimidated ethnic group (ethnicpurpose01-ethnicpurpose03)

Whenever the purpose of a PGM is coded as intimidating an ethnic group, we record the name of this intimidated ethnic group, using information from the Ethnic Power Relations (EPR) core dataset 2019 (Vogt et al. 2015). The potentially three different ethnic groups that are intimidated are captured with these three variables.

6.9.10 EPR id of intimidated ethnic group (cowgroupid_p01-cowgroupid_p03)

These three variables code the cowgroupid from the Ethnic Power Relations (EPR) core dataset 2019 (Vogt et al. 2015) for the ethnic group names in ethnicpurpose0*.

6.9.11 Quality of information for intimidated ethnic group (info_ethnic_purp)

If the PGM is recorded as having ethnically based membership and we identify the ethnic group, this variable codes whether the name of the ethnic group was given by the PGMD sources or inferred through a link, such as the leader of the group. This variable takes on the following values:

given (1)

`inferred (2)`

`not applicable (3)`

6.9.12 Intimidate political opposition (purp_intimidate_opposition)

If the purpose of the PGM is reported to intimidate unarmed political opposition, this variable is coded `1` and `0` otherwise.

6.9.13 Fight insurgents (purp_fight_insurgents)

If the purpose of the PGM is reported to fight armed insurgents, this variable is coded `1` and `0` otherwise.

6.9.14 Fight crime (purp_anti_crime)

If the purpose of the PGM is reported to fight crime, this variable is coded `1` and `0` otherwise.

6.9.15 Destabilise neighbouring countries (purp_destabilise_neighbour)

If the purpose of the PGM is reported as destabilising neighbouring countries, this variable is coded `1` and `0` otherwise.

6.9.16 Purpose (purpose_text)

This text field summarises the intended main purpose of the PGM as reported in the sources.

6.10 Relative benefit

With nine binary variables we record the reported benefit the PGM provides relative to the regular forces. We do not code activities if it if not clear why regular forces were not used for the task. For example, if the PGM is used to commit violence and it is not mentioned why regular forces were not used for this task, we cannot code any information on relative benefit. Note: Sources often do not provide this information, resulting in many missing values for this binary variables.

6.10.1 Knowledge (rel_benefit_knowledge)

This variable is coded `1` if the PGM is reported to provide superior (often local) knowledge compared to the military or the police and `0` otherwise.

6.10.2 Local presence (rel_benefit_local_presence)

This variable is coded `1` if the PGM is reported to being better able to fufill tasks, usually in remote areas, because they were present locally, whereas the military or police were not.

6.10.3 Local support (rel_benefit_local_support)

This variable is coded `1` if the PGM is reported to being supported by the local population.

6.10.4 Faster mobilization (`rel_benefit_faster_mobil`)

This variable is coded 1 if the PGM can mobilize faster than regular forces; this also applies to cases where the government creates or uses the PGM as force multiplier for the regular forces, for example by increasing the number of fighters.

6.10.5 Cheaper mobilization (`rel_benefit_cheaper_mobil`)

This variable is coded 1 if the mobilization of the PGM was less costly than mobilizing or increasing the numbers of regular forces.

6.10.6 Deniability (`rel_benefit_deniability`)

This variable is coded 1 if the government is reported to use the PGM to be able to deny the violence the group commits and to avoid accountability for their actions.

6.10.7 Coup-proofing (`rel_benefit_coup_proof`)

This variable is coded 1 if the government is reported to use the PGM to avoid or minimize the risk of a coup by the military or to counterbalance the military.

6.10.8 Violence (`rel_benefit_use_violence`)

This variable is coded 1 if the government is reported to use the PGM because the military was unwilling to use (excessive) violence.

6.10.9 Loyalty (`rel_benefit_loyalty`)

This variable is coded 1 if the PGM is perceived to be particularly loyal to the political regime, government, or individual members of the ruling elite.

6.11 Target

With several binary variables we identify a wide range of types of individuals or institutions that are reported to have been targeted by the PGM. The variables in this category are coded 1 whenever there is any indication that this type of individual or group has been targeted, at some point, by the PGM. Therefore, they *do not* capture patterns or the strength of patterns in whom the PGM targets, but the breadth of its targets.

6.11.1 Target unarmed government critics (`target_govcritic`)

This variable codes whether the PGM targeted unarmed government critics and members of unarmed political opposition.

6.11.2 Target armed group (`target_rebel`)

This variable codes whether the PGM targeted armed opposition, such as rebels or insurgents.

6.11.3 Target religious group (target_religious)

This variable codes whether the PGM targeted and harmed civilians because they belong to a religious group.

6.11.4 Target ethnic group (target_ethnic)

This variable codes whether the PGM targeted and harmed civilians because they belong to an ethnic group.

6.11.5 Ethnic target name (ethnictarget01–ethnictarget03)

If the PGM targeted individuals because of their ethnicity, we identify the name of the ethnic group based on the Ethnic Power Relations (EPR) core dataset 2019 (Vogt et al. 2015). The potentially three different ethnic groups that are targeted by the PGM are captured with these three variables. If the group given by our sources is mentioned as a sub-group of an EPR-identified group, we list the EPR group as the main group and provide the name of the sub-group in the field `other`. If we cannot identify a group, we use the Ethnolog website to try to match the name of the ethnic group. If our sources do not provide a specific name for the ethnic group, we try to establish a link via other connections, such as the ethnicity of the opposition group that is targeted

6.11.6 EPR id of targeted ethnic group (cowgroupid_t01–cowgroupid_t03)

These three variables code the `cowgroupid` from the Ethnic Power Relations (EPR) core dataset 2019 (Vogt et al. 2015) for the ethnic group names in `ethnictarget0*`.

6.11.7 Target students (target_student)

This variable codes whether the PGM targeted and harmed students.

6.11.8 Quality of information for targeted ethnic group (info_ethnic_targ)

If the PGM is recorded as targeting an ethnic group (`target_ethnic = 1` and we identify the name of the ethnic group, we code whether the name of the ethnic group was given by the PGMD sources or inferred through a link, such as the leader of the group. This variable takes on the following values:

```
given (1)

inferred (2)

not applicable (3)
```

6.11.9 Target peasants (target_peasant)

This variable codes whether the PGM targeted and harmed peasants.

6.11.10 Target journalists (target_journalist)

This variable codes whether the PGM targeted and harmed journalists.

6.11.11 Target foreigners (target_foreigner)

This variable codes whether the PGM targeted and harmed students.

6.11.12 Target foreigners (target_aidworker)

This variable codes whether the PGM targeted and harmed aid workers.

6.11.13 Target immigrants(target_immigrant)

This variable codes whether the PGM targeted and harmed immigrants.

6.11.14 Target criminals (target_criminal)

This variable codes whether the PGM targeted and harmed criminals.

6.11.15 Target PGMs (target_pgm)

This variable codes whether the PGM targeted another PGM.

6.11.16 Target military (target_military)

This variable codes whether the PGM targeted and harmed members of the regular military force of the same country.

6.11.17 Target foreign military (target_foreign_military)

This variable codes whether the PGM targeted and harmed members of a foreign military.

6.11.18 Target peacekeepers (target_pko)

This variable codes whether the PGM targeted and harmed members of an international peacekeeping force.

6.11.19 Target civilians (target_civilian)

This variable is coded 1 if the PGM has targeted and harmed unarmed civilians that are not further identified by a specific group affiliation or characteristic.

6.11.20 Target unclear (target_unclear)

This variable is coded 1 if the information on the PGM target is unclear.

6.11.21 Treatment of civilians (treatment_civilians)

This text field describes how the PGM treated civilians. If relevant information is available, this text field identifies the predominant forms of violence used and temporal patterns in the treatment of civilians.

6.11.22 Type of violence: beating (violencetype_beat)

This binary variable is coded 1 if the PGM was reported to commit beatings. Due to the likely under-reporting of the groups' use of violence, we code this type of violence if one source refers to beatings. If more fine-grained information is available about patterns of the PGM's use of violence, it is described in the text field treatment_civilians.

6.11.23 Type of violence: kidnapping, abductions (violencetype_kidnap)

This binary variable is coded 1 if the PGM was reported to commit kidnappings or abductions. Due to the likely under-reporting of the groups' use of violence, we code this type of violence if one source refers to kidnappings or abductions. If more fine-grained information is available about patterns of the PGM's use of violence, it is described in the text field treatment_civilians.

6.11.24 Type of violence: torture (violencetype_torture)

This binary variable is coded 1 if the PGM was reported to commit torture. Due to the likely under-reporting of the groups' use of violence, we code this type of violence if one source refers to torture. If more fine-grained information is available about patterns of the PGM's use of violence, it is described in the text field treatment_civilians.

6.11.25 Type of violence: sexual violence (violencetype_sexual)

This binary variable is coded 1 if the PGM was reported to commit sexual violence. Due to the likely under-reporting of the groups' use of violence, we code this type of violence if one source refers to sexual violence. If more fine-grained information is available about patterns of the PGM's use of violence, it is described in the text field treatment_civilians. We use the definition by Cohen and Nordås (2014) to identify acts of sexual violence as "(1) rape, (2) sexual slavery, (3) forced prostitution, (4) forced pregnancy, and (5) forced sterilization/abortion (ICC, 2000) [...] (6) sexual mutilation, and (7) sexual torture. Importantly, the definition is gender neutral and does not preclude the existence of female perpetrators or male victims. [...] The definition reflects current legal understandings, but does not include acts such as sexual humiliation, sexualized insults or forced undressing" (Cohen and Nordås 2014, 419).

6.11.26 Type of violence: killings (violencetype_kill)

This binary variable is coded 1 if the PGM was reported to commit killings. Due to the likely under-reporting of the groups' use of violence, we code this type of violence if one source refers to killings. If more fine-grained information is available about patterns of the PGM's use of violence, it is described in the text field treatment_civilians.

6.12 Types of support

We record several potential sources of support that a PGM is reported to have received. We focus on material support and support in training when provided by actors not part of the government the PGM is linked to. We capture these non-exclusive categories with different binary variables and an additional text field.

6.12.1 Support from the domestic government (support_gov)

This binary variable indicates if sources explicitly mention that the PGM received material support from the government it is linked to.

6.12.2 Support from the military (support_military)

This binary variable codes if sources indicate that the PGM received material support from the military of the country the PGM is located in.

6.12.3 Support from business corporations (support_corporation)

This binary variable codes if sources indicate that the PGM received material support from business corporations.

6.12.4 Support from landowners (support_land)

This binary variable codes if sources indicate that the PGM received material support from landowners.

6.12.5 Support from villagers (support_village)

This binary variable codes if sources indicate that the PGM received material support from villagers.

6.12.6 Support from a foreign government (support_foreign_gov)

This binary variable codes if sources indicate that the PGM received material support from a foreign government.

6.12.7 Foreign government sponsor (sponsor_NAMEOFCOUNTRY)

If the variable support_foreign_gov is coded 1, then another binary variables indicates whether the PGM received support from this particular country.

6.12.8 Support from a foreign actor (support_foreign)

This binary variable measures whether the PGM received material from a foreign non-government actor.

6.12.9 Support through drugs (support_drugs)

This variable measures whether the PGM supported itself by dealing with drugs.

6.12.10 Support through plunder and loot (support_loot)

This variable measures whether the PGM supported itself with plundering and looting activities.

6.12.11 Support through crime (support_crime)

This variable measures whether the PGM supported itself through unspecified criminal activities.

6.12.12 PGM being self-maintained (support_self)

This variable codes if sources suggest that the PGM was self-maintained.

6.12.13 Support unclear (support_unclear)

This variable captures if there is conflicting or unclear evidence about how the PGM was supported.

6.12.14 Other connections (other_connections)

This text field captures any other reported connection not identified under the above categories.

6.13 PGM activities (activities)

This variable indicates the years for which we found evidence of PGM activity. Examples of such activities include anything from training exercises, to operations and acts of violence. This variable is used in the online version of the database to shade squares for each year a group was found to be active. This is visible on the country pages of the database that list all PGMs coded for that specific country. This variable is also available in the Country-Year Data file of the PGMD, but not in the Group Data file.

6.14 Other (other)

This text field records any information about the PGM that is not captured by the above variables. If the group is known by alternative names, they are also listed in this field. If the sources reveals changes in the group's characteristics or behaviour across time, they are described in this field.

7 References for meta-analysis

This text field is only available in the online version of the PGMD. It list references other than news sources that have been used for coding the variables for the PGM.

8 Pieces of evidence

To improve transparency of our coding decisions, each coding decision is supported with one piece of information. For reasons of parsimony we did not include all reports on PGMs that we found but only included in the database one item for each (or multiple) coding decisions. The pieces of evidence can be accessed for each PGM via the online PGMD only. It contains over 6,000 key sources that have informed the coding decisions. The database stores only brief quotes from publicly available news sources and reports. If academic sources informed the coding, we list these in 7 in the online database. These pieces of evidence for each group can only be accessed via the online database.

9 Interrupted PGMs

Some PGMs fall out of our definition as being a PGM and then become a PGM again. These special cases are listed here, since the setup of our dataset only allows for one start and one end date. The dates indicate the times when the group was a PGM. The country-year dataset accounts for these temporal interruptions of the subsequently listed PGMs.

9.1 Afghanistan

- Forces of Ismail Khan / Herat militia

  2001 – 2004
  2012-11-14 – 2013-01-01

- Uzbek-Junbesh-e-Milli

  1989-04-20 – 1992
  1992-04-28 – 1993-06
  1996-08 – 1996-09-26
  2001-12-22 – 2007-05
  2009-06-16 –

- Ittehad Islami

  1992-06-29 – 1996-09
  2003-12 –

- Jamaat Islami

  1992-04-28 – 1996-09-27
  2001-12-22 –

9.2 Bangladesh

- Bangadesh Chhatra League

  1996-06-23 – 2001-07-12
  2009-01-06 –

- Jatiyatabadi Chhatra Dal (JCD)

  1979-01-01 – 1982-03-24
  1991-03-20 – 1996-03-30
  2001-10-10 – 2006-10-29

9.3 Congo, Democratic Republic of

- FDLR

  1998-01 – 2002-07
  2007-09 – 2008-12
  2013-07 –

9.4 Georgia

- Forest Brothers

  1998-01-01 – 2004-01-25
  2005-11-17 –

9.5 Haiti

- Ton Ton Macoutes informal

  1986-02-12 – 1991-02-97

  1991-09-30 – 1994-10-19

9.6 Iraq

- Awakening Groups / Sahwa

  2008-11 – 2011-12

  2014 –

- Popular Army/Jerusalem Army

  1970-01-01 – 1991-02-28

  1992-07-02 – 2003-04-09

- Saddam Kurdish Militia

  1987-09-22 – 1991-03-01

  1991-04-05 – 2003-04-09

9.7 Indonesia

- Halilintar

  1975-01-01 -– 1982-12-31

  1995-01-01 – 2003-02-25

9.8 Libya

- Al-Zintan Military Council

  2011-08-23 – 2014-07-01

  2014-08-04 –

- Tripoli Military Council

  2012 – 2012-08-08

  2014-03-01 – 2014-08-04

9.9 Malawi

- Youth Cadets (Democratic Progressive Party)

  2011-04-20 — 2012-04-07

  2014-05-31 — 2020-06-28

9.10 Nepal

- Young Communist League

  2008-08-18 — 2009-05-25
  2011-08-29 — 2013-03-14

9.11 Pakistan

- Jamaat-i-Islami

  1977-07 – 1988-06
  1990-11 – 1992-05

- Muttahida Quami Movement

  1987-11 – 1989-10-23
  1990-11-06 – 1992-06
  1997-02-17 – 1998-10
  2002-10 – 2013-02

9.12 Philippines

- Tagum Death Squads

  1998-07-01 — 2001-12-01
  2004-01-01 — 2013-12-01

9.13 Somalia

- Raskamboni/Raas Kaambooni

  2012-08-20 – 2013-01-31
  2013-09-01 –

9.14 Sudan

- Fertit Militiamen

  1986-01-01 — 1990-11-10
  1997-11-22 — 2007-11-17

9.15 Zimbabwe

- War Vets

  2000-04-21 — 2016-06-01
  2017-11-24 —

- Youth Service Brigade/Green Bombers

  2000-02-12 – 2009-02-01
  2010-01-01 – 2012-01-01
  2017-01-01 – 2017-12-31

10 Country notes

This section provides additional information about specific countries, for example for which time periods some countries were not coded due to the lack of a clearly identifiable government. It also lists certain groups that were not coded in the PGMD and provides the rationale behind it.

10.1 Afghanistan

For Afghanistan, we assume that our data suffer from under-reporting and therefore underestimates the number of PGMs in the country.

A number of private security/military companies have had contracts in Afghanistan to provide a range of security and security-related services, including training of the Afghan National Police and Afghan National Army, logistical support, maintenance, intelligence, de-mining, and poppy eradication. These companies include DynCorp, World Security Initiatives, MPRI, Aegis, Global Risks Group, Armour Group, Ronco, Saladin Security, Gardaworld. One source has identified 90 Afghan and foreign companies by name, but estimates the total number of companies operating in Afghanistan is as high as 140 (Joras and Schuster 2008). These companies are not the principal focus of this research project and search procedures. By June 2008 the government claimed to have disbanded 300 illegal armed groups (see UN, Afghanistan's New Beginnings Programme). To avoid these programmes, one option for government officials and heads of militias is the device of a private security company – see Sherzai in the database.

10.2 Burundi

The militia *Sans Echec* is not included in the database. This group is a Tutsi militia, supporting the Tutsi military but fighting against the Hutu government. They supported Buyoya's election campaign in 1993, but he came second in the election. The group was dissolved when Buyoya took power in a coup in 1996, so they ceased to exist at the time when they would have become pro-government.

10.3 Brazil

Batalhão de Operacões Policiais Especiais (BOPE) and the *Pacification Police Units* (UPPs) were not included in the database as they appear to be part of the regular forces. Off-duty policy militias, such as the *League of Justice* and the *Rio Das Pedras* were not included in the database because the national government does not support them (despite some support from some local governments).

10.4 Chile

The militia *Trizano* is excluded due to insufficient information.

10.5 Colombia

The militia *Black Eagles/Aguilas Negras* is excluded due to insufficient evidence on link to government.

10.6 Comoros Islands

The Presidential Guard is not coded due to lack of information. Appears to be regular forces and, at time, anti-government.

10.7 Cote d'Ivoire

The militia *Patriotic Grouping for Peace* is not included in the PGMD. Although the group is linked to a member within the government, the group was banned by the government in 2003 (US State Department Report 2006), which is prior to the first date of the group's mention.

Not included are the *Dozos* and the *Amadé Ouérémi* militia due to insufficient information on the link to the government.

The following groups are not included due to insufficient information:

- *Patriotic Group for Peace / Group of Patriots for Peace (GPP)*
- *Front for the Liberation of the Greater West (FIGO / FLGO)*
- *Centre-West Security Front (FSCO)*
- *Association of We Patriots (AP-WE)*
- *Ivoirian Movement for the Liberation of Western Cote d'Ivoire (MILOCI)*
- *The Greater West Resistance Union (UPRGO)*
- *Resistance Front of the Greater West (FRGO)*
- *Anti-Terrorist Forces (FAT)*
- *MI - 24*
- *Rapid Intervention Force (FIR)*
- *Special Force for the Liberation of the African World (FS-LIMA)*
- *Patriots' Union for the Total Liberation of Cote d'Ivoire*
- *National Coalition of Cote d'Ivoire's Resistance Fighters (Conareci)*
- *People's Liberation Force*
- *Union of Resistant Patriots of the Greater West (UPERGO)*
- *Union of Patriots for the Total Liberation of Cote d'Ivoire (UPLTCI)*
- *Group of Fighters for the Liberation of Cote d'Ivoire (GCLCI)*
- *CNLB*

10.8 Croatia

Hrvatske Obrambene Snage (HOS) is not included because it was the armed wing of an opposition party and they did not appear to be active within Croatia.

10.9 Egypt

The *Highlander Militia* is excluded due to insufficient information.

10.10 Ethiopia

The *Jailbreakers* are not included due to insufficient information. It is not clear whether the group was armed or organized, or whether it was linked to the government.

Pro-Mubarak vigilantes are not included because it is unclear whether they were armed.

Army-backed vigilantes are not included due to insufficient information.

Green Shirts are not included due to insufficient information.

10.11 Indonesia

These groups were active during the fight for independence in East Timor, but they were not included in the PGMD due to insufficient information:

- *Rajawali (Eagle)*
- *Firmi Merah Putih*
- *Saka Loromonu*
- *ARMUI Merah Putih*
- *Guntur Merah Putih*
- *Guntur Kailak*
- *Harimau Merah Putih*
- *Hametin Merah Putih*
- *Kaer Metin Merah Putih*
- *Darah Merah*
- *Jati Merah Putih*
- *Meo*
- *Morok*
- *Mahadomi*
- *57/75 Junior*

10.12 Iraq

The *Mujahideen Khalq* organization is not included as it is an Iranian exile organization and the target is the Iranian regime.

10.13 Kenya

Kenya has a large number of political militias, which are generally linked to individual politicians, but are not clearly pro-government. These groups are not included. This applies to *Mungiki*, which was supposed to be financially supported by members of the cabinet in 2007/08, but they are generally labelled as “outlawed”, are fought by the policy and the President is explicitly in opposition to the group.

The group *Kanga Squad* is excluded as the group seems to be part of regular police force.

10.14 Mexico

The *Political Investigations* Group is not included as it is based on the testimony of one individual only.

10.15 Mozambique

Community Police is not included due to insufficient evidence of being pro-government.

10.16 Myanmar

Border Guard Force is not included since it appears to be a regular group within the military.

Ka Kwe Ye-kky is not included because this group was dissolved in 1973, prior to the begin of our database.

10.17 Nepal

We used 16 May 2006 as the end date of the King's rule, when the legislature removed the King's executive powers.

10.18 Nigeria

Multiple groups are excluded due to insufficient information:

- *Awkuzu Community Vigilante*: not included due to insufficient information about the link to government.
- *Biu Community Vigilante Group*: not included due to insufficient information about the link to government.
- *Delta State Vigilante Group*: insufficient evidence on being armed.
- *Hisba Guards*: no indication that the group is armed.
- *Kaduna State Vigilante Group*: insufficient evidence on being armed
- *Kano State Vigilante Group*: insufficient evidence on being armed.
- *Operation Sweep*: Not included as it appears to be a special force within the regular security apparatus.
- *Operation Zaki*: Not included as it appears to be a special force within the regular security apparatus.
- *Osun State Vigilante Group*: insufficient evidence on being armed.
- *Robe State Vigilante Group*: insufficient evidence on being armed.

10.19 Philippines

These groups were not included due to insufficient information on whether they were armed, linked to the government or existed after 1981:

- *Negros' Concerned Citizens' Group to Fight Insurgency*

- *Remnants of God*
- *Ituman (The BlackOne)*
- *Putian (The White Ones)*
- *Pula-Puti (Red-White)*
- *Knights of Rizal*
- *Philippine Divine Missionaries for Christ*
- *Rock Christ*
- *White Rock*
- *Power of the Spirit*
- *Alpha Omega*
- *Walay Sapatos (No Shoes)*
- *Tres Cantos (Three Corners)*
- *Salvatore*
- *Dalan sa Langit (Way to Heaven)*
- *Missionaries for Social Transcendental Inc. (MSTI)*
- *Philippine Benevolent Missionaries Association*
- *Counter-Insurgency Command (CIC)*
- *Special Anti-Terrorist Group (SAGT)*
- *Kilusan Laban sa Komunismo (Movement Against Communism)*
- *New Filipino Crusaders*
- *Alsa Bayan*
- *Guerilyang Tagapagtanggol ng Masa (Guerrillas Protecting the Masses)*
- *Kilusan ng Pilipinong Magsasaka (Filipino Farmers Movement)*
- *People's Militia Force (PMF) in Central Luzon*
- *Guardian Angels*
- *Bagong Alyansang Makabansa (Patriotic New Alliance, BANSA)*
- *Alyansa ng Magsasaka Labansa Komunismo (Alliance of Farmers Against Communism, ALMAKA) in Nueva Ecija*
- *Peace and Order Brigade in Pampanga*
- *KILCOM (Movement Against Communism)*
- *Movement Against Violence in Bicol*
- *Yellow Union, an Aquino-loyalist group in Masbate*
- *KATAD (Kapatiran ng mga MindorefioTungo sa Adhikaing Demokratiko*

- *Brotherhood of Mindorefios Advocating Democratic Principles*
- *ANAKPAWIS on Mindoro*
- *Magic Eye*
- *Youth Intelligence Group (organized by the Quezon City OIC Mayor)*
- *Association of Democratic Vigilantes and Concerned Entrepreneurs (ADVANCE)*
- *United Vigilantes Association*
- *People's Liberation Organization (PLO)*

10.20 Romania

The *Vigilante Miners* are not included due to insufficient evidence for some form of organisation. They appear to be an ad-hoc group.

10.21 Serbia

Serbia is coded from 2006 onwards.

The group *Bridge Watchers* is not included. They were a pro-Serbian government group active on the bridge dividing Kosovska Mitrovica and Mitrovica Hospital, Kosovo, just after Kosovo's independence (17 February 2008, based on the Archigos dataset).

RS Ravna Gora Chetnik Movement is not include. This group was only active in Kosovo and founded in March 2008, after Kosovo's independence.

10.22 Kosovo

The *Civil Defense Force* included various ethnic militias (*Gbethis*, *Donsos*, *Kapras*), which are not separately listed.

The *Tamaporo* and *Kamajors* are separately listed due to some diverging characteristics that could be identified from the sources.

Between May 1997 and March 1998 President Kabbah – allied with CDF – was not in the capital. Given the annual coding, CDF is listed as pro-government in 1997 and 1998.

From May 1997 until Feb 1998 the Armed Forces Revolutionary Council controlled the capital and were backed by *RUF*. Therefore, during this time period, the *RUF* is considered to be a pro-government armed group.

Sandline employed to counter *RUF* during this period, but not included in the dataset as *RUF* formed part of the government during this period.

The *Movement of Concerned Kono Youth MOCKY* not included. Though anti-RUF, it is not clearly identified as pro-government.

10.23 Somalia

Somalia is coded in the PGMD from 20 August 2012 onwards, when the Federal Government of Somalia took over power from Somalia's Transitional Federal Government.

The United Rahweyn Movement is not coded as a PGM due to insufficient information about the link to the government.

10.24 Sudan

In 2005, the Comprehensive Peace Agreement (CPA) granted South Sudan a semi-autonomous status and established peace between the Khartoum government and the *Sudan People's Liberation Movement/Army* (SPLM/A). Nevertheless, the SPLM/A does not join the Sudanese army and is not pro-government after 2005, as it still supports the idea of an independent South Sudan and does not work for the Khartoum government. Hence, it is considered to be a rebel group and not a PGM.

Some of the PGMs in Sudan merge with the SPLM/A after 2005 (*South Sudan Unity Movement/Army, South Sudan Defence Forces, Murle Forces, Fertit Militiamen*) and consequently become part of this rebel group. In this instance we code the PGMs as becoming rebel groups and the termination type as *PGM defects*. The reported UCDP Actor ID in these cases is the actor ID of the SPLM/A. For further information, see also the pieces of evidence provided in the database for the *South Sudan Unity Movement/Army, South Sudan Defence Forces, Murle Forces* and *Fertit Militiamen*.

Thourjikany Forces are not included due to insufficient evidence. They are possibly a small splinter group of the *South Sudan Defence Forces* (SSDF).

10.25 South Africa

31 Battalion and *32 Battalion* are excluded as they appear to be part of regular forces.

Koevoet is not included. It is unclear whether this is a special unit within the regular forces; the group was integrated into police force in March 1990. It was active since 1979, but evidence refers to activities only outside of South Africa.

The *Black Cats* are not included, as there is no evidence of them being linked to the government before 1994 and no activities are recoded post-1994.

10.26 Syria

The *Houthis* are not included because we found no evidence that they have any kind of organizational structure in Syria.

Iranian Revolutionary Guards Corps (IRGC) are not included because we found no evidence that they are linked to the Syrian government.

Kataib Hezbollah is not included because we found no evidence that it is linked to the Syrian government.

10.27 Uganda

The *Black Mambas* are not included as they appear to be part of the regular forces.

Operation Wembley is not included as this appears to refer to a mission of regular forces and not a separate group.

The *Violent Crime Crack Unit* (VCCU) is not included as this appears to refer to a mission of regular forces and not a separate group.

10.28 Ukraine

Since the fall of President Yanukovych in February 2014, many splinter groups and battalions were fighting on the side of the Ukraine government against the separatists in the east of the Ukraine. For most of them, however, we could not establish a link to the Ukraine government and were therefore unable to include them in the PGMD.

10.29 United Kingdom

There is evidence of collusion between security force units and members of the loyalist paramilitary group *Ulster Defence Association* in Northern Ireland in late 1980s in particular. Some individuals were linked to the British government at Force Research Unit level, but it is not clear that the group was. The group is also not identified as pro-government by the sources. This is clearly a difficult case but due to the absence of support for the group itself being pro-government, we did not include the UDA. Although there were government informers in the groups, the groups are not identified as pro-government and there is no evidence of government ministerial knowledge of the collusion (Guardian 18 April 2003).

10.30 United States of America

Evidence suggests that the State Defence Forces are unarmed, except for in Alaska. There is no documentary evidence of operations in which armed groups have been active.

10.31 Venezuela

Francisco Miranda Front was not included due to insufficient information.

10.32 Yugoslavia

The *Krajina militia / Marticevci* is not included as it was active in Croatia and Bosnia after they were already independent countries (see Archigos). The Krajina Militia acted as the armed forces of the Republic of Serbian Krajina. This was a pro-Serbian de-facto regime that covered around 1/3 of Croatia, but was not a recognized state.

10.33 Zambia

The *MMD Cadres* are not included because of insufficient evidence of the group being armed.

10.34 Zimbabwe

The *Upfumi Kuvadiki* (Wealth to the Young) is not included due to insufficient evidence of the group being armed.

References

Cohen, Dara Kay and Ragnhild Nordås. 2014. "Sexual violence in armed conflict: Introducing the SVAC dataset, 1989–2009." *Journal of Peace Research* 51(3):418–428.

Gleditsch, Kristian S and Michael D Ward. 1999. "Interstate system membership: A revised list of the independent states since 1816." *International Interactions* 25(4):393–413.

Goemans, Henk E., Kristian Skrede Gleditsch and Giacomo Chiozza. 2009. "Introducing Archigos: A Dataset of Political Leaders." *Journal of Peace Research* 46(2):269–283.

Joras, Ulrike and Adrian Schuster. 2008. "Private Security Companies and Local Populations: An Exploratory Study of Afghanistan and Angola.".

Pettersson, Therese and Magnus Öberg. 2020. "Organized violence, 1989–2019." *Journal of peace research* 57(4):597–613.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* 59(7):1327–1342.