Transborder Ethnic Ties and Repression of Ethnic Minorities

Codebook

Dyad ID: The unique ID for each Government-External Kin dyad (Government's COW Code_External Kin's Code).

External Kin Country Code: COW country code of the External Kin's country (Data Source:) (Sarkees and Wayman, 2010).

Year: Year.

Government Country Code: COW code of the Minority group's country (Sarkees and Wayman, 2010).

Government's Name: Name of the Minority group's country.

Minority Group ID: EPR-TEK ID of the Minority group (Vogt et al., 2015).

Minority's Name: Name of the Minority group (Vogt et al., 2015).

Minority's Size: The percentage of the Minority group's population to total population of its country (Vogt et al., 2015).

Minority's Status: Status of the Minority group. A disadvantaged/weak minority group can have one of the following status (Vogt et al., 2015, p. 5):

- Powerless: Elite representatives hold no political power (or do not have influence on decision making) at the national level of executive power - although without being explicitly discriminated against.

- Discriminated: Group members are subjected to active, intentional, and targeted discrimination by the state, with the intent of excluding them from political power. Such active discrimination can be either formal or informal, but always refers to the domain of public politics (excluding discrimination in the socio-economic sphere).

- Self-exclusion: The special category of self-exclusion applies to groups that have excluded themselves from central state power, in the sense that they control a

particular territory of the state which they have declared independent from the central government.

Minority has a powerful EK: A dichotomous variable coded as "1" if the Minority group has at least one powerful (in the sense of being the major group with access to political power) ethnic kin in another country, and "0" otherwise (Vogt et al., 2015).

EK Country Name: Name of the powerful ethnic kin's country (**?**).

EK's Country ID': EPR-TEK ID of the External Kin's country (Vogt et al., 2015).

EK Group Name: Name of the powerful ethnic kin group (Vogt et al., 2015).

EK's Status: Status of the External Kin in its own country. External Kin can have access to government power at varying levels (Vogt et al., 2015, p. 5):

- Monopoly: Elite members hold monopoly power in the executive to the exclusion of members of all other ethnic groups.

- Dominance: Elite members of the group hold dominant power in the executive but there is some limited inclusion of members of other groups who however do not have real influence on decision making.

- Senior Partner: Representatives of the group participate as senior partners in a formal or informal power-sharing arrangement. By power sharing, we mean any arrangement that divides executive power among leaders who claim to represent particular ethnic groups and who have real influence on political decision making.

EK's Status Start Year: The year that *EK's Status* started (Vogt et al., 2015).

EK's Status End Year: The year that *EK's Status* ended (Vogt et al., 2015).

EK's Size: The percentage of the external kin group's population to total population of its country (Vogt et al., 2015).

G-EK Contiguity: A dichotomous variable coded as "1" if the the External Kin's

country and the Minority's country are contiguous in any of the following forms (Douglas et al., 2002, p. 3):

- Separated by a land or river border

- Separated by 12 miles of water or less

- Separated by 24 miles of water or less (but more than 12 miles)

- Separated by 150 miles of water or less (but more than 24 miles)

- Separated by 400 miles of water or less (but more than 150 miles)

G-EK DefensiveAlliance: A dichotomous variable coded as "1" if the Government and External Kin are in a defensive alliance, and "0" otherwise (Gibler, 2008).

G-EK NeutralityAlliance: A dichotomous variable coded as "1" if the Government and External Kin signed terms that included neutrality toward one or more states in the alliance, and "0" otherwise (Gibler, 2008).

G-EK NonaggressionAlliance: A dichotomous variable coded as "1" if the Government and External Kin signed terms that included a promise of non-aggression toward one or more states in the alliance, and "0" otherwise (Gibler, 2008).

G-EK Entente: A dichotomous variable coded as "1" if the Government and External Kin signed terms that included an understanding that they would consult with one or more states in the alliance if a crisis occurred, and "0" otherwise (Gibler, 2008).

G-EK Smoothtotrade: Smoothed total trade values between the Government and External Kin. Total trade values is calculated as the sum of imports of Country A form Country B and imports of Country B from Country A, in US millions of current dollars (Barbieri, Keshk and Pollins, 2008).

Total Trade: The natural logarithm of *G-EK Smoothtotrade.*

Government's CINC Score: CINC Score of the Government. The Composite Index

of National Capability (CINC) score aggregates the six individual measured components of national material capabilities (total population, urban population, military personnel, military expenditures, primary energy consumption, and iron and steel production) into a single value per state-year. The CINC reflects an average of a state's share of the system total of each element of capabilities in each year, weighting each component equally. In doing so, the CINC will always range between 0 and 1. "0.0" would indicate that a state had 0% of the total capabilities present in the system in that year, while "1.0" would indicate that the state had 100% of the capabilities in a given year (Singer, Bremer and Stuckey, 1972, p. 8).

EK's CINC Score: CINC Score of the External Kin. Coding rules are same with the *Government's CINC Score* (Singer, Bremer and Stuckey, 1972).

Total Trade between the G and EK (imputed): Same as *G-EK Smoothtotrade*. Missing values (for 2014-2017) are imputed.

Government's CINC Score (imputed): Same as *Government's CINC Score*. Missing values (for 2012-2017) are imputed.

EK's CINC Score (imputed): Same as *EK's CINC Score*. Missing values (for 2012-2017) are imputed.

G-EK CINC Ratio: Equals to Government's CINC Score (imputed)/Government's CINC Score (imp EK's CINC Score (imputed). It measures the ratio of material power of the Government and External Kin. The higher the *EK's CINC Score* compared to *Government's CINC Score*, the smaller the *G-EK CINC Ratio*. In other words, smaller *G-EK CINC Ratio* means more powerful EK.

Government's Democracy Level: Regime type of the Government. It is coded as follows (Coppedge et al., 2018):

- 1= Electoral autocracy: De-jure multiparty elections for the chief executive and the legislature, but failing to achieve that elections are free and fair, or de-facto

multiparty, or a minimum level of Dahl's institutional prerequisites of polyarchy as measured by V-Dem's Electoral Democracy Index.

- 2= Electoral democracy: De-facto free and fair multiparty elections and a minimum level of Dahl's institutional prerequisites for polyarchy as measured by V- Dem's Electoral Democracy Index, but either access to justice, or transparent law enforcement, or liberal principles of respect for personal liberties, rule of law, and judicial as well as legislative constraints on the executive not satisfied as measured by V-Dem's Liberal Component Index.

- 3= Liberal democracy: De-facto free and fair multiparty elections and a minimum level of Dahl's institutional prerequisites for polyarchy as measured by V- Dem's Electoral Democracy Index are guaranteed as well as access to justice, transparent law enforcement and the liberal principles of respect for personal liberties, rule of law, and judicial as well as legislative constraints on the executive satisfied as measured by V-Dem's Liberal Component Index.

EK's Democracy Level: Regime type of the External Kin. Coding rules are same with the *Government's Democracy Level* (Coppedge et al., 2018)

Both Sides Are Democratic: A dichotomous variable coded as "1" if both the Government and External Kin are democratic (electoral or liberal according to the definitions of V-dem data), and "0" otherwise (Coppedge et al., 2018).

Dispersion: The initial reason that the minority and external kin live in different countries:

- 1= Migration

- 2= Collapse of an Empire

- 3= Colonial Borders

- 4= Never lived together (but historically share the same ethnic origins)

Shared Language/Religion between the EK and M: In addition to the shared ethnic origins, the minority and external kin can also share the same language and/or religion:

- 0=no language or religious ties

- 1=only language or religious

- 2=both language and religious ties

Minority's Goal: Minority's goals/demands from the government

- 1=policy change (improvement of minority rights and/or economic grievance)

- 2=autonomy (or greater autonomy)

- 3=independence

Minority's Coherence: To what extent the minority group is centrally institutionalized:

- 0=unorganized/not institutionalized (no NGO/party/militant organizations)

- 1=moderately centralized (NGO-Civil Society Organizations)

- 2=highly centralized and organized (NGO-Civil Society Organizations+Political Party or Militant organization)

Level of Government Repression: The level of government repression the minority suffers:

- 0= no repression

- 1= minor violations of rights

- 2= major violations of rights

- 3= genocide

Cultural Repression: A dichotomous variable coded as "1" if the Government violates the cultural rights of the minority and as "0" otherwise.

Economic Repression: A dichotomous variable coded as "1" if the Government violates the economic rights of the minority and as "0" otherwise.

Political Repression: A dichotomous variable coded as "1" if the Government violates the political rights of the minority and as "0" otherwise.

Physical/Violent Repression: A dichotomous variable coded as "1" if the Government violates the physical integrity rights of the minority and as "0" otherwise.

Coup Risk for the Government: Is there any coup risk, security, or bureaucratic defection?

- 0=No coup risk

- 1=Only security or bureaucratic defection

- 2=Both security and bureaucratic defection

Third-Party Support to Minority: Does any other third party (a State/IO/INGO) support the Minority against the Government:

- 0=No State/IO/INGO support

- 1= Low-level diplomatic support by a State/IO/INGO

- 2=Diplomatic support + Economic sanctions by a State/IO/INGO

Third-Party Support to Government: Does any other third party (a State/IO/INGO) support the Government:

- 0=No State/IO/INGO support

- 1= Low-level diplomatic support by a State/IO/INGO

- 2=Diplomatic support + Economic sanctions by a State/IO/INGO

Diplomatic Support by EK: A dichotomous variable coded as "1" if the External Kin provides diplomatic support to the Minority, and as "0" otherwise.

Economic Support by EK: A dichotomous variable coded as "1" if the External Kin provides economic support to the Minority, and as "0" otherwise.

Military Support by EK : A dichotomous variable coded as "1" if the External Kin provides military support to the Minority, and as "0" otherwise.

Level of EK Support: Level of the support that the External Kin provides to Minority (if any):

- 0=No support

- 1=Low

- 2=Medium

- 3=High

Magnitude of the EK's Support: A function of the type and level of support provided by the External Kin. I assume that military support is more effective than economic support; and economic support is more effective than diplomatic support. Accordingly, this variable is coded as follows:

$$\text{External Kin's Support} = \text{Level of Support (Diplomatic Support} + \\ 2.(\text{Economic Support}) + 3.(\text{Military Support})) \tag{1}$$

Dual citizenship: Do the members of the Minority hold dual citizenship (both in their host country and the external kin's country)?

- 0=No dual citizenship

- 1=Dual citizenship

Diaspora: Does the Minority have a diaspora organization in the External Kin's country:

- 0= M does not have any diaspora organization in the EK's country

- 1= M has a diaspora organization in the EK's

Revisionist External Kin: A dichotomous variable coded as "1" if the EK is revisionist according to COW Dyadic MID (v3.0) (Maoz et al., 2019), and "0" otherwise.

Is MID related to Minority's Status: A dichotomous variable coded as "1" if the MID between the Government and EK is related to the Minority's status.

Government's Ethnic Fractionalization: The probability that two individuals selected at random from Minority and Government's country will be from different ethnic groups [0-1] (Fearon, 2003).

EK's Ethnic Fractionalization: the probability that two individuals selected at random from External Kin's country will be from different ethnic groups [0-1] (Fearon, 2003).

Magnitude of the Repression: A composite variable measuring the overall magnitude of government repression. It is calculated as follows:

$$
\begin{aligned}
\text{Magnitude of the Repression} = \text{Level of Government Repression(Cultural Repression+} \\
\text{Economic Repression} + \text{Political Repression} + 2(\text{Physical/Violent Repression}))
\end{aligned}
\tag{2}
$$

Physical/Violent Repression is multiplied by two as it is expected have stronger effects on the dependent variables.

Domestic Violence: A dichotomous variable coded as "1" if there was a resistance by the Minority according to NAVCO (v.3) (Chenoweth and Lewis, 2013) or a civil war according to COW War Datat (v.4.0) (Sarkees and Wayman, 2010), and "0" otherwise.

Government's GDP: Expenditure-side real GDP at chained PPPs, to compare relative living standards across countries and over time (Feenstra, Inklaar and Timmer, 2015).

Government's GDP (Logged): The natural logarithm of *Government's GDP*.

EK's GDP: Expenditure-side real GDP at chained PPPs, to compare relative living standards across countries and over time (Feenstra, Inklaar and Timmer, 2015).

Captial-to-capital Distance: The distance between the capital cities of the Government and External Kin based on the GeoDist data (Mayer and Zignago, 2011).

Distance (Logged): The natural logarithm of *Captial-to-capital Distance*.

G-EK DefensiveAlliance (z-score): z-score of *G-EK DefensiveAlliance*.

G-EK NeutralityAlliance (z-score): z-score of *G-EK NeutralityAlliance*.

G-EK NonaggressionAlliance (z-score): z-score of *G-EK NonaggressionAlliance*.

G-EK Entente (z-score): z-score of *G-EK Entente*.

G-EK Total Trade (z-score): z-score of *G-EK Smoothtotrade*

Dispersion (z-score): z-score of *Dispersion*.

M-EK Shared Language/Religion (z-score): z-score of *Shared Language/Religion between the EK and M*.

Dual citizenship (z-score): z-score of *Dual citizenship*.

Diaspora (z-score): z score of *Diaspora*.

EK's Democracy Level (z-score): z-score of *EK's Democracy Level*.

EK's Ethnic Fractionalization (z-score): z-score of *EK's Ethnic Fractionalization*.

Third-Party Support to Minority (z-score): z-score of *Third-Party Support to Minority*.

G-EK CINC Ratio (z-score): z-score of *G-EK CINC Ratio*.

Minority's Size (z-score): z-score of *Minority's Size*.

Minority's Coherence (z-score): z-score of *Minority's Coherence*.

Minority's Goal (z-score): z-score of *Minority's Goal*.

Magnitude of the EK's Support (z-score): z-score of *Magnitude of the EK's Support*.

p_e threshold: A composite variable which increases when (1) the Government and the External Kin are contiguous, (2) the Minority and External Kin share the same language and religion, (3) M has an active diaspora organization in the External Kin's country, (4) the External Kin is a democratic country, (5) initial reason of dispersion of the ethnic group is migration or collapse of an empire, and (6) the members of the Minority group can have dual citizenship from the External Kin's state. It decreases when (1) there is an alliance between the Government and External Kin, (2) the size of total trade between the Government and External Kin is high, and (3) the level of ethnic fractionalizaton in the External Kin's country is high. It is calculated as the sum/extraction of the *z scores* of these variables, rather than their actual values, to avoid a biased estimation.

p_e threshold (normalized): It is the normalized version of *p_e threshold*: $p_e - min(p_e)/max(p_e) - min(p_e)$.

g_1 threshold: A composite variable which increases when (1) the Government's material capability is superior to the External Kin, and (2) there is third-party support the Government. It decreases when there is third-party support (other than the External Kin) to the Minority. It is calculated as the sum/extraction of the *z scores* of these variables, rather than their actual values.

g_1 threshold (normalized): It is the normalized version of *g_1 threshold*: $g_1 - min(g_1)/max(g_1) - min(g_1)$.

m_s threshold: A composite variable of which increases when the Minority (1) has bigger population, (2) is more coherent and has centralized organizations, (3) has a more ambitious goal such as autonomy or independence, (4) receives substantial support by the External Kin and/or another third party. It is calculated as the sum/extraction of the *z scores* of these variables, rather than their actual values.

m_s threshold (normalized): It is the normalized version of *m_s threshold* : $m_s - $

$min(m_\mathrm{s})/max(m_\mathrm{s}) - min(m_\mathrm{s})$.

Peace Years (Until MID): Number of years until the occurrence of MID.

Peace Years^2 (Until MID): Square of *Peace Years (Until MID)*.

Peace Years^3 (Until MID): Cube of *Peace Years (Until MID)*

Peace Years (Until Repression): Number of years until the occurrence of *Repression.*

Peace Years^2 (Until Repression): Square of *Peace Years (Until Repression).*

Peace Years^3 (Until Repression): Cube of *Peace Years (Until Repression).*

Peace Years (Until Domestic Violence): Number of years until the occurrence of Domestic Violence.

Peace Years^2 (Until Domestic Violence): Square of *Peace Years (Until Domestic Violence).*

Peace Years^3 (Until Domestic Violence): Cube of *Peace Years (Until Domestic Violence).*

Status-Quo: A dichotomous variable coded as "1" if the Minority is *Powerless* according to the EPR data (Vogt et al., 2015). A Minority is *Powerless* if "Elite representatives hold no political power (or do not have influence on decision making) at the national level of executive power - although without being explicitly discriminated against" (Vogt et al., 2015, p. 5). *Status-Quo* is coded as "0", if the Minority is *Discriminated* in the EPR data (Vogt et al., 2015). A Minority is *Discriminated* if "Group members are subjected to active, intentional, and targeted discrimination by the state, with the intent of excluding them from political power. Such active discrimination can be either formal or informal, but always refers to the domain of public politics (excluding discrimination in the socio-economic sphere)." (Vogt et al., 2015, p. 5).

References

Barbieri, Katherine, Omar Keshk and Brian Pollins. 2008. "Correlates of War Project Trade Data Set Codebook." *Codebook Version* 2.

Chenoweth, Erica and Orion A Lewis. 2013. "Nonviolent and Violent Campaigns and Outcomes Dataset, v. 2.0." *University of Denver* .

Coppedge, Michael, John Gerring, Staffan I Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman and Jeffrey Staton. 2018. "V-Dem Codebook v7.1 Varieties of Democracy (V-Dem) Project.".

Douglas, Stinnett, Tir Jaroslav, Schafer Philip and Gochman Charles. 2002. "The Correlates of War Project Direct Contiguity Data." *Conflict Management and Peace Science* 19(2):59–68.

Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of economic growth* 8(2):195–222.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105(10):3150–82.

Gibler, Douglas M. 2008. *International military alliances, 1648-2008.* CQ Press.

Maoz, Zeev, Paul L. Johnson, Jasper Kaplan, Fiona Ogunkoya and Aaron P. Shreve. 2019. "The Dyadic Militarized Interstate Disputes (MIDs) Dataset Version 3.0: Logic, Characteristics, and Comparisons to Alternative Datasets." *Journal of Conflict Resolution* 63(3):811–835.

Mayer, Thierry and Soledad Zignago. 2011. Notes on CEPII's Distances Measures: The GeoDist Database. Working Papers 2011-25 CEPII.
URL: *http://www.cepii.fr/CEPII/en/publications/wp/abstract.asp?NoDoc=3877*

Sarkees, Meredieth Reid and Frank Wayman. 2010. *Resort to War: 1816-2007. Correlates of War.* Washington DC: CQ Press.

Singer, J. David, Stuart Bremer and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In *Peace, War, and Numbers*, ed. Bruce Russett. Beverly Hills: Sage pp. 19–48.

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.