Minorities at Risk

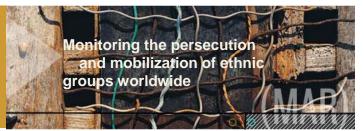Monitoring the persecution and mobilization of ethnic groups worldwide

All Minorities at Risk (AMAR) Codebook Version 8/2016

A project of

CIDCM

Center for International Development and Conflict Management

Contents

All Minorities at Risk (AMAR)

Phase I Codebook

August 2016

I.   AMAR Project

Project Overview

The All Minorities at Risk (AMAR) project is a university-based research project that monitors and analyzes the status and conflicts of politically-active communal groups in countries with a current population of at least 500,000. The project is designed to provide information in a standardized format that aids comparative research and contributes to the understanding of conflicts involving relevant groups.

The project was founded in 1986 as the Minorities at Risk data (MAR) by Ted Robert Gurr, one of the preeminent scholars of political violence and ethnic conflict. Since 1988, the Center for International Development and Conflict Management (CIDCM) at the University of Maryland has hosted the project. The MAR Project has been led by Ted Robert Gurr (1986-2004), Christian Davenport (2004-2006), and Jonathan Wilkenfeld (2007-2010). The AMAR project is led by Jóhanna Birnir (2010-present). The MAR and AMAR project have been funded by various agencies, including the Carnegie Corporation, the Hewlett Foundation, the National Consortium for the Study of Terrorism and Responses to Terrorism (START), the National Science Foundation, the State Failure (now Political Instability) Task Force, the United States Institute of Peace, and the U.S. Department of Homeland Security, and the Department of Government at the University of Maryland.

The MAR dataset is the original core component of the project. In addition to the dataset of quantitative indicators, narrative risk assessments, analytic summaries, and chronologies of events for included groups are available on the project website. The MAR website also provides information on related projects and a bibliography of published materials utilizing the dataset for analysis. Users of MAR data (and from now on the AMAR data) are encouraged to provide copies and citations of publications to project staff for listing on the website. Users are also encouraged to provide narratives of groups they study to be included in a collection of user narratives on the website.

MAR and AMAR have grown through the tireless effort of many graduate assistants and several faculty associates. The project coordinators who have played a key role in sustaining the project since the mid-1980s include Monty G. Marshall, Scott McDonald, Shin-wha Lee, Michael Haxton, Anne Pitsch, Randi Mack, Michael Johns, Amy Pate, Carter Johnson, Mary Michael, Chad Drummond, Molly Inman, and Agatha S. Hultquist. To provide guidance on groups to be included, new indicators, and data quality control issues, an advisory board was established for the Minorities project in 1999. Several faculty affiliates – many of them members of the advisory board – have also provided leadership in procuring funding and in initiating data collection projects. These include Steve Saideman

(Carleton University), Victor Asal (SUNY Albany) and Jóhanna Birnir (University of Maryland). Information on board members and activities is posted on the AMAR web site. Information about the current project staff can also be found on the AMAR web site.

Users can communicate with the AMAR project by email at amar-cidcm@umd.edu, an address that is managed by the AMAR Project Coordinator. Additional contact information may be found on the AMAR website.

Project History

The MAR dataset developed over six distinct phases with the transition to AMAR constituting the last phase. MAR Phase I covered 227 communal groups, which met the criteria for classification as a minority at risk for the years 1945-1989. MAR Phase II covered 275 groups from 1990-1995, MAR Phase III covered 275 groups from 1996-1998, and MAR Phase IV covered 287 groups from 1998-2003. The release of the 2004-2006 data marked the beginning of MAR Phase V, which included a reformulation of the codebook. AMAR Phase I is the All-Minorities at Risk (AMAR) project, an independent continuation of the MAR project.

The AMAR Phase I project is a Sample Frame, consisting of a list of commonly recognized socially relevant ethnic groups that are minorities and majorities alike (see Birnir et al. 2015).[1] The MAR compatible criteria for inclusion of groups in AMAR are as follows:

1. Membership in the group is determined primarily by descent by both members and non-members.[2]
2. Membership in the group is recognized and viewed as important by members and/or non-members. The importance may be psychological, normative, and/or strategic.
3. Members share some distinguishing cultural features, such as common language, religion, occupational niche, and customs.
4. One or more of these cultural features are either practiced by a majority of the group or preserved and studied by a set of members who are broadly respected by the wider membership for so doing.
5. The group has at least 100,000 members or constitutes one percent of a country's population.

Based on these selection criteria, the AMAR Phase I project identified 1,202 ethnic groups. Of these groups, 291 are groups that appeared in MAR Phase V and 911 are new groups. For these 1,202 AMAR groups, AMAR project staff collected data on country population (CPOP), group population (GPOP), and group proportion (GPRO) data. This data was collected and calculated based on population estimates for 2001 or 2007, depending on the

---

[1] Birnir, Johanna K., Jonathan Wilkenfeld, James D. Fearon, David Laitin, Ted Robert Gurr, Dawn Brancati, Stephen Saideman, Amy Pate, and Agatha S. Hultquist. 2015. "Socially relevant ethnic groups, ethnic structure and AMAR." Journal of Peace Research 52(1): 110-115.

[2] The group may be a caste determined by descent.

availability of data. The Sample Frame, which includes the list of groups and corresponding population data, is described in Birnir et al. 2016 and available on the AMAR web site. AMAR project staff also catalogued MAR groups that were merged to create new groups in AMAR, MAR groups that were split to create new groups in AMAR, and MAR groups whose names were changed slightly in AMAR.

For all 1,202 AMAR groups, AMAR project staff also collected data on whether there was one or more politically dominant ethnic group (based on control of the executive branch) in the country.   The Sample Frame of 1,202 AMAR groups was also compared to ethnic group lists in three prominent datasets – Alesina et al. 2003, Ethnic Power Relations (EPR) versions 1.1 and 3.0, and Fearon 2003 – and match variables were coded for each list.

For a list of all the new AMAR variables, see *here* in the codebook.

The AMAR Phase I project corrects selection issues in the original MAR data (Phases I –V). Specifically, 74 new groups were selected randomly from the Sample Frame of socially relevant groups outlined by Birnir et al. 2015, according to the new AMAR criteria. The original 291 MAR groups and the newly coded 74 selection bias groups were then joined in an appropriately weighted AMAR sample (for details on the weighting see Birnir et al. 2016[3]) that can be used to answer questions about ethnic politics without concerns about biases related to selection criteria of political mobilization.

In addition, AMAR Phase I reconciles some of the MAR variables that were reformulated during the 2006 review of the codebook to create a single, integrated dataset.

The AMAR Phase I data therefore consists of the following components:

- For all 1,202 AMAR ethnic groups, group proportion, political dominance, matching variables, and corresponding weights.

- For 365 AMAR ethnic groups (291 original MAR Phase V groups and 74 selection bias groups), all variables listed in the remainder of this codebook.

Research Protocol

Coders for the project are primarily graduate and undergraduate students who have undergone a rigorous training procedure. All coding is then reviewed by senior editors and by the research director before public release. The project has not yet assessed inter-coder reliability. All coding is conducted using open-source information. As much as is possible, coders rely on multiple sources for each code assigned. Details on coding conventions for specific variables are contained within the description of variables below.

Researchers are encouraged to carry out their own consistency and validity checks on indicators they use or adapt from the MAR dataset. Project staff would greatly appreciate

---

[3] Birnir Jóhanna K, David D. Laitin, Jonathan Wilkenfeld, Agatha S. Hultquist, David M. Waguespack, and Ted R. Gurr. 2016. "Socially Relevant Identity: Addressing Selection Bias Issues and Introducing the AMAR (All Minorities at Risk) Data." CIDCM working paper.

being appraised of the results of such analysis. Furthermore, MAR users are encouraged to notify project staff of any discrepancies found in the data for further evaluation.

Resources and Documentation

The MAR Project maintains hard copy files for MAR Phases I through III of the project and computer records for MAR Phases IV and V (although the documentation files for the first two phases are incomplete for some groups). Computer records are also kept for the AMAR selection bias data and AMAR Phase I data. The project archives currently contain:

- MAR Phase I codesheets and group summaries
- MAR Phase II chronologies for a subset of Phase I groups
- MAR Phase III codesheets, overviews, chronologies and risk assessments
- Maps of most MAR Phase III groups, showing areas of geographic concentration
- Selected hard-copy source materials for MAR Phases I and II
- Codebooks and coding conventions for MAR Phases I, III and IV
- Hard copies of selected publications using MAR data
- Access database of MAR Phase IV coding
- Access database of MAR Phase V coding
- Access database of AMAR selection bias coding

Individual researchers can gain access to these materials by arrangement with the project coordinator.

The MAR Phase III 1999-2000 and MAR Phase IV 2001-2003 data were released contained with the MARGene program. With the reformulation of the codebook in MAR Phase V, the MARGene program was discontinued. The project is moving to a web-based platform for the creation of customized datasets. However, the data contained with MARGene and the program itself will be archived.

Citing AMAR

For use of the AMAR quantitative data, please cite as follows:

Birnir Jóhanna K., David D. Laitin, Jonathan Wilkenfeld, Agatha S. Hultquist, David M. Waguespack, and Ted R. Gurr. 2016. "Socially Relevant Identity: Addressing Selection Bias Issues and Introducing the AMAR (All Minorities at Risk) Data." CIDCM working paper.

The following paper introduces the AMAR Sample Frame:

Birnir, Johanna K., Jonathan Wilkenfeld, James D. Fearon, David Laitin, Ted Robert Gurr, Dawn Brancati, Stephen Saideman, Amy Pate, and Agatha S. Hultquist. 2015. "Socially relevant ethnic groups, ethnic structure and AMAR." Journal of Peace Research 52(1): 110-115.

The remainder of this codebook outlines all the AMAR Phase I variables and details the reconciliation process for each variable that was reconciled. Users are encouraged to scrutinize each reconciliation decision to make sure it adheres to their research objectives.

We have included recoding syntax that was written to correct coding mistakes found during the MAR Phase V update. This syntax also contains some coding to back-code groups for the MAR Phase V variables.

The following terms are used in reference to the different versions of the data:

- MAR Phase IV – MAR data for the years 1940-2003 coded according to the Phase IV codebook.[4]

- MAR Phase V – MAR data for the years 2004-2006 and AMAR selection bias data for the years 1980-2006 coded according to the Phase V codebook.

- AMAR Phase I – Reconciled MAR Phase IV 1940-2003 data, MAR Phase V 2004-2006 data, and AMAR 1980-2006 selection bias data for 365 ethnic groups, with weights included for analysis. Also includes group proportion, political dominance, and matching variables for all 1,202 AMAR ethnic groups.

We wish to thank the following interns and volunteers whose assistance was instrumental in helping to produce the AMAR Phase IV data: Andrew Axthelm, Nathalia Cibotti, Eric Dunford, Nevina Jakopin, Rebecca Kim, Michelle Likier, Sam Likier, Christina Lynch, Ronen Plechnin, Jamil Scott, and Delisha Thompson.

---

[4] In the description of the reconciliation process that follows in the codebook, these variables are listed in lower-case (versus upper-case for the MAR Phase V and AMAR Phase I variables) in order to draw attention to the fact that some of the variables and coding conventions for MAR Phase IV differ substantially from those in MAR Phase V and AMAR Phase I.

II.   Reconciling MAR Phases IV and V into AMAR Phase I

In 2006, under the leadership of MAR Project Director Christian Davenport, MAR project staff conducted a review of the approximately 400 variables that had been part of the various phases of the MAR project, with special attention to the 125 variables that were the focus of Phase IV. Project staff identified several categories of variables: those most frequently used in scholarly analysis; those somewhat used in scholarly analysis; and those only rarely used or not used at all in scholarly analysis. Additionally, project staff noted which variables were central to multiple theories of ethnic conflict. Based on this analysis, a total of 71 variables were selected as being "core" variables. Of the "core" variables, some, including the most frequently used variables in past analyses, were not changed from previous phases of the MAR dataset for the Phase V data. Other variables were reformulated for Phase V to facilitate either collection or statistical analysis of the data. Finally, several new variables were added to the dataset for Phase V.

Based on this reformulation, the variables in the MAR Phase V codebook were noted as falling into one of four categories. These categories apply to the variables in the AMAR Phase I codebook.

* *Category 1:* Variable was unchanged from MAR Phase IV to Phase V, and subsequently, for AMAR Phase I. Variable name, variable levels, and specification of levels are the same. Coding of these variables was considered unchanged for Phase V, and subsequently, is considered unchanged for AMAR Phase I.   This category of variables was not reconciled for AMAR Phase I because these variables did not change from Phase IV to Phase V.

** *Category 2*: Variable levels were re-specified from MAR Phase IV to Phase V and reconciled for AMAR Phase I. The variable name and levels were largely unchanged from Phase IV to Phase V. However, levels were better specified, with more specific information to denote different coding levels. This category also includes variables where the levels have been shifted.   Coding of these variables for Phase V was not considered unchanged with Phase IV data, although they only required minor changes to be considered continuous. This category of variables was reconciled for AMAR Phase I.

*** *Category 3*: Variable was reformulated from MAR Phase IV to Phase V. The reformulated Phase V variables measure concepts previously measured by other variables in Phase IV. However, they were reformulated in Phase V in order to facilitate either collection or analysis of the data. Coding of these variables for Phase V was not considered unchanged with Phase IV data.   This category of variables was reconciled for AMAR Phase I.

**** *Category 4:* New variable for MAR Phase V. These Phase V variables measure concepts not previously measured by other variables in Phase IV. However, the concepts measured were identified as being of interest by users of MAR data and project staff.   This category of variables was not reconciled for AMAR Phase I because these variables were not coded in Phase IV.

In order to release a single, integrated dataset with continuous coding of the core variables for the years 1940-2006, MAR project staff systematically reviewed and reconciled category 2 and 3 variables. As a result, all re-specified category 2 variables and three re-formulated category 3 variables were reconciled. The remaining category 3 variables could not be reconciled with the Phase IV data due to funding limitations. Back-coding of category 4 variables was also not completed at this time due to funding limitations.

******Category 5**:* New variable for AMAR Phase I. These variables identify AMAR groups that were in MAR proper (MAR Phase V) and the AMAR selection bias data. The include group population estimates for all 1,202 AMAR groups and statistical weights for the Sample Frame. They also include variables that identify politically dominant groups (based on control of the executive branch of government), matching codes that compare the groups in AMAR to those in three prominent ethnic group datasets – Alesina et al. 2003, Ethnic Power Relations (versions 1 and 3), and Fearon 2003, and variables that identify groups whose names changed from MAR Phase V to AMAR Phase I, as well as groups from MAR Phase V that were merged or split for AMAR Phase I. These variables are included in the AMAR Phase I data.

CATEGORY 1 VARIABLES

As indicated above, category 1 variables were not changed from MAR Phase IV to Phase V. Since they were not changed, there was no need to reconcile these variables. As such, these variables can be considered as continuously coded for 1940-2006 in the AMAR Phase I data.   The following thirty-three variables are considered category 1 variables.

- NUMCODE
- GROUP
- CCODE
- COUNTRY
- REGION
- YEAR
- GPOP
- CPOP
- GPRO
- CUSTOM
- GROUPCON
- GC119
- AUTLOST
- AUTONEND
- TRANSYR
- SEPKIN
- POLDIS*
- ECDIS*
- CULPO1
- GOJPA
- AUTON2

- [AUTGAIN](#)
- [AUTPRO](#)
- [INTRACON](#)
- [FACTCC1](#), [FACTCC2](#), [FACTCC3](#)
- [INTERCON](#)
- [CCGROUP1](#) , [CCGROUP2](#), [CCGROUP3](#)
- [PROT**](#)
- [REB**](#)

*POLDIS and ECDIS were updated with annual coding for 1980-1989 by Dr. Victor Asal (SUNY-Albany); previously, these variables were coded for 1980 and 1985 only. The Asal data was integrated into the AMAR Phase I dataset and replaces the MAR Phase IV POLDIS and ECDIS coding for 1980-1989.

**PROT and REB were unchanged from MAR Phase IV to Phase V. However, the MAR Phase IV data includes two additional variables that coded quinquennial data for protest and rebellion from 1940-1999 – proti and rebel, respectively. These two variables were merged into the PROT and REB variables as part of the AMAR Phase I data in order to create integrated protest and rebellion variables for 1940-2006. The reconciled data codes protest and rebellion every five years from 1940-1984 and annually from 1985-2006.

CATEGORY 2 VARIABLES

The following twenty category 2 variables were re-specified from MAR Phase IV to Phase V. They were reconciled for AMAR Phase I. As such, they can be considered as continuously coded for 1940-2006 in the AMAR Phase I data.

- [LANG](#)
- [BELIEF](#)
- [RELIGS1](#)
- [RACE](#)
- [GC2](#)
- [GC7](#)
- [GC10](#)
- [GC11](#)
- [EMIG](#)
- [FACTSEV1](#), [FACTSEV2](#), [FACTSEV3](#)
- [CCGROUPSEV1](#), [CCGROUPSEV2](#), [CCGROUPSEV3](#)
- [GC6B](#) (not reconciled: see below)
- [YEARWT](#) (not reconciled; see below)
- [MAGN](#) (not reconciled; see below)
- [PRSTAT](#) (not reconciled; see below)
- [SEPX](#) (not reconciled; see below)

Note: For Phase V, GC6B includes an additional coding level – "0 – no regional base." Since the remaining coding levels are otherwise the same for Phase IV and Phase V, this variable was not reconciled for AMAR Phase I.

YEARWT, MAGN, and PRSTAT are constituent components of the AUTLOST index. In Phase IV, these variables were not coded separately. In Phase V, these variables were disaggregated and coded individually. Since they appear as distinct variables in the MAR Phase V dataset only, they were not reconciled for AMAR Phase I.

SEPX could not be reconciled without substantial additional research. Due to a lack of funding, this variable was not reconciled for AMAR Phase I.

CATEGORY 3 VARIABLES

The following four category 3 variables were changed from MAR Phase IV to Phase V. They were reconciled for AMAR Phase I. As such, they can be considered as continuously coded for 1940-2006 in the AMAR Phase I data.

- POLGR
- ECGR
- CULGR
- CULPO2

The following fiftenn category 3 variables were not reconciled as part of the AMAR Phase I data. These variables could not be reconciled from the 1945-2003 MAR Phase IV data without substantial research and will not be reconciled until future funding is obtained for this project. However, these variables were coded for 2004-2006 for MAR groups and for 1980-2006 for the 74 groups in the AMAR selection bias data. As such, they can be considered as continuously coded from 2004-2006 for MAR groups and from 1980-2006 for the selection bias group in the AMAR Phase I data.

- KINSUP
- KINMATSUP
- KINPOLSUP
- KINMILSUP
- STATSUP
- STAMATSUP
- STAPOLSUP
- STAMILSUP
- NSASUP
- NSAMATSUP
- NSAPOLSUP
- NSAMILSUP
- REPGENSIV
- REPNVIOL
- REPVIOL

CATEGORY 4 VARIABLES

The following four category 4 variables were new for MAR Phase V and coded for MAR groups from 2004-2006 only. They were also coded for 1980-2006 for the 74 groups in the AMAR selection bias data.   These variables are included in the AMAR Phase I data.

- DISPLACE
- LEGISREP
- EXECREP
- GUARREP

CATEGORY 5 VARIABLES

The following fourteen category 5 variables are new to the AMAR Phase I data. They include a variable for group population estimates for all 1,202 AMAR ethnic groups. They also include variables that identify AMAR groups that were in MAR proper (MAR Phase V) and the AMAR selection bias data, variables that identify groups whose names changed or were merged or split from MAR Phase V to AMAR Phase I, and variables that identify politically dominant groups (based on control of the executive branch of government). They also include a weighted group proportion variable that can be used for statistical analysis. In addition, other new variables include three matching variables that compare the groups in AMAR to those in three prominent ethnic group datasets – Alesina et al. 2003, Ethnic Power Relations (versions 1 and 3), and Fearon 2003. These variables are included in the AMAR Phase I data.

- GPROAMAR
- AMARWEIGHTS
- MARPROPER
- SELECTIONBIAS
- NAMECHANGED
- PREVIOUSNAME
- SPLITGROUP
- MERGEDGROUP
- ONEDOMGROUP
- ALLDOMGROUPS
- ALESINAMATCH
- EPRV1MATCH
- EPRV3MATCH
- FEARONMATCH

III.   AMAR Phase I Codebook

I. Group characteristics

A. Group Identity

i. NUMCODE*   Ethnic group case identifier (country code + group id)

ii. AMAR_GROUP*   Full name of AMAR ethnic group

iii. CCODE*   Country ID number
                           (The Correlates of War (Singer and Small) country
                           identification number

iv. COUNTRY*   Country in which the group resides

v. AMAR_REGION*   AMAR regions
                                   Asia (formerly MAR region 3)
                                   Europe (formerly MAR regions 0 and 2)
                                   India (formerly MAR region 3)
                                   Latin America and the Caribbean (formerly MAR
                                   region 7)
                                   Middle East and North Africa (formerly MAR region
                                   5)
                                   North America (formerly MAR region 0)
                                   Sub-Saharan Africa (formerly MAR region 6)

vi. YEAR*   Year of Observation

vii. MARPROPER*****   Group coded in MAR proper data (MAR Phase I-V)
                                   0   Not coded in MAR proper
                                   1   Coded in MAR proper

viii. SELECTIONBIAS*****   Group coded in AMAR selection bias data
                                   0   Not coded in selection bias data
                                   1   Coded in selection bias data

ix. NAMECHANGED*****   Group's name changed from MAR to AMAR
                                   0   Group's name did not change
                                   1   Group's name changed

x.   PREVIOUSNAME*****   Name of group as appeared in MAR proper

xi. SPLITGROUP*****   Group split from MAR to AMAR
                                   0   Group was not split into a new group
                                   1   Group was split into a new group

xii. MERGEDGROUP*****   Group merged from MAR to AMAR

        0   Group was not merged with another group to create a new group

        1   Group was merged with another group to create a new group

xiii. ONEDOMGROUP*****   One politically dominant group in the country

Identifies the ethnic group that has been the sole politically dominant ethnic group. Political dominance is coded based on control of the executive branch of government.

        0   Group has not been the sole politically dominant ethnic group

        1   Group has been the sole politically dominant ethnic group

xiv. ALLDOMGROUPS*****   Politicallly dominant groups

Identifies all groups that have been politically dominant. Political dominance is coded based on control of the executive branch of government.

        0   Group has never been politically dominant

        1   Group has been politically dominant for a least one year over time

B. Group Weights

i. AMARWEIGHTS*****   Weights used for statistical analysis of the data.

C. Group Matching

The AMAR Phase I data was matched by country and group name to groups in three prominent ethnic group datasets: Alesina et al. 2003, EPR 2010 (version 1.1) and 2012 (version 3), and Fearon 2003. These variables identify AMAR groups that were matched to these datasets. For more information about how groups were matched (e.g., whether an AMAR group was matched as a subgroup or aggregate group, whether an Alesina, EPR, or Fearon group was not included in AMAR because it doesn't meet the AMAR population threshold criteria, etc.), see the AMAR web site .

i.   ALESINAMATCH*****

AMAR group matched to ethnic group in Alesina et al. 2003 data

        0   AMAR group not matched to Alesina et al. group

| | 1 | AMAR group matched to Alesina et al. group |

ii. **EPRV1MATCH*******

> AMAR group matched to ethnic group in Ethnic Power Relations (EPR) v1 2010 data
> 0   AMAR group not matched to EPR v1 group
> 1   AMAR group matched to EPR v1 group

iii. **EPRV3MATCH*******

> AMAR group matched to ethnic group in Ethnic Power Relations (EPR) v1 2010 data
> 0   AMAR group not matched to EPR v3 group
> 1   AMAR group matched to EPR v3 group

iv. **FEARONMATCH*******

> AMAR group matched to ethnic group in Fearon 2003 data
> 0   AMAR group not matched to Fearon 2003 group
> 1   AMAR group matched to Fearon 2003 group

D. Group Population

Group population estimates include at least one year of data. These variables are updated periodically, rather than annually. Group population estimates frequently differ based on sources and how group boundaries are determined. The average of disparate sources is generally reported.

i. GPOP*   Group's population

ii. CPOP*   Country's population

iii. GPRO*   Group proportion of country population (to 4 decimal places, e.g., .1729)

iv. GPROAMAR***** Group proportion of country population to 4 decimal places, e.g., .1729) for all 1,202 AMAR groups
Based on population estimates for 2001 or 2007 only.

E. Measures of Distinctiveness

These are static variables and are generally the same for all years being coded. If a change in codes is noted from one update to another, further research is conducted to determine the turning point year. The comparison group for the following indicators is the largest (plurality or majority) ethnic group in the state.

i.   LANG** (RECONCILED VARIABLE)
          Different language group
          0   Linguistic assimilation with plurality group
                Group has same language as plurality (e.g., Arab Shi'a and Arab
                Sunni in Iraq) or most of the group (>90%) no longer speaks
                native language but has assimilated to language of dominant
                group (e.g., German Americans, native Hawaiians).
          1   Group speaks multiple languages, at least one different from
                plurality group
                Members of group speak different languages (e.g., Southern
                Sudanese in Sudan) or part of group is assimilated to plurality
                but part still speaks native language.
          2   Group speaks primarily one language, different from plurality
                group Plurality of group speaks the same language AND it is
                different from plurality group language (e.g., Kurds in Turkey
                or Iraq)
          -99   No basis for judgment
          -88   Unknown

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

ii.   CUSTOM*
          Different group customs (marriage, family, dress, etc.)
          0   Same social customs as majority
          1   Different social customs from majority
                At least a significant minority (>25%) of ethnic group population
                follows different social customs from the majority group in the
                country. Examples of different social customs include polygamy vs.
                monogamy; nomadic or semi-nomadic lifestyles vs. settled, etc. Also
                coded here are groups that are nominally the same religion as the
                plurality group but have significantly different practice (e.g.,
                incorporation of traditional religion into Islamic or Christian
                practice).
          -99   No basis for judgment

iii.   BELIEF** (RECONCILED VARIABLE)
          Different group religion
          0   Same religion as plurality
                The plurality of the group (>90%) is the same religion as the
                plurality group (e.g., Basques in Spain are Roman Catholic like
                most Spaniards).
          1   Different sect within same religion as plurality
                Same major religion, but different sect (e.g., Roman Catholic
                Irish in Northern Ireland vs. Protestants; Sunni Arabs in Iraq
                vs. Shi'a Arabs in Iraq).

2 Different religion
  Totally distinct religions (e.g., Palestinians are Muslim or Christian vs. Jewish Israelis; Christians in Iran vs. Shi'a Muslims).
-99 No basis for judgment
-88 Unknown
-77 Multiple Sects; some different from dominant group.

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

 iv. RELIGS1** (RECONCILED VARIABLE)
  Specific religion: Plurality religion of group
  1 Roman Catholic
  2 Orthodox
  3 Protestant
  4 Other Christian sect
  5 Sunni Islam
  6 Shi'a Islam
  7 Other Islamic sect
  8 Buddhist
  9 Animist
  10 Other
  -99 No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

 v. RACE** (RECONCILED VARIABLE)
  Different physical appearance
  0 No physical differences in appearance
  1 Physically distinguishable subtype of same racial stock
   (e.g., Korean vs. Japanese; Greek vs. German)
  2 Different racial stock from the dominant group with substantial intermixture
   (e.g., Chinese v. Malay; Black or Indio v. European)
  3 Different racial stock, little or no intermixture
  -99 No basis for judgment
  -88 Unknown

For the variable RACE (which like most of the Group Characteristics variables is a way to measure how easily distinguishable members of the minority are from members of the majority), AMAR uses the concept of continental (or geographic) race. This identifies the geographic origins of the group. Evolutionary biology has found little support for the concept of race in a strictly genetic sense; however, as a social construct with some basis in differences in physical appearance, it is useful. The five racial types AMAR uses are:

16

Asiatic
 Mongolian, Chinese, Japanese, Malay, SE Asian peoples, Polynesians, Micronesians

African
Indigenous peoples of sub-Saharan Africa

Europoid
European peoples, indigenous peoples of North Africa (Berbers, Egyptians), Middle Eastern peoples (Arabs, Persians), some Central and South Asian peoples (Pashtuns, Baluchis)

Indio/Amerindian
Indigenous peoples of North and South America
Oceanic/Pacific
Melanesians, Papuans, aboriginals of Australia and New Zealand
The main sources used to develop these guidelines are as follows:

Nei, Masatoshi and Arun K Roychoudhury. 1993. "Evolutionary Relationships and Human Populations on a Global Scale." *Molecular Biology and Evolution*. 10:5. 927-943,

Zhivotovsky, Lev A., Noah A. Rosenberg, and Marcus W. Feldman. 2003. "Features of Evolution and Expansion of Modern Humans, Inferred from Genomewide Microsatellite Markers." *American Journal of Human Genetics*. 72:1171-1186.

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

F. Group Concentration

These variables are relatively static and are updated periodically, rather than annually. If a change in codes is noted from one update to another, further research is conducted to determine the year of the turning point.

i. GROUPCON*

Group spatial distribution
0 Widely dispersed
1 Primarily urban or minority of group lives in one region
2 Majority of group lives in one region, others dispersed (at least 50% of group lives in one region, but not more than 75%)
3 Concentrated in one region (at least 75% of group lives in one region)
-99 No basis for judgment

ii. GC119*

Urban/rural distribution
1 Mainly rural (>80%)

        2   Mostly rural (60-80%)
        3   Mixed urban/rural
        4   Mostly urban (60-80%)
        5   Mainly urban (>80%)
      -99   No basis for judgment

    iii.   GC2** (RECONCILED VARIABLE)
              Regional base:
              A spatially contiguous region larger than an urban area that is part of
              the country, in which 25% or more of the minority resides and in
              which the minority constitutes the predominant proportion of the
              population.
              0 = No
              1 = Yes
              -99 = No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

    iv. GC6B*
         Regional base--proportion of group members in regional base
             0   No regional base
             1   <50%
             2   50-75%
             3   >75%
           -99   No basis for judgment

    v.   GC7** (RECONCILED VARIABLE)
              Proportion of group living outside regional base
             0   No regional base
              1   >50%
              2   25-50%
              3   <25%
           -99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

    vi.   GC10** (RECONCILED VARIABLE)
              Transnational dispersion -- kindred groups
             0   The group has no close kindred across an international border
             1   The group has close kindred across a border which does not
                    adjoin its regional base (including groups that have
                    transnational kindred but not a regional base)

2   The group has close kindred in one country which adjoins its regional base

3   The group has close kindred in more than one country which adjoins its regional base

-99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

   vii.   GC11** (RECONCILED VARIABLE)

Transnational dispersion -- kindred groups in power

0   Kindred have no access to political power (no kindred abroad)

1   Kindred are outside political ruling coalition but are not barred from power

2   Kindred are in ruling coalition

3   Kindred dominate state coalition

-99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

II. Group status

A. Historical autonomy and separatism indicators

   i.   AUTLOST*

Index of lost political autonomy, based on YEARWT, MAGN, PRSTAT; AUTLOST = (MAGN+PRSTAT-1)/YEARWT.

   ii.   YEARWT**

Based on year of the most recent less of autonomy. If no loss of autonomy, then based on most recent transfer of centralized authority,

0   No history of autonomy or transfer

1   <25 years ago

2   25-49 years ago

3   50-74 years ago

4   75-99 years ago

5   >100 years ago

   iii.   MAGN**

Magnitude of change

0   No history of autonomy or transfer

1   transfer only centralized authority

2   Loss of short-term autonomy (<10 years) under colonial rule

3　　　　Loss of long-term autonomy

iv. PRSTAT**

　　　Group status prior to change
　　　0　　　　No history of autonomy or transfer
　　　1　　　　Autonomous but acephalous
　　　　　　　Groups that lack centralized authority structures (e.g.,
　　　　　　　many indigenous groups whose highest level of political
　　　　　　　structure was the village) or acted autonomously from
　　　　　　　centralized political structures.
　　　2　　　　Part of larger segment of group OR province in another
　　　　　　　state or territory
　　　　　　　Groups that were part of a larger segment with
　　　　　　　independence or autonomy (e.g., Albanians in Kosovo and
　　　　　　　Macedonia) or groups that have their own province in
　　　　　　　another state or colonial territory.
　　　3　　　　Traditional centralized authority OR autonomous region or
　　　　　　　province OR autonomous people under colonial rule
　　　4　　　　State or republic
　　　　　　　Groups that previously had independent states (e.g., Tuva
　　　　　　　in Russia).

v. AUTONEND*

　　　Year/decade/century autonomy was lost

vi. TRANSYR*

　　　Year/decade/century transferred

vii. SEPX**

　　　Separatism index
　　　0　　　　None
　　　1　　　　AUTLOST>0 but no separatist (independence/revanchist)
　　　　　　　or autonomy movements in past 50 years
　　　2　　　　Separatist or autonomy movement that persisted as an
　　　　　　　active political force for at least 5 years in the past 50
　　　　　　　years, but not in the past 25 years
　　　3　　　　Active separatist or autonomy movements in the past 25
　　　　　　　years
　　　-99　　　No basis for judgment

viii.　　SEPKIN*

　　　Active separatism among kin groups
　　　0　　　　No
　　　1　　　　Yes
　　　-99　　　No basis for judgment

B. Disadvantages

i.   EMIG** (RECONCILED VARIABLE)
>   Emigration for political or economic reasons
>   0   Condition not present
>   1   Condition minor
>   Affects less than 1% of group population
>   2   Condition of medium significance
>   Affects between 1 and 10% of group population
>   3   Condition serious
>   Affects more than 10%
>   -99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

ii. DISPLACE ****
>   Internal displacement for political or economic reasons
>   0   Condition not present
>   1   Condition minor
>   Condition is minor if affects less than 1% of group
>   population
>   2   Condition of medium significance
>   Condition is of medium significance if affects between 1
>   and 10% of group population
>   3   Condition serious
>   Condition is serious if affects more than 10%
>   -99   No basis for judgment

iii. POLDIS*

>   Political discrimination index
>   0   No discrimination
>   1   Neglect/remedial polices
>   Substantial under-representation in political office and/or
>   participation due to historical neglect or restrictions.
>   Explicit public policies are designed to protect or improve
>   the group's political status.
>   2   Neglect/no remedial policies
>   Substantial under-representation due to historical neglect
>   or restrictions. No social practice of deliberate exclusion.
>   No formal exclusion. No evidence of protective or remedial
>   public policies.
>   3   Social exclusion/neutral policy
>   Substantial under-representation due to prevailing social
>   practice by dominant groups. Formal public policies

toward the group are neutral or, if positive, inadequate to offset discriminatory social practices.

4 Exclusion/repressive policy
Public policies (formal exclusion and/or recurring repression) substantially restrict the group's political participation by comparison with other groups. (Note: This does not include repression during group rebellion. It does include patterned repression against group members when the group is not openly resisting state authority.)

-99 No basis for judgment

iv. ECDIS*

Economic discrimination index

0 No discrimination
1 Neglect/remedial polices
Significant poverty and under-representation in desirable occupations due to historical marginality, neglect, or restrictions. Public policies are designed to improve the group's material well being.

2 Neglect/no remedial policies
Significant poverty and under-representation due to historical marginality, neglect, or restrictions. No social practice of deliberate exclusion. Few or no public policies aim at improving the group's material well being.

3 Social exclusion/neutral policy
Significant poverty and under-representation due to prevailing social practice by dominant groups. Formal public policies toward the group are neutral or, if positive, inadequate to offset active and widespread discrimination.

4 Exclusion/repressive policy
Public policies (formal exclusion and/or recurring repression) substantially restrict the group's economic opportunities by contrast with other groups.

-99 No basis for judgment

v. CULPO1*

Restrictions on religion

0 No restrictions
1 Activity informally restricted
The activity is restricted by widespread but informal social practice (e.g., by discrimination against people who follow group's religious practices).

2 Activity somewhat restricted
3 Activity sharply restricted
-99 No basis for judgment

vi.   CULPO2*** (RECONCILED VARIABLE)
                    Restrictions on use of language or language instruction
                    0   No restrictions
                    1   Activity informally restricted.
                            The activity is restricted by widespread but informal social
                            practice (e.g., by discrimination against people who speak the
                            group's language)
                    2   Activity somewhat restricted
                    3   Activity sharply restricted
                    -99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

For CULPO1 and CULPO2, the following guidelines apply:
1. These items are included only if the communal group is treated differently than
    others. These items are not included if restrictions apply to everyone in the
    population because of the type of regime or other factors.
2. Public restrictions that apply to all citizens because they are necessary for the
    common good are not restrictions even if they violate the religious norms of the
    communal group, (e.g. requirements that families have only one child, or that all
    children be vaccinated).
3. Lack of public support for group cultural activities is not a restriction unless public
    support is provided to similar activities by other groups.
4. Discrimination anywhere in the country is sufficient to code a discrimination
    variable--even if the group is not discriminated against in the region where they are
    concentrated.
5. These variables are often time invariant and are inferred to exist if they existed in
    the recent past. However, variable values are changed based on events that have
    altered the situation.

C. Group organization and representation

The Minorities at Risk Organizational Behavior (MAROB) data includes additional
information on organizations for selected ethnic groups. For more information please see
the data page of the MAR website: http://www.cidcm.umd.edu/mar/data.asp.

i.   GOJPA*
                    Group organization for joint political action
                    0   No political movements or organizations represent group
                            interests
                    1   Group interests promoted by umbrella organizations
                    2   Group interests promoted by one or more conventional
                            political parties or movements
                    3   Group interests promoted mainly by conventional
                            movements or parties but also by militant organizations

with limited support
4   Group interests promoted mainly by militant organizations
        but also by some conventional organizations
5   Group interests promoted only by militant organizations
-99   No basis for judgment

For GOJPA, the following guidelines apply:
1. For each year, the highest type of group representation is reported (none, umbrella, conventional, militant).
2. The variable reports whether the organization acts on behalf of the minority group and whether it does so from within or outside the country. Actions by international organizations/international nongovernmental organizations are not reported (e.g., Cultural Survival, development organizations, Amnesty International).
3. Minority-based NGOs, (e.g., community organizations working for health care, education, other social services, etc.) are reported as "group interests are promoted by umbrella organizations..."
4. If there is an even split between militant and conventional organizations, the higher one is reported (militant).
    Conventional movements and parties are those that rely mainly on non-coercive political techniques such as organization-building; education and consciousness-raising symbolic action; participating in electoral politics; interest representation to officials; organizing peaceful protests, etc.
    Militant movements and parties are those that rely substantially on coercive political techniques such as obtaining funds, supplies, and members by use or threat of force; use of threats and violence against state officials and rival organizations; guerrilla and civil war; defending and administering rebel-held zone, etc.

ii.   AUTON2*
                Group autonomy status
                Does group have administrative autonomy (i.e., control of political and bureaucratic structures in an autonomous region). Autonomy must be legally recognized by the home government.
                0   No
                1   Yes
                -99   No basis for judgment

iii. AUTGAIN*
                Year group gained autonomy

iv.   AUTPRO*
                Percentage of group in autonomous region

v.   LEGISREP****
                Group representation in legislative branch of central government
                0   No
                1   Yes

Representation may be through individual group members who belong to non-ethnically based parties or by representation through ethnically based parties.

-99   No basis for judgment

vi. EXECREP****

Group representation in executive branch of central government

0   No

1   Yes

Representation may be through individual group members who belong to non-ethnically based parties or by representation through ethnically based parties.

-99   No basis for judgment

vii. GUARREP****

Group is guaranteed representation in central government

0   No

1   Yes

Guaranteed positions in the central government (e.g., appointed positions in cabinet, appointed positions in legislature, guaranteed elected positions in legislature, etc.)

-99   No basis for judgment

D. Grievances

For each type of grievance, the HIGHEST level of grievance expressed by group representatives is reported (e.g., if the majority of a group desires autonomy but a radical faction desires independence, the code under POLGR is 3, NOT 2). Values are based on statements and actions by group leaders and members or observations of grievances by third parties.

i.   POLGR*** (RECONCILED VARIABLE)

Highest level of political grievance

0   No political grievances expressed

1   Political grievances focused on ending discrimination or creating or strengthening remedial policies

2   Political grievances focused on creating or strengthening autonomous status

3   Political grievances focused on creating separate state for group or revanchist change in borders (union with kindred requires revanchist change in borders).

-99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see here and here in Appendix II.

ii.   ECGR*** (RECONCILED VARIABLE)

             Highest level of economic grievance

             0   No economic grievances expressed

             1   Economic grievances focused on ending discrimination

             2   Economic grievances focused on creating or strengthening remedial policies

             -99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

iii.   CULGR*** (RECONCILED VARIABLE)

             Highest level of cultural grievance

             0   No cultural grievances expressed

             1   Cultural grievances focused on ending discrimination

             2   Cultural grievances focused on creating or strengthening remedial policies

             -99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

III.   External support

A. Kindred group support

Kindred group support includes support from diaspora members (e.g., Kurds from Turkey now working in Germany are part of the Kurdish diaspora) and support from close kindred outside the country (e.g., Albanians in Albania supporting Kosovar Albanians or Albanians in Macedonia).

i.   KINSUP***

             Any kindred group support

             0   No

             1   Yes

                 For each year in which a subtype of support is reported.

             -99   No basis for judgment

ii. KINMATSUP***

             Kindred group material, non-military, support

             Any financial or material support that is not used for military purposes. Examples include humanitarian aid (in the form of money or goods); development aid (in the form of money or goods); funding for civic, cultural or political associations, etc.

             0   No

             1   Yes

<pre>-99   No basis for judgment</pre>

iii. KINPOLSUP***

        Kindred group political support
        Reported when members of kindred groups provide members for
        monitoring elections, ceasefires, etc. on behalf of group members;
        when exile or diaspora members vote for ethnically based parties in
        expatriate elections; when members of kindred groups hold protests
        on behalf of group members; when members of kindred groups host
        talks/ negotiations between ethnic kin and government of home
        country.

0	No
1	Yes
-99	No basis for judgment

iv. KINMILSUP***

        Kindred group military support
        Reported when kindred group members provide funds for military
        supplies, sanctuaries or safe havens for armed fighters; military
        training in exile; advisory military personnel; active combat units; or
        cross-border raids or rescue missions for ethnic group.

0	No
1	Yes
-99	No basis for judgment

B. Foreign state and state-led actor (IGOs) support

State support includes support given by the governments of individual states (e.g., the
United States, Sweden, South Africa) AND support given by intergovernmental
organizations and their agencies (e.g., the United Nations, the European Union, NATO, the
World Bank, the World Food Programme, UNICEF, etc.).

i. STASUP***

        Any foreign state or IGO support

0	No
1	Yes
	For each year in which a subtype of support is reported.
-99	No basis for judgment

ii. STAMATSUP***

        Foreign state/IGO material, non-military, support
        Any financial or material support that is not used for military
        purposes (e.g. humanitarian aid (in the form of money or goods);
        development aid (in the form of money or goods); funding for civic,
        cultural or political associations; etc.).

0	No

1	Yes
-99	No basis for judgment

iii. STAPOLSUP***

Foreign state/IGO political support
Reported when state actors provide personnel for monitoring elections, ceasefires, etc. on behalf of group members; when state actors host talks/ negotiations between ethnic group and government of home country; when state actors levy sanctions against government for treatment of ethnic minority.

0	No
1	Yes
-99	No basis for judgment

iv. STAMILSUP***

Foreign state/IGO military support
Reported when state actors provide funds for military supplies, sanctuaries or safe havens for armed fighters; military training in exile; advisory military personnel; active combat units; or cross-border raids or rescue missions for ethnic group.

0	No
1	Yes
-99	No basis for judgment

C. Non-state actor (non-kindred) support

Non-state actors are predominantly non-governmental organizations (e.g., the Red Cross, Amnesty International, Gates Foundation), but also include prominent individuals (e.g., Jimmy Carter, Bono), religious organizations (e.g., the Catholic Church), and transnational criminal and terrorist networks (e.g., al Qaeda). These variables do not include cases when the non-state actor is predominantly made up of kindred group members – those cases would be coded under the appropriate kindred support variable above.

i. NSASUP***

Any non-state actor support

0	No
1	Yes
	For each year in which support was recorded.
-99	No basis for judgment

ii. NSAMATSUP***

Non-state actor material, non-military, support
Any financial or material support that is not used for military purposes. Examples include humanitarian aid (in the form of money

or goods); development aid (in the form of money or goods); funding for civic, cultural or political associations; etc.

0   No
1   Yes
-99   No basis for judgment

iii. NSAPOLSUP***

Non-state actor political support

Reported when non-state actors provide personnel for monitoring elections, ceasefires, etc. on behalf of group members; when state actors host talks/negotiations between ethnic group and government of home country; when non-state actors levy sanctions against government for treatment of ethnic minority.

0   No
1   Yes
-99   No basis for judgment

iv. NSAMILSUP***

Non-state actor military support

Reported when non-state actors provide funds for military supplies, sanctuaries or safe havens for armed fighters, military training in exile, advisory military personnel, active combat units, or cross-border raids or rescue missions for ethnic group.

0   No
1   Yes
-99   No basis for judgment

IV.   Group conflict behavior

A. Intracommunal conflict

i. INTRACON*

Presence of intracommunal conflict

0   No
1   Yes

For each year in which intracommunal conflict is reported.

-99   No basis for judgment

The following variables are based on which pair of antagonists had the highest level of conflict for a given year. The highest level of conflict for any one year is reported in FACTCC1 and the name of the conflicting factions of the ethnic group is reported for that year in FACTSEV1. The second-highest level of conflict for any one year is reported in FACTCC2, and the name of the conflicting faction is reported for that year in FACTSEV2, etc. An antagonistic pair can move between variables (e.g., an antagonistic faction can be

reported in FACTCC1 for one year and in FACTCC3 for another). Only the 3 pairs of groups with the highest levels of conflict are reported.

It is important to note that the MAR Phase IV fccs1-3 variables are not ordered in descending order of severity, whereas AMAR Phase I FACTSEV1-3 are. We are noting the difference in ordering between the MAR Phase IV and AMAR Phase I and MAR Phase V data, but are leaving the variables as is.

ii. FACTCC1*
     Names of intracommunal antagonists with highest level of conflict

iii. FACTSEV1** (RECONCILED VARIABLE)
     Severity of conflict for first pair of antagonists
     0   No conflict
     1   Sporadic violent attacks
             Attacks without weapons (e.g., brawls), knives, or few
             small arms (e.g., one or two handguns).
     2   Series of bombings/assassinations
     3   Substantial rioting
     4   Sporadic armed clashes
             Attacks with multiple firearms, automatic weapons, or
             heavy weaponry (mortars, shelling, etc.)
     5   Protracted communal warfare
             More than 6 clashes a year between antagonists
     -99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

iv. FACTCC2*
     Names of intracommunal antagonists with second-highest level of
     conflict

v.   FACTSEV2** (RECONCILED VARIABLE)
     Severity of conflict for second pair of antagonists
     0   No conflict
     1   Sporadic violent attacks
             Attacks without weapons (e.g., brawls), knives, or few
             small arms (e.g., one or two handguns).
     2   Series of bombings/assassinations
     3   Substantial rioting
     4   Sporadic armed clashes
             Attacks with multiple firearms, automatic weapons, or
             heavy weaponry (mortars, shelling, etc.)
     5   Protracted communal warfare
             More than 6 clashes a year between antagonists

30

-99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

    vi. FACTCC3*

                  Names of intracommunal antagonists with third-highest level of
                  conflict

    vii. FACTSEV3** (RECONCILED VARIABLE)

                  Severity of conflict for third pair of antagonists

0	No conflict
1	Sporadic violent attacks
	Attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns).
2	Series of bombings/assassinations
3	Substantial rioting
4	Sporadic armed clashes
	Attacks with multiple firearms, automatic weapons, or heavy weaponry (mortars, shelling, etc.)
5	Protracted communal warfare
	More than 6 clashes a year between antagonists
-99	No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

B. Intercommunal conflict

For intercommunal conflict, the following are noted:
1. For each year, open hostilities between the minority group and other communal groups are reported.
2. Open conflicts with other minorities and the majority or dominant group are reported, and not conflicts with the state or with dominant groups exercising state power except when the state cannot control such groups.
3. The minority being coded usually is the target of attacks of the kinds specified below, but in some instances may initiate them.

    i. INTERCON*

                  Presence of intercommunal conflict

0	No
1	Yes
	For each year in which intercommunal conflict is reported.
-99	No basis for judgment

The following variables are based on which antagonist had the highest level of conflict for a given year. The highest level of conflict for any one year is reported in CCGROUP1SEV and the name of the antagonist is reported for that year in CCGROUP1. The second-highest level of conflict for any one year is reported in CCGROUP2SEV, and the name of the antagonist is reported for that year in CCGROUP2 etc. An antagonist can move between variables (e.g., one organization can be recorded in CCGROUP1 for one year and in CCGROUP3 for another). If more than 3 antagonists in a given year, mention in notes but do not code. Only the 3 pairs of groups with the highest levels of conflict are reported.

It is important to note that MAR Phase IV cc1x-cc3x and gcc1-gcc3 are not ordered in descending order of severity, whereas AMAR Phase I CCGROUPSEV1-3. We are noting the difference in ordering between the MAR Phase IV and AMAR Phase I and MAR Phase V data, but leaving these variables as is.

    ii. CCGROUP1*
                Name of group with highest level of conflict

    iii. CCGROUP1SEV** (RECONCILED VARIABLE)
                Level of conflict with CCGROUP1

0	No conflict
1	Individual acts of harassment, no fatalities
2	Political agitation, campaigns urging authorities to impose restrictions on group
3	Sporadic violent attacks by gangs or other small groups, some fatalities Attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns) involving fewer than 20 people.
4	Anti-group demonstrations, rallies, marches
5	Communal rioting, armed attacks Attacks with multiple firearms, automatic weapons, or heavy weaponry (mortars, shelling, etc.) OR attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns) involving more than 20 people
6	Communal warfare More than 6 clashes a year between antagonists
-99	No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

    iv. CCGROUP2*
                Name of group with second-highest level of conflict

v. CCGROUP2SEV** (RECONCILED VARIABLE)

Level of conflict with CCGROUP2

0   No conflict

1   Individual acts of harassment, no fatalities

2   Political agitation, campaigns urging authorities to impose restrictions on group

3   Sporadic violent attacks by gangs or other small groups, some fatalities

Attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns) involving fewer than 20 people.

4   Anti-group demonstrations, rallies, marches

5   Communal rioting, armed attacks

Attacks with multiple firearms, automatic weapons, or heavy weaponry (mortars, shelling, etc.) OR attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns) involving more than 20 people

6   Communal warfare

More than 6 clashes a year between antagonists

-99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

vi. CCGROUP3*

Name of group with second-highest level of conflict

vii. CCGROUP3SEV** (RECONCILED VARIABLE)

Level of conflict with CCGROUP3

0   No conflict

1   Individual acts of harassment, no fatalities

2   Political agitation, campaigns urging authorities to impose restrictions on group

3   Sporadic violent attacks by gangs or other small groups, some fatalities

Attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns) involving fewer than 20 people.

4   Anti-group demonstrations, rallies, marches

5   Communal rioting, armed attacks

Attacks with multiple firearms, automatic weapons, or heavy weaponry (mortars, shelling, etc.) OR attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns) involving more than 20 people

6   Communal warfare

More than 6 clashes a year between antagonists

-99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

C. Protest

Protest and rebellion follow these guidelines:
1. Protest and rebellion initiated by organizations that claim to represent the group's interests and directed against governments that claim to exercise authority over the group is reported for each year in which it occurs.
2.   The de facto government is recognized for the purposes of these data. "Government" is defined as the body that exercises authority/control over the majority of the country. For example, Taliban was the de facto government in Afghanistan in 2000 even though it was not recognized by the world community.
3.   Protests on behalf of the group that take place outside of the group's home country are not included.
4. The most serious manifestation of each type of protest or rebellion is reported for each year.
5. Positive evidence is used to report these variables. These data do not assume that action from one year carries over into the next.
6. Protest and rebellion are distinct and may occur without the other.
7. General protests or rebellions (i.e. those carried out by the general populace as opposed to only group members) are reported if a) group members are present in substantial numbers, and b) the anti-regime action includes issues of particular concern to the group.
8. General protest or rebellion severity counts all participants, not just group members.

i.   PROT* (RECONCILED VARIABLE)
            Protest
            0   None reported
            1   Verbal opposition
                  Public letters, petitions, posters, publications, agitation, court action etc). Requests by a minority-controlled regional group for independence.
            2   Symbolic resistance
                  Sabotage, symbolic destruction of property OR political organizing activity on a substantial scale (e.g. sit-ins, blockage of traffic). Mobilization for autonomy/secession by a minority-controlled regional government.
            3   Small demonstrations
                  A few demonstrations, rallies, strikes, and/or riots, the largest of which has total participation of less than 10,000
            4   Medium demonstrations
                  Demonstrations, rallies, strikes, and/or riots, the largest of

34

which has total participation between 10,000 and 100,000
5   Large demonstrations
Demonstrations, rallies, strikes, and/or riots, , the largest of which has total participation over 100,000
-99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

D. Rebellion

i. REB* (RECONCILED VARIABLE)

Rebellion
0   None reported
1   Political banditry, sporadic terrorism (fewer than 6 events)
2   Campaigns of terrorism (more than 6 events)
3   Local rebellions
Armed attempts to seize power in a locale. If they prove to be the opening round in what becomes a protracted guerrilla or civil war during the year being coded, code the latter rather than local rebellion.
4   Small-scale guerrilla activity
Includes all three of the following traits:
- fewer than 1000 armed fighters
- sporadic armed attacks (less than 6 reported per year)
- attacks in a small part of the area occupied by the group or in one or two other locales
5   Intermediate guerrilla activity
Includes one or two of the defining traits of large-scale activity and one or two of the defining traits of small-scale activity
6   Large-scale guerrilla activity
Includes all three of the following:
- more than 1000 armed fighters
- frequent armed attacks (more than 6 reported per year)
- attacks affecting large part of the area occupied by group
7   Civil war
Protracted civil war fought by rebel military
Has all the characteristics of large-scale guerilla activity, plus rebels control large scale base areas that are secure over time
-99   No basis for judgment

For a detailed description of how this variable was reconciled, see Appendix I.
For STATA code to reconcile this variable, see Appendix II.

E. Government repression of group

Government repression follows these guidelines:
1. Each variable is reported at the highest level of repression directed at the relevant part of the group's population.
2. The de facto government is recognized here. "Government" is defined as the body that exercises authority/control over the majority of the country. For example, Taliban was the de facto government in Afghanistan even though it was not recognized by the world community.
3. These tactics may be used by any government agencies, at any level, including but not limited to the military, police, and special security services.
4. Only those actions that are carried out are reported. Threats of action are not reported.

i. REPGENCIV***

        Repression of group civilian populations (those not engaging in violent or nonviolent political activities)

    0   None reported

    1   Surveillance
        e.g., domestic spying, wiretapping, etc.

    2   Harassment/containment
        e.g., saturation of police/military presence; militarized checkpoints targeting members of groups; curfews, states of emergency

    3   Nonviolent coercion
        e.g., arrests, show-trials, property confiscation, exile/deportation

    4   Violent coercion, short of killing
        e.g., forced resettlement, torture

    5   Violent coercion, killing
        e.g., systematic killings, ethnic cleansing, reprisal killings

    -99   No basis for judgment

ii. REPNVIOL***

        Repression of group members engaged in nonviolent collective action (e.g., politicians, human rights leaders, nonviolent protesters, etc.)

    0   None reported

    1   Surveillance:
        e.g., domestic spying, wiretapping, etc.

    2   Harassment/containment
        e.g., saturation of police/military presence; militarized checkpoints targeting members of groups; curfews, states of emergency, closing down political publications/offices

    3   Nonviolent coercion
        e.g., arrests, show-trials, property confiscation, exile/deportation

4 Violent coercion, short of killing
e.g., forced resettlement, torture, non-lethal force used against protesters
5 Violent coercion, killing
e.g., systematic killings, ethnic cleansing, reprisal killings, lethal force used against protesters
-99 No basis for judgment

iii. REPVIOL***

Repression of group members engaged in violent collective action (e.g., guerrillas, rioters)
0 None reported
1 Surveillance:
e.g., domestic spying, wiretapping, etc.
2 Harassment/containment
e.g., saturation of police/military presence; militarized checkpoints targeting members of groups; curfews, states of emergency
3 Nonviolent coercion
e.g., arrests, show-trials, property confiscation, exile/deportation
4 Violent coercion, short of killing
e.g., forced resettlement, torture
5 Violent coercion, killing
e.g., systematic killings, ethnic cleansing, reprisal killings, military campaigns against rebels
-99 No basis for judgment

IV. Appendix I – Description of reconciliation process

LANG (RECONCILED VARIABLE)
Different language group
- 0   Linguistic assimilation with plurality group
  Group has same language as plurality (e.g., Arab Shi'a and Arab Sunni in Iraq) or most of the group (>90%) no longer speaks native language but has assimilated to language of dominant group (e.g., German Americans, native Hawaiians).
- 2   Group speaks multiple languages, at least one different from plurality group
  Members of group speak different languages (e.g., Southern Sudanese in Sudan) or part of group is assimilated to plurality but part still speaks native language.
- 2   Group speaks primarily one language, different from plurality group
  Plurality of group speaks the same language AND it is different from plurality group language (e.g., Kurds in Turkey or Iraq)
- -99   No basis for judgment
- -88   Unknown

Description of reconciliation
- AMAR Phase I LANG 0 = MAR Phase V LANG 0 = MAR Phase IV lang 3
- AMAR Phase I LANG 1 = MAR Phase V LANG 1 = MAR Phase IV lang 2
- AMAR Phase I LANG 2 = MAR Phase V LANG 2 = MAR Phase IV lang 1
- AMAR Phase I LANG -99 = MAR Phase V LANG -99 = MAR Phase IV lang -99
- AMAR Phase I LANG -99 = MAR Phase V LANG N/A = MAR Phase IV lang 0

Note: According to the MAR Phase IV codebook, MAR Phase IV lang 0 is defined as unknown. However, this does seem correct, as the language of some groups can be easily determined. For example, the Catholics in Northern Ireland are coded as MAR Phase IV lang 0, but sources indicate that this group has the same language as the plurality. Since language is relatively static, it is likely that the MAR Phase V values can be used to update the MAR Phase IV data. However, we are leaving this up to the discretion of researchers and recoding these cases as -88. See the appendix for a list of these groups.

Further, while MAR Phase IV lang 2 (group speaks multiple languages) is reconciled as AMAR Phase I and MAR Phase V LANG 1 (group speaks multiple languages, at least one different from plurality group), there are five groups that are not coded LANG 1 in the MAR Phase V data. It is possible that the language of these groups changed over time and the 2004-2006 update reflects this. Since it is not clear when their languages changed, we are noting these cases but leaving them coded as LANG 1 for 1940-2003. This pertains to the following groups: Indigenous Peoples in El Salvador coded 0 for 2004-2006; Roma in Slovakia coded 2 for 2004-2006; Bakongo in Angola coded 2 for 2004-2006; Ovimbundu in Angola coded 0 for 2004-2006; and Biharis in Bangladesh coded 2 for 2004-2006.

Finally, while MAR Phase IV lang 1 (group speaks same language) is reconciled as AMAR Phase I and MAR Phase V LANG 2 (group speaks primarily one language, different from plurality group), not all groups that were coded MAR Phase IV lang1 are coded LANG 2 in the 2004-2006 update. Specifically, of the 143 groups that were coded MAR Phase IV lang 1, 38 groups are coded as LANG 2 for 2004-2006 and 91 are coded as a value other than LANG 2 for 2004-2006. For these 91 groups, it is possible that some groups changed their language over time. It is also likely that these groups not coded LANG 2 for 2004-2006 because Phase IV lang 1 does not specify whether a plurality of the ethnic minority group speaks the language, whereas Phase V LANG 2 does. Since Phase IV lang 1 doesn't distinguish whether a plurality of the group speaks the language, groups that are coded Phase IV lang 1 for 1940-2003 may be coded LANG 1 or LANG 2 for 2004-2006. In either case, it is possible that the Phase V 2004-2006 values can be used to update the 1940-2003 data. However, we are leaving it up to the discretion of researchers to update the 1940-2003 coding with the 2004-2006 values.

List of groups coded LANG -88 for AMAR Phase IV:
Jews in Argentina; Shi'is in Bahrain; Hindus in Bangladesh; Mons in Burma; Rohingya (Arakanese) in Burma; Hutus in Burundi; Tutsis in Burundi; Hui Muslims in China; Blacks in Colombia; Slovaks in Czech Republic; Blacks in Ecuador; Copts in Egypt; Adzhars in Georgia; Acehnese in Indonesia; Chinese in Indonesia; Baha'is in Iran; Shi'is in Iraq; Sunnis in Iraq; Palestinians in Jordan; Luo in Kenya; Honamese in Korea, South; Druze in Lebanon; Maronite Christians in Lebanon; Palestinians in Lebanon; Shi'is in Lebanon; Sunnis in Lebanon; Black Moors in Mauritania; Ahmadis in Pakistan; Blacks (Afro-Peruvians) in Peru; Hutus in Rwanda; Tutsis in Rwanda; Shi'is in Saudi Arabia; Issaq in Somalia; Asians in South Africa; Jurassians in Switzerland; Alawi in Syria; Zanzibaris in Tanzania; Chinese in Thailand; Catholics in Northern Ireland; Scots in the United Kingdom; African-Americans in the United States of America; Blacks in Venezuela; and Chinese in Vietnam.

List of groups coded LANG 2 for 1940-2003 but LANG 1 for 2004-2006 for AMAR Phase IV:
Hispanics in United States of America; Quebecois in Canada; French Canadians in Canada; Zapotecs in Mexico; Black Karibs in Honduras; Indigenous Peoples in Honduras; Antillean Blacks in Costa Rica; Blacks in Panama; Indigenous Peoples in Panama; Indigenous Highland Peoples in Ecuador; Indigenous Peoples in Chile; Catalans in Spain; Turks in Germany; Hungarians in Slovakia; Greeks in Albania; Albanians in Macedonia; Serbs in Macedonia; Muslims in Greece; Turks in Bulgaria; Roma in Bulgaria; Gaguaz in Moldova; Magyars (Hungarians) in Romania; Chechens in Russia; Karachay in Russia; Roma in Russia; Ingush in Russia; Buryat in Russia; Tuvinians in Russia; Yakut in Russia; Poles in Lithuania; Crimean Tatars in Ukraine; Abkhazians in Georgia; Ossetians (South) in Georgia; Tuareg in Mali; Dioals in Casamance in Senegal; Kewri in Mauritania; Fulani in Guinea; Susu in Guinea; Temne in Sierra Leone; Ashanti in Ghana; Ewe in Ghana; Ewe in Togo; Kabre in Togo; Westerners in Cameroon; Ibo in Nigeria; Ogani in Nigeria; Yoruba in Nigeria; Ijaw in Nigeria; Lari in the Republic of Congo; M'boshi in the Republic of Congo; Luba in the Democratic Republic of Congo; Lunda, Yeke in the Democratic Republic of Congo; Ngbandi in the Democratic Republic of Congo; Tutsis in the Democratic Republic of Congo; Acholi in Uganda; Kikuyu in Kenya; Kalenjin in Kenya; Luhya in Kenya; Kisii in Kenya; Afars in Eritrea; Lozi in Zambia; Xhosa in South Africa; Zulus in South Africa; Azerbaijanis in Iran;

39

Bakhtiari in Iran; Kurds in Iran; Turkmen in Iran; Arabs in Iran; Pashtuns in Afghanistan; Tajiks in Afghanistan; Uzbeks in Tajikistan; Uzbeks in Kyrgyzstan; Tajiks in Uzbekistan; Germans in Kazakhstan; Turkmen in China; Tibetans in China; Kashmiris in India; Muslims in India; Sikhs in India; Mizos in India; Tripuras in India; Assamese in India; Lhotshampas in Bhutan; Baluchis in Pakistan; Hindus in Pakistan; Sindhis in Pakistan; Shans in Burman; Chinese in Malaysia; Malays in Singapore; Maori in New Zealand; and Fijians in Fiji.

For STATA code to reconcile this variable, see Appendix II.

BELIEF (RECONCILED VARIABLE)
Different group religion
- 0   Same religion as plurality
  The plurality of the group (>90%) is the same religion as the plurality group (e.g., Basques in Spain are Roman Catholic like most Spaniards).
- 1   Different sect within same religion as plurality
  Same major religion, but different sect (e.g., Roman Catholic Irish in Northern Ireland vs. Protestants; Sunni Arabs in Iraq vs. Shi'a Arabs in Iraq).
- 2   Different religion
  Totally distinct religions (e.g., Palestinians are Muslim or Christian vs. Jewish Israelis; Christians in Iran vs. Shi'a Muslims).
- -99   No basis for judgment
- -88   Unknown
- -77   Multiple Sects; some different from dominant group.

Description of reconciliation
- AMAR Phase I BELIEF 0 = MAR Phase V BELIEF 0, MAR Phase IV belief N/A
- AMAR Phase I BELIEF 1 = MAR Phase V BELIEF 1, MAR Phase IV belief 1
- AMAR Phase I BELIEF 2 = MAR Phase V BELIEF 2 = MAR Phase IV belief 3
- AMAR Phase I BELIEF -99 = MAR Phase V BELIEF -99, MAR Phase IV belief -99
- AMAR Phase I BELIEF -88 = MAR Phase V BELIEF N/A, MAR Phase IV belief 0
- AMAR Phase I BELIEF -77 = MAR Phase V BELIEF N/A, MAR Phase IV belief 2

Note: Since there is no direct match for MAR Phase IV belief 0 (unknown) and MAR Phase IV belief 2 (Multiple Sects; some different from dominant group) in MAR Phase V, these values are being recoded as -88 and -77 respectively in the AMAR Phase I data. Although the belief variable is relatively static and it is likely that the MAR Phase V values can be used to replace these codes, we are leaving this up to the discretion of researchers.

For STATA code to reconcile this variable, see Appendix II.

RELIGS1 (RECONCILED VARIABLE)

Specific religion: Plurality religion of group
1   Roman Catholic
2   Orthodox
3   Protestant
4   Other Christian sect
5   Sunni Islam
6   Shi'a Islam
7   Other Islamic sect
8   Buddhist
9   Animist
10   Other
-99   No basis for judgment

Description of reconciliation
- AMAR Phase I RELIGS1 1= MAR Phase V RELIGS1 1 = MAR Phase IV religs1 1
- AMAR Phase I RELIGS1 2 = MAR Phase V RELIGS1 2, MAR Phase IV religs1 N/A
- AMAR Phase I RELIGS1 3 = MAR Phase V RELIGS1 3 = MAR Phase IV religs1 2
- AMAR Phase I RELIGS1 4 = MAR Phase V RELIGS1 4 = MAR Phase IV religs1 3
- AMAR Phase I RELIGS1 5 = MAR Phase V RELIGS1 5 = MAR Phase IV religs1 4
- AMAR Phase I RELIGS1 6 = MAR Phase V RELIGS1 6 = MAR Phase IV religs1 5
- AMAR Phase I RELIGS1 7 = MAR Phase V RELIGS1 7 = MAR Phase IV religs1 6
- AMAR Phase I RELIGS1 8 = MAR Phase V RELIGS1 8 = MAR Phase IV religs1 7
- AMAR Phase I RELIGS1 9 = MAR Phase V RELIGS1 9 = MAR Phase IV religs1 8
- AMAR Phase I RELIGS1 10 = MAR Phase V RELIGS1 10 = MAR Phase IV religs1 9
- AMAR Phase I RELIGS1 -99 = MAR Phase V RELIGS1 -99, MAR Phase IV religs1 -99
- AMAR Phase I RELIGS1 -88 = MAR Phase V RELIGS1 N/A, MAR Phase IV religs1 0

Note: Since the MAR Phase IV Religs1 variable includes an extra category – level 2 Orthodox – RELIGS1 values of 2-9 were shifted correspondingly for AMAR Phase I. Researchers should note that since MAR Phase IV did not code Orthodox, this category only appears in the MAR Phase V and AMAR Phase I data for 2004-2006 and for 1980-2006 for the selection bias groups.

Also, although there is no 0 value in the MAR Phase IV data, 112 observations were still coded 0 for this value. Since religion is relatively static from year to year, it is likely that these groups can be recoded based on their values in the MAR Phase V data. However, we are leaving this decision up to researchers and recoding these cases as -88. This applies to the following groups: Other Indigenous Peoples in Mexico; Zapotecs in Mexico; Lari in the Republic of Congo; M'boshi in the Republic of Congo; Uzbeks in Tajikistan; and Tajiks in Uzbekistan.

For STATA code to reconcile this variable, see Appendix II.

RACE (RECONCILED VARIABLE)

Different physical appearance

0	No physical differences in appearance
1	Physically distinguishable subtype of same racial stock (e.g., Korean vs. Japanese; Greek vs. German)
2	Different racial stock from the dominant group with substantial intermixture (e.g., Chinese v. Malay; Black or Indio v. European)
3	Different racial stock, little or no intermixture
-99	No basis for judgment
-88	Unknown

For the variable RACE (which like most of the Group Characteristics variables is a way to measure how easily distinguishable members of the minority are from members of the majority), AMAR uses the concept of continental (or geographic) race. This identifies the geographic origins of the group. Evolutionary biology has found little support for the concept of race in a strictly genetic sense; however, as a social construct with some basis in differences in physical appearance, it is useful. The five racial types AMAR uses are:

Asiatic
 Mongolian, Chinese, Japanese, Malay, SE Asian peoples, Polynesians, Micronesians

African
Indigenous peoples of sub-Saharan Africa

Europoid
European peoples, indigenous peoples of North Africa (Berbers, Egyptians), Middle Eastern peoples (Arabs, Persians), some Central and South Asian peoples (Pashtuns, Baluchis)

Indio/Amerindian
Indigenous peoples of North and South America

Oceanic/Pacific
 Melanesians, Papuans, aboriginals of Australia and New Zealand

The main sources used to develop these guidelines are as follows:

Nei, Masatoshi and Arun K Roychoudhury. 1993. "Evolutionary Relationships and Human Populations on a Global Scale." *Molecular Biology and Evolution*. 10:5. 927-943,

Zhivotovsky, Lev A., Noah A. Rosenberg, and Marcus W. Feldman. 2003. "Features of Evolution and Expansion of Modern Humans, Inferred from Genomewide Microsatellite Markers." *American Journal of Human Genetics*. 72:1171-1186.

Description of reconciliation
- AMAR Phase I RACE 0 = MAR Phase V RACE 0, MAR Phase IV race N/A

- AMAR Phase I RACE 1 = MAR Phase V RACE 1, MAR Phase IV race 1
- AMAR Phase I RACE 2 = MAR Phase V RACE 2, MAR Phase IV race 2
- AMAR Phase I RACE 3 = MAR Phase V RACE 3, MAR Phase IV race 3
- AMAR Phase I RACE -99 = MAR Phase V -99, MAR Phase IV race -99
- AMAR Phase I RACE -88 = MAR Phase V RACE N/A, MAR Phase IV race 0

Note: Per the MAR Phase IV codebook, race 0 is defined as "unknown", even though there is already a -99 "missing" value. Also, the MAR Phase V codebook has a "no physical differences in appearance" (RACE=0) value, whereas the MAR Phase IV codebook does not. Since race is a static variable, it is likely that the MAR Phase V codes can be used to replace the MAR Phase IV code of 0 for these groups. However, we are recoding these cases -88 and leaving it up to the discretion of researchers to impute the 2004-2006 values for 1940-2003.

For STATA code to reconcile this variable, see Appendix II.

GC2 (RECONCILED VARIABLE)

Regional base:
A spatially contiguous region larger than an urban area that is part of the country, in which 25% or more of the minority resides and in which the minority constitutes the predominant proportion of the population.
0 = No
1 = Yes
-99 = No basis for judgment
NOTE: Originally collected for a collaboration between James Fearon, David Laitin, and MAR. Original data points were collected only for 1960 and 1990.

Description of reconciliation
- AMAR Phase I GC2 1 = MAR Phase V GC2 1, MAR Phase IV gc2 1
- AMAR Phase I GC2 0 = MAR Phase V GC2 0 = MAR Phase IV gc2 2

For STATA code to reconcile this variable, see Appendix II.

GC7 (RECONCILED VARIABLE)

Proportion of group living outside regional base
0   No regional base
1   >50%
2   25-50%
3   <25%
-99   No basis for judgment

Description of reconciliation
- AMAR Phase I GC7 0 = MAR Phase V GC7 0 = MAR Phase IV gc7 N/A

- AMAR Phase I GC7 1 = MAR Phase V GC7 1 = MAR Phase IV gc7 3
- AMAR Phase I GC7 2 = MAR Phase V GC7 2, MAR Phase IV gc7 2
- AMAR Phase I GC7 3 = MAR Phase V GC7 3 = MAR Phase IV gc7 1
- AMAR Phase I GC7 -99 = MAR Phase V GC7 -99, MAR Phase V Phase IV gc7 -99

For STATA code to reconcile this variable, see Appendix II.

GC10 (RECONCILED VARIABLE)
Transnational dispersion -- kindred groups
0   The group has no close kindred across an international border
1   The group has close kindred across a border which does not adjoin its regional base (including groups that have transnational kindred but not a regional base)
2   The group has close kindred in one country which adjoins its regional base
3   The group has close kindred in more than one country which adjoins its regional base
-99   No basis for judgment
NOTE: Originally collected for a collaboration between James Fearon, David Laitin, and MAR. Original data points were collected only for 1960 and 1990.

Description of reconciliation
- AMAR Phase I GC10 0 = MAR Phase V GC10 0 = MAR Phase IV gc10 1 & MAR Phase IV gc10 3
- AMAR Phase I GC10 1 = MAR Phase V GC10 1 = MAR Phase IV gc10 2
- AMAR Phase I GC10 2 = MAR Phase V GC10 2 = MAR Phase IV gc10 4
- AMAR Phase I GC10 3 = MAR Phase V GC10 3 = MAR Phase IV gc10 5

Note: MAR Phase IV gc10 3 is being reconciled as MAR Phase V GG10 0 since this value is a subset of MAR Phase V GC10 0.

For STATA code to reconcile this variable, see Appendix II.

GC11 (RECONCILED VARIABLE)
Transnational dispersion -- kindred groups in power
0   Kindred have no access to political power (no kindred abroad)
1   Kindred are outside political ruling coalition but are not barred from power
2   Kindred are in ruling coalition
3   Kindred dominate state coalition
-99   No basis for judgment

Description of reconciliation
- AMAR Phase I GC11 0 = MAR Phase V GC11 0 = MAR Phase IV gc11 1

- AMAR Phase I GC11 1 = MAR Phase V GC11 1 = MAR Phase IV gc11 2
- AMAR Phase I GC11 2 = MAR Phase V GC11 2 = MAR Phase IV gc11 3
- AMAR Phase I GC11 3 = MAR Phase V GC11 3 = MAR Phase IV gc11 4
- AMAR Phase I GC11 -99 = MAR Phase V GC11 -99, MAR Phase IV gc11 -99

For STATA code to reconcile this variable, see Appendix II.

EMIG (RECONCILED VARIABLE)
Emigration for political or economic reasons
0   Condition not present
1   Condition minor
     Affects less than 1% of group population
2   Condition of medium significance
     Affects between 1 and 10% of group population
3   Condition serious
     Affects more than 10%
-99   No basis for judgment

Description of reconciliation
- AMAR Phase I EMIG 0 = MAR Phase V EMIG 0 = MAR Phase IV DMEMEC 0 & DMEMPO 0
- AMAR Phase I EMIG 1 = MAR Phase V EMIG 1 = MAR Phase IV DMEMEC 1 or DMEMPO 1
- AMAR Phase I EMIG 2 = MAR Phase V EMIG 2 = MAR Phase IV DMEMEC 2 or DMEMPO 2
- AMAR Phase I EMIG 3 = MAR Phase V EMIG 3 = MAR Phase IV DMEMEC 3 or DMEMPO 3
- AMAR Phase I EMIG -99 = MAR Phase V EMIG -99, MAR Phase IV DMEMEC -99 & DMEMPO -99

Note: There are groups for which MAR Phase IV DMEMEC is -99 and MAR Phase IV DEMEPO is 0 and vise versa. For these cases, we are recoding based on the value of 0, rather than -99; therefore, groups are reconciled as AMAR Phase I EMIG 0 if at least one of the constituent MAR Phase IV variables is coded 0.

For STATA code to reconcile this variable, see Appendix II.

CULPO2 (RECONCILED VARIABLE)
Restrictions on use of language or language instruction
0   No restrictions
1   Activity informally restricted.
     The activity is restricted by widespread but informal social practice (e.g., by discrimination against people who speak the group's language)
2   Activity somewhat restricted

3   Activity sharply restricted
-99   No basis for judgment

Description of reconciliation
- AMAR Phase I CULOP2 0 = MAR Phase V CULPO2 0, MAR Phase IV CULPO2 0 & CULPO3 0
- AMAR Phase I CULPO2 1 = MAR Phase V CULPO2 1 = MAR Phase IV CULPO2 1 or CULPO3 1
- AMAR Phase I CULPO2 2 = MAR Phase V CULPO2 2 = MAR Phase IV CULPO2 2 or CULPO3 2
- AMAR Phase I CULPO2 3 = MAR Phase V CULPO2 3 = MAR Phase IV CULPO2 3 or CULPO3 3
- AMAR Phase I CULPO2 -99 = MAR Phase V CULPO2 -99, MAR Phase IV CULPO2 -99 & CULPO3 -99

Note: There are cases in which MAR Phase IV culpo2 is -99 and MAR Phase IV culpo3 is 0 and vise versa. For these cases, we are recoding based on the value of 0, rather than -99; therefore, groups are reconciled as AMAR Phase I CULPO2 0 if at least one of the constituent MAR Phase IV variables is coded 0.

For STATA code to reconcile this variable, see Appendix II.

POLGR (RECONCILED VARIABLE)
Highest level of political grievance
0   No political grievances expressed
1   Political grievances focused on ending discrimination or creating or strengthening remedial policies
2   Political grievances focused on creating or strengthening autonomous status
3   Political grievances focused on creating separate state for group or revanchist change in borders (union with kindred requires revanchist change in borders).
-99   No basis for judgment

Description of reconciliation
- AMAR Phase I POLGR 0 = MAR Phase V POLGR 0, MAR Phase IV POLGR 0
- AMAR Phase I POLGR 1 = MAR Phase V POLGR 1, POLGR 2 = MAR Phase IV Polgr1, Polgr2, Polgr3, Polgr4, Polgr5, Polgr6
- AMAR Phase I POLGR 2 = MAR Phase V POLGR 3 = MAR Phase IV Autgr1, Autgr4, Autgr5, Autgr6
- AMAR Phase I POLGR3 = MAR Phase V POLGR 4 = MAR Phase IV Autgr2, Autgr3

Note: Reconciling the MAR Phase IV autonomy variables was straightforward, as they could be directly matched to the coding levels of the AMAR Phase I POLGR 2 value.

It was also straightforward to reconcile the MAR Phase IV Polgr4 variable, since demands for equal civil rights can be considered calls for ending discrimination, which corresponds to AMAR Phase I POLGR 1 value.

Reconciling the other MAR Phase IV non-autonomy variables – Polgr1, Polgr2, Polgr3, Polgr5, and Polgr6 – was not as straightforward, since it was not clear whether these variables should have been recoded as MAR Phase V POLGR level 1 (group calling for ending discrimination) or MAR Phase V POLGR level 2 (group calling for creating or strengthening remedial policies or programs).[5] For instance, just knowing that the group seeks greater political rights is not enough to determine whether a Phase V code of 1 or 2 is appropriate, since seeking greater political rights may mean gaining access to political participation (which would be reconciled as Phase V code of 1 – ending discrimination) but it could also mean calling for guaranteed representation in the legislature (which would be reconciled as Phase V code of 2 – seeking remedial policies).

Rather than re-examining the notes and sources for the approximately 900 groups that are coded as having MAR Phase IV Polgr1, Polgr2, Polgr3, Polgr5, and Polgr6 grievances only (and if the notes and sources are not sufficiently detailed to determine which MAR Phase V code should be used, checking external sources for more information), we collapsed the MAR Phase V POLGR coding values of 1 and 2 into one category for AMAR Phase I, so that ending discrimination and calling for remedial policies are now both coded as AMAR Phase I POLGR value of 1. Our rationale for doing so is that very few groups in the MAR Phase V data advocate for remedial policies and programs. Of the 852 group-year observations for POLGR in the MAR Phase V data, only 71 or about 8% advocate for remedial policies and programs only; subsequently, collapsing the MAR Phase V POLGR 1 and 2 levels into one category for AMAR Phase I results in a minimal loss of variation. Therefore, MAR Phase IV Polgr1, Polgr2, Polgr3, Polgr5, and Polgr6 variables are reconciled as AMAR Phase I POLGR value of 1.

Further, since we collapsed MAR Phase V POLGR 1 and POLGR 2 into one category – POLGR 1 – for AMAR Phase I, we also relabeled the remaining MAR Phase V POLGR values. MAR Phase V POLGR level 3 (autonomy) is now relabeled as AMAR Phase I POLGR level 2 and MAR Phase V POLGR level 4 (political independence) is now relabeled as AMAR Phase I POLGR level 3. Therefore, for AMAR Phase I, we collapsed MAR Phase V categories 1 and 2

[5] Examples of a code of 1 are: claiming to be marginalized; claiming laws are discriminatory (e.g., citizenship or other laws); asking the government to repeal discriminatory laws; asking the government to lift restrictions on political participation (e.g., identity documents that are required for voting); and asking the government to implement rights already granted in constitution and other laws; asking for equal civil and political rights. Examples of a code of 2 are groups that ask for: a multi-party government cabinet, proportional representation, power-sharing arrangement, or other guaranteed representation; special voting constituencies; quotas in government posts; implementation of or a change in affirmative action programs; citizenship or dual citizenship; the government to pay more attention to the ethnic group.

into AMAR Phase I category 1 and re-ordered the remaining MAR Phase V categories 3 and 4 as AMAR Phase I categories 2 and 3, respectively.

Finally, if a group was coded 0 and -99 for one of the collapsed MAR Phase IV variables, we reconciled it as AMAR Phase I value of 0, rather than -99, since 0 is a more definitive coding than -99. As a result, groups that were reconciled as AMAR Phase I 0 either have 0 values for all of the collapsed MAR Phase IV variables or 0 values for at least one of the collapsed MAR Phase IV variables.

For STATA code to reconcile this variable, see here and here in Appendix II.

ECGR (RECONCILED VARIABLE)
Highest level of economic grievance
0   No economic grievances expressed
1   Economic grievances focused on ending discrimination
2   Economic grievances focused on creating or strengthening remedial policies
-99 No basis for judgment

Description of reconciliation
- AMAR Phase I ECGR 0 = MAR Phase V ECGR 0, MAR Phase IV Econgr0
- AMAR Phase I ECGR 1 = MAR Phase V ECGR 1 = Phase IV Econgr1
- AMAR Phase I ECGR 2 = MAR Phase V ECGR 2 = MAR Phase IV Econgr2, Econgr3, Econgr4, Econgr5
- AMAR Phase I ECGR -99 = MAR Phase V ECGR -99, MAR Phase IV Econgr1 -99 & Econgr2 -99 & Econgr3 -99 & Econgr4 -99 & Econgr5 -99 & Econgr6 -99
- Note: MAR Phase IV Econgr6 reconciled on case by case basis (see below)

Note: It was straightforward to reconcile MAR Phase IV Econgr2 (share of public funds), Econgr3 (greater economic opportunities), Econgr4 (improving working conditions and protective regulations), and Econgr5 (protection of land, jobs, and resources) as AMAR Phase I ECGR level of 2, since these variables explicitly call for remedial policies or programs.

It was not straightforward to reconcile MAR Phase IV Econgr1 and Econgr6, since it was not immediately clear whether diffuse economic concerns and other economic concerns translate into AMAR Phase I ECGR codes of 1 or 2. These coding values were therefore reconciled as follows.

MAR Phase IV Econgr1 was reconciled as AMAR Phase I ECGR 1. Since this variable indicates that the group has an economic grievance but there is no additional information to indicate whether the grievance is related to ending discrimination or remedial policies, we conservatively recoded these cases as AMAR Phase I ECGR 1 in order to not inflate the severity of the grievance. There are 48 groups that were coded as MAR Phase IV Econgr1 only that are being reconciled as AMAR Phase I ECGR 1.

There are 18 observations that were coded as MAR Phase IV Econgr6 only. These groups are being reconciled on a case by case basis. See below for detailed information about each of these cases.

Further, if a group was coded 0 and -99 for one of the collapsed MAR Phase IV variables, we reconciled it as an AMAR Phase I value of 0, rather than -99, because 0 is a more definitive coding than -99. As a result, groups that were reconciled as AMAR Phase I 0 either have 0 values for all of the collapsed MAR Phase IV variables or 0 values for at least one of the collapsed MAR Phase IV variables.

1. Hill Tribes in Burma in 1985 – Coding notes state, "Stop migration of Bengalis into the Hill Tracts (and remove illegal settlers)." Since grievance requires government action, recoding as ECGR 2.

2. Rohingya (Arakanese) in Burma 1985 - Coding notes state, "Nationalizations of property due to citizenship laws forced thousands to migrate." Since grievance suggests discriminatory policy affected group status but no indication group is asking the government to implement any policy to remedy grievance, recoding as ECGR 1.

3. Rohingya (Arakanese) in Burma 1995 - Coding notes state, "There is substantial forced labor, porterage, and extortion. Little is being done to reduce forced labor." Since grievance suggests discriminatory policy affected group status but no indication group is asking the government to implement any policy to remedy grievance, recoding as ECGR 1.

4. Turkmen in China in 1985 -Coding notes state, "Control Han migration." Since grievance suggests the group is asking the state to implement a remedial policy to control migration, recoding as ECGR 2.

5. Eritreans in Ethiopia in 1985 - Coding notes state, "More equitable, egalitarian society; land reform." Since grievance indicates the group is asking the government for land reform, recoding as ECGR 2.

6. Palestinians in Lebanon in 1985 - Coding notes state, "Refugee status makes economic integration problematic." Since grievance suggests refugee status is economically problematic but no indication group is calling for government to implement policies to ameliorate status, conservatively coding as ECGR 1.

7. Merina in Madagascar in 1990 - Coding notes state, "Maintaining economic advantages for Merina elites." Since grievance suggests the group is asking the government to continue implementing policies that maintain the economic status of the group, recoding as ECGR 2.

8. Kewri in Mauritania in 1985 - Coding notes state, "They protest Arabic as official language, Islamic education, and Islamic law." Since grievance indicates the group is protesting discriminatory policies, recoding as ECGR 1.

9. Pashtuns (Pushtuns) in Pakistan in 1985 - Coding notes state, "Problems associated with millions of Afghan refugees from the war." Since grievance suggest group is concerned about refugee issue but no further information indicating group is asking the government for policies to ameliorate grievance, coding as ECGR 1.

10. Chinese in Panama in 1990 - Coding notes state, "Removal of economic restrictions on 'non-citizens'." Since grievance indicates the group is asking the government to remove a discriminatory policy, recoding as ECGR 1.

11. Basques in Spain in 1985 - Coding notes state, "Some separatists have anti-capitalist demands, directed mainly at the Basque industrialist/banker group." Since grievance suggests the group may be asking for the implementation of anti-capitalist policies, recoding as ECGR 2.

12. Northern Hill Tribes in Thailand in 1990 - Coding notes state, "It seems [in addition to] treatment of drug addiction, the tribes people want better economic opportunities. Far Eastern Economic Review, 12/2/1993." Since grievance indicates the group wants better economic opportunities but not clear if group is asking for government to end discriminatory policies or implement remedial ones to ameliorate grievance, conservatively recoding as ECGR 1.

13. Europeans in Zimbabwe in 1985 - Coding notes state, "Fewer restrictions on activities; stronger guarantees and privileges." Since grievance indicates group seeks stronger guarantees and privileges, which suggest remedial policies, recoding as ECGR 2.

14. Avars in Russia in 1995 – Coding notes state, "Other = continuation of heavy subsidies from Moscow. From previous coding." Since grievance indicates the group is asking for the continuation of a remedial policy, recoding as ECGR 2.

15. Roma in Greece in 1995 - Coding notes state, "Protection from being removed from their homes and protection from police brutality. http://errc.org/publications/factsheets/greece.shtml". Since grievance indicates the group is asking the state to take action, recoding as ECGR 2.

16. Roma in Greece in 2001-2003 – Coding notes state, "Want end of the practice of forced resettlement. BBC May 31 2003, Amnesty International, ERRC Report." Since grievance indicates the group is asking the state to take action to change a practice, recoding as ECGR 2.

For STATA code to reconcile this variable, see Appendix II.

CULGR (RECONCILED VARIABLE)
Highest level of cultural grievance
0   No cultural grievances expressed
1   Cultural grievances focused on ending discrimination

2   Cultural grievances focused on creating or strengthening remedial policies

-99   No basis for judgment

Description of reconciliation
- AMAR Phase I CULGR 0 = MAR Phase V CULGR0, MAR Phase IV Culgr1 & Culgr2 & Culgr3 & Culgr4 & Culgr5
- AMAR Phase I CULGR 1 = MAR Phase V CULGR 1 = MAR Phase IV Culgr1
- AMAR Phase I CULGR 2 = MAR Phase V CULGR 2 = MAR Phase IV Culgr2, Culgr3, Culgr4, Culgr5
- Note: Phase IV Culgr6 reconciled on case by case basis (see below)

Note: It was straightforward to reconcile MAR Phase IV Culgr1 (freedom of religion) as AMAR Phase I CULGR level of 1, since this variable is indicative of a call for ending discrimination. It was also straightforward to reconcile MAR Phase IV Culgr2 (promoting culture), Culgr3 (instruction in group language), Culgr4 (official use of group language), and Culgr5 (protection from attacks) as AMAR Phase I CULGR level 2, since these grievances require some sort of state funding or state program.

It was not straightforward to reconcile MAR Phase IV Culgr6, since it was not clear whether "other social/cultural concerns" warranted an AMAR Phase I CULGR value of 1 or 2. As such, the eight groups that were coded as MAR Phase IV Culgr6 only were reconciled on a case by case basis. See below for detailed notes.

Further, if a group was coded 0 and -99 for one of the collapsed MAR Phase IV variables, we reconciled it as an AMAR Phase I value of 0, rather than -99, because 0 is a more definitive coding than -99. As a result, groups that were reconciled as AMAR Phase I 0 either have 0 values for all of the collapsed MAR Phase IV variables or 0 values for at least one of the collapsed MAR Phase IV variables.

1. Haitain Blacks in the Dominican Republic - Coding notes state, "Haitian workers and their supporters have long protested the DR government's deportation policies, and the denial of citizenship to the children of Haitian laborers." Since grievance suggests group wants the government to change the current citizenship policy, recoding as CULGR 2.

2. Afro-Guyanans in Guyana - Coding notes state "Recognition of the impact of slavery on Afro-Guyanans." Since grievance indicates group wants the government to recognize status but no further mention of group calling for a remedial policy, conservatively recoding as CULGR 1.

3. Blacks in Panama - Coding notes state "An end to discrimination at nightclubs.
(June 1999) Inter Press Service reported the Pro-Dignity Committee was formed by a group of Afro-Panamanians to formally protest the entrance policies of clubs and discos in the cities. Clubs regularly used "the right of admission" to prevent blacks and other ethnic

groups from entering." Since grievance indicates the group wants to end discrimination, recoding as CULGR 1.

4. Roma in Russia - Coding notes state, "Change in the popular prejudice against the group." Since grievance suggests group wants the government to take some sort of action to change prejudice against the group, recoding as CULGR 2.

5. Nagas in India - Since previous coding is -99 and subsequent coding is mostly 0, conservatively recoding as CULGR 1.

6. Asians in South Africa - Since previous coding is -99 and subsequent coding is recoded as CULGR 2, conservatively recoding as CULGR 1.

7. Coloreds in South Africa - Since previous coding is -99 and subsequent coding is recoded as CULGR 2, conservatively recoding as CULGR 1.

8. Antillean Blacks in Costa Rica - Since previous coding is -99 and subsequent coding is recoded as CULGR 2, conservatively recoding as CULGR 1.

For STATA code to reconcile this variable, see Appendix II.

FACTSEV1-3 (RECONCILED VARIABLES)
Severity of conflict for first-third pair of antagonists

0	No conflict
1	Sporadic violent attacks
	Attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns).
2	Series of bombings/assassinations
3	Substantial rioting
4	Sporadic armed clashes
	Attacks with multiple firearms, automatic weapons, or heavy weaponry (mortars, shelling, etc.)
5	Protracted communal warfare
	More than 6 clashes a year between antagonists
-99	No basis for judgment

Description of reconciliation
Since the coding levels for the MAR Phase IV fccs1-3 and AMAR Phase I FACTSEV1-3 variables are the same, we renamed the MAR Phase IV variable to match the AMAR Phase I variable name.

For STATA code to reconcile these variables, see Appendix II.

Level of conflict with CCGROUP1-CCGROUP3

0	No conflict
1	Individual acts of harassment, no fatalities
2	Political agitation, campaigns urging authorities to impose restrictions on group
3	Sporadic violent attacks by gangs or other small groups, some fatalities
	Attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns) involving fewer than 20 people.
4	Anti-group demonstrations, rallies, marches
5	Communal rioting, armed attacks
	Attacks with multiple firearms, automatic weapons, or heavy weaponry (mortars, shelling, etc.) OR attacks without weapons (e.g., brawls), knives, or few small arms (e.g., one or two handguns) involving more than 20 people
6	Communal warfare
	More than 6 clashes a year between antagonists
-99	No basis for judgment

Description of reconciliation

- AMAR Phase I CCGROUPSEV1-3 0 = MAR Phase V CCGROUPSEV1-3 0 = MAR Phase IV cc1x-cc3x 0 & year <1990 & MAR Phase IV gcc1-gcc3 0
- AMAR Phase I CCGROUPSEV1-3 1 = MAR Phase V CCGROUPSEV1-3 1 = MAR Phase IV cc1x-cc3x 1 & year <1990 & MAR Phase IV gcc1-gcc3 1
- AMAR Phase I CCGROUPSEV1-3 2 = Phase IV cc1x-cc3x 2 & year <1990 = Phase V CCGROUPSEV1-3 2 & MAR Phase IV gcc1-gcc3 2
- AMAR Phase I CCGROUPSEV1-3 3 = Phase IV cc1x-cc3x 3 & year <1990 = Phase V CCGROUPSEV1-3 3 & MAR Phase IV gcc1-gcc3 3
- AMAR Phase I CCGROUPSEV1-3 4 = Phase IV cc1x-cc3x 4 & year <1990 = Phase V CCGROUPSEV1-3 4 & MAR Phase IV gcc1-gcc3 4
- AMAR Phase I CCGROUPSEV1-3 5 = Phase IV cc1x-cc3x 5 & year <1990 = Phase V CCGROUPSEV1-3 5 & MAR Phase IV gcc1-gcc3 5
- AMAR Phase I CCGROUPSEV1-3 6 = Phase IV cc1x-cc3x 6 & year <1990 = Phase CCGROUPSEV1-3 6 & MAR Phase IV gcc1-gcc3
- AMAR Phase I CCGROUPSEV1-3 -99 = MAR Phase V CCGROUPSEV1-3 -99, MAR Phase IV cc1x-cc3x -99 & year <1990 & MAR Phase IV gcc1-gcc3 -99

Note: It is important to note that in the MAR Phase IV 1940-2003 data, the first inter-communal conflict severity variable, cc1x-cc3x, is coded every 10 years from 1940-1990 and the second inter-communal conflict severity variable, gcc1-gcc3, is coded annually from 1990-2000 (although there is one group that is coded from 1981-1989, Darfur Muslims in Sudan). As such, there is one year of coding overlap between these two variables: cc1x-cc3x for 1990 records the highest level of conflict for the 1990s and gcc1-

gcc3 for 1990 records the highest level of conflict for 1990 specifically. Since both variables are coded in the same way (i.e., the exact same coding levels) and gcc1-gcc3 codes annual data beginning in 1990, AMAR Phase I CCGROUPSEV1-3 was reconciled with decennial data from MAR Phase IV cc1x-cc3x for 1940-1980 and annual data from MAR Phase IV gcc1-gcc3 for 1990-2003.

Further, MAR Phase IV cc1x-cc3x and gcc1-gcc3 are not ordered in descending order of severity, whereas AMAR Phase I CCGROUPSEV1-3 are (i.e., CCGROUPSEV1 lists the group with the highest severity of inter-communal conflict, CCGROUPSEV2 lists the group with the second-highest level of inter-communal conflict, and CCGROUPSEV3 lists the group with the third-highest level of inter-communal conflict, whereas cc1x-cc3x and gcc1-gcc3 do not). We are noting the difference in ordering between the MAR Phase IV and AMAR Phase I and MAR Phase V data, but leaving these variables as is.

For STATA code to reconcile this variable, see Appendix II.

PROT (RECONCILED VARIABLE)
Protest
0	None reported
1	Verbal opposition
	Public letters, petitions, posters, publications, agitation, court action etc). Requests by a minority-controlled regional group for independence.
2	Symbolic resistance
	Sabotage, symbolic destruction of property OR political organizing activity on a substantial scale (e.g. sit-ins, blockage of traffic). Mobilization for autonomy/secession by a minority-controlled regional government.
3	Small demonstrations
	A few demonstrations, rallies, strikes, and/or riots, the largest of which has total participation of less than 10,000
4	Medium demonstrations
	Demonstrations, rallies, strikes, and/or riots, the largest of which has total participation between 10,000 and 100,000
5	Large demonstrations
	Demonstrations, rallies, strikes, and/or riots, , the largest of which has total participation over 100,000
-99	No basis for judgment

Description of reconciliation
- AMAR Phase I PROT 0 = MAR Phase V PROT 0, MAR Phase IV proti 0 & year<1985 or MAR Phase IV prot 0
- AMAR Phase I PROT 1 = MAR Phase V PROT 1, MAR Phase IV proti 1 & year <1985 or MAR Phase IV prot 1

- AMAR Phase I PROT 2 = MAR Phase V PROT 2, MAR Phase IV proti 2 & year<1985 or MAR Phase IV prot 2
- AMAR Phase I PROT 3 = MAR Phase V PROT 3, MAR Phase IV proti 3 & year<1985 or MAR Phase IV prot 3
- AMAR Phase I PROT 4 = MAR Phase V PROT 4, MAR Phase IV proti 4 & year<1985 or MAR Phase IV prot 4
- AMAR Phase I PROT 5 = MAR Phase V PROT 5, MAR Phase IV proti 5 & year <1985 or MAR Phase IV prot 5
- AMAR Phase I PROT =99 = MAR Phase V PROT -99, MAR Phase IV proti -99 & year <1985 or MAR Phase IV prot -99

Note: MAR Phase IV Proti codes data every five years from 1940-1999, whereas MAR Phase IV prot codes data annually from 1985. To create a single, integrated variable, the AMAR Phase I PROT variable was reconciled to include quinquiennial proti data from Phase IV 1940-1984 and annual prot data from Phase IV 1985-2003.

For STATA code to reconcile this variable, see Appendix II.

REB (RECONCILED VARIABLE)
Rebellion
0   None reported
1   Political banditry, sporadic terrorism (fewer than 6 events)
2   Campaigns of terrorism (more than 6 events)
3   Local rebellions
      Armed attempts to seize power in a locale. If they prove to be the opening round in what becomes a protracted guerrilla or civil war during the year being coded, code the latter rather than local rebellion.
4   Small-scale guerrilla activity
      Includes all three of the following traits:
- fewer than 1000 armed fighters
- sporadic armed attacks (less than 6 reported per year)
- attacks in a small part of the area occupied by the group or in one or two other locales
5   Intermediate guerrilla activity
      Includes one or two of the defining traits of large-scale activity and one or two of the defining traits of small-scale activity
6   Large-scale guerrilla activity
      Includes all three of the following:
- more than 1000 armed fighters
- frequent armed attacks (more than 6 reported per year)
- attacks affecting large part of the area occupied by

group
    7   Civil war
            Protracted civil war fought by rebel military
            Has all the characteristics of large-scale guerilla activity,
            plus rebels control large scale base areas that are secure
            over time
    -99   No basis for judgment

Description of reconciliation
- AMAR Phase I REB 0 = MAR Phase V REB 0 = MAR Phase IV rebel 0 & year<1985 or MAR Phase IV reb 0
- AMAR Phase I REB 1 = MAR Phase V REB 1= MAR Phase IV rebel 1 & year<1985 or MAR Phase IV reb 1
- AMAR Phase I REB 2 = MAR Phase V REB 2 = MAR Phase IV rebel 2 & year<1985 or MAR Phase IV reb 2
- AMAR Phase I REB 3 = MAR Phase V REB 3 = MAR Phase IV rebel 3 & year<1985 or MAR Phase IV reb 3
- AMAR Phase I REB 4 = MAR Phase V REB 4 = MAR Phase IV rebel 4 & year<1985 or MAR Phase IV reb 4
- AMAR Phase I REB 5 = MAR Phase V REB 5 = MAR Phase IV rebel 5 & year<1985 or MAR Phase IV reb 5
- AMAR Phase I REB 6 = MAR Phase V REB 6 = MAR Phase IV rebel 6 & year<1985 or MAR Phase IV reb 6
- AMAR Phase I REB 7 = MAR Phase V REB 7 = MAR Phase IV rebel 7 & year<1985 or MAR Phase IV reb 7
- AMAR Phase I REB -99 = MAR Phase V REB -99 = MAR Phase IV rebel-99 & year<1985 or MAR Phase IV reb -99

Note: MAR Phase IV Rebel is coded every five years from 1945-1999 and MAR Phase IV reb is coded annually from 1985. To create a single, integrated rebellion variable, the AMAR Phase I REB variable was reconciled to combine quinquiennial rebel data from Phase IV 1945-1984 and annual reb data from Phase IV 1985-2003.

For STATA code to reconcile this variable, see Appendix II.

V. Appendix II – STATA code for reconciliation process

```
/*This do file includes syntax to implement the following changes:

*MAR 1940-2003 data*
use MAR data (through 2003) available here - http://www.cidcm.umd.edu/mar/mar_data.asp
I. MAR 1940-2003 data: cleaning and updating the MAR 1940-2003 data so that it can be
merged with the MAR 2004-2006 and selection bias data.
II. MAR 1940-2003 data: Reconciling the 1940-2003 data to match the coding conventions
of the 2004-2006 data.
III. MAR 1940-2003 data: Replacing 1940-2003 poldis and ecdis data with 1980-1989
poldis and ecdis data coded by Victor Asal,
which updates these variables for 1980 and 1985-1989 and adds new coding for 1981-
1984.

*MAR 2004-2006 data*
use MAR data (2004-2006) available here - http://www.cidcm.umd.edu/mar/mar_data.asp
IV. MAR 2004-2006 data: cleaning the MAR 2004-2006 data
V. MAR 2004-2006 data: appending the MAR 2004-2006 data to the 1940-2003 data

*MAR selection bias data*
VI. MAR 1980-2006 selection bias data: cleaning the MAR 1980-2006 selection bias data
VII. MAR selection bias data: appending the MAR selection bias data to the 1940-2003
data (note: this data is referred to as the AMAR 1940-2006 data)

*AMAR 1940-2006 data
VIII. AMAR 1940-2006 data: recoding the POLGR variable
IX. AMAR 1940-2006 data: implementing approved recode requests
X. AMAR 1940-2006 data: changing name of Yugoslavia to correspond to de jure name
changes (as indicated in Serbs in Yugoslavia coding from selection bias data)
XI. AMAR 1940-2006 data: changing name of Kosovo Albanians in Yugoslavia to reflect
fact that group's name changed in AMAR

*AMAR sample frame data
XII. AMAR sample frame: adding updated 2001/2007 gpro figures from the AMAR sample
frame*/

********************************************************************************

/*I. The following code cleans and updates the MAR 1940-2003 data so that it can be
merged with the MAR 2004-2006 and selection bias data.

This mostly consists of making sure type and names of groups and countries match,
getting rid of duplicate files, and getting rid of MAR variables that are not coded in
the update.*/

*1) Reconciling MAR 1940-2003 country names (capitalization and spelling) to that of
MAR 2004-2006 and selection bias (latest update) country names.

set mem 1g

use "C:\Users\Aga\Desktop\AMAR Reconciliation_2014\AMAR Phase VI Reconciled
Data_FINAL\AMAR Phase IV beta.UPDATED NOV 11\mardatav02.2005.dta", clear

sort country group year

replace country=proper(country)

sort country group year

by country group: gen n1 = (_n == 1)
```

```
*note the first record within the group
list country if n1==1

replace country = "Czech Republic" if country == "Czechrep"
replace country = "Costa Rica" if country == "Costarica"
replace country = "Dominican Republic" if country == "Dominican Rep."
replace country = "Democratic Republic of Congo" if country == "Dem. Rep. Congo"
replace country = "Republic of Congo" if country== "Rep. Of Congo"
replace country = "Korea, South" if country == "S. Korea"
replace country = "Papua New Guinea" if country == "Papua N.G."
replace country = "United Kingdom" if country == "Uk"
replace country = "United States of America" if country == "Usa"
replace country = "Russia" if country == "Ussr"

*2) Reconciling MAR 1940-2003 group names (capitalization and spelling) to that of
latest update group names.

replace group=proper(group)

list group if n1==1

replace group = "Indigenous Highland Peoples" if group == "Indigenous Highland Ppls"
replace group = "Lowland Indigenous Peoples" if group == "Lowland Indigenous Ppls"
replace group = "Kirdi" if group == "Kirdis" & country == "Cameroon"
replace group = "Kalenjin" if group == "Kalenjins" & country == "Kenya"
replace group = "Maasai" if group == "Maasais" & country == "Kenya"
replace group = "Catholics in Northern Ireland" if group == "Catholics In N. Ireland"
replace group = "Bemba" if group == "Bemebe" & country == "Zambia"
replace group = "Malinke" if group =="Malinka" & country == "Guinea"
replace group = "Baha'is" if group =="Baha'Is" & country == "Iran"
replace group = "M'boshi" if group =="M'Boshi" & country == "Republic of Congo"
replace group = "Diolas in Casamance" if group == "Diolas In Casamance" & country ==
"Senegal"
replace group = "Crimean Tatars" if group == "Crimean Tartars" & country == "Ukraine"
replace group = "Shi'is" if group == "Shi'Is"

*As a result of changing all capitalized to proper (see above command) Shi'ia changed
to Shi'Ia in Iran Lebanon Bahrain Saudi Arabia

*3) Reconciling MAR 1940-2003 regions to that of latest update regions.

gen region1="."

replace region1 = "Western Democracies and Japan" if region==0
replace region1 = "Post-communist States" if region==2
replace region1 = "Asia" if region==3
replace region1 = "Middle East and North Africa" if region==5
replace region1 = "Sub-Saharan Africa" if region==6
replace region1 = "Latin America and the Caribbean" if region==7

drop region

rename region1 region

*Note: Post-communist States used to be E. Europe and former Soviet Union.  Middle
East and North Africa was N. Africa and Middle East

sort country group year

/*4. Removing 2 records that are duplicated in the updated MAR2004-2006 data and
should not be in this data
(Sudan Darfur Black Muslims in 2004 and Syria Kurds in 2004)*/
```

```
drop if year == 2004

/* 5. In the first phase Tonga in Zambia and Mande in Mali were coded but were dropped
from later phases.  According to Amy Pate, former research director, Mande were likely
dropped because they were considered "Advantaged minority, 43% of population,
dominates government and economy".  There is no clear explanation for why the Tonga
were dropped.  However, since neither group was part of phase II they were eligible
and included in the selection bias.  Because the availability of data for coding of
selection bias is better than Phase I, we are dropping them from the data here and
including them in the selection bias data*/

drop if group=="Tonga" & country=="Zambia"
drop if group=="Mande" & country=="Mali"

/*6. Keep only the cases that are current (coded continuously from the time they enter
the data) and will be used for the eventual analysis, including Kurds in Syria and
Black Muslims in Darfur Sudan*/

replace current=1 if country=="Syria" & group=="Kurds" & year>1979

replace current=1 if country=="Sudan" & group=="Darfur Black Muslims" & year>1979

/*Note from 09-12.  Some groups most notably in the USSR are not current after 1990
because their country name changed in the 1990s.
Three of these cases, Chechens, Ingush, and Tatars in the USSR were coded in phase I
but the Chechen and Ingush were coded as a single group.
When the USSR broke up these groups were added in separately as Phase II current
groups and designated non current as coded in the USSR prior to 1990.
In the interest of retaining data, we change the designation of non-currency for
Tatars groups prior to 1990 but retain a noncurrency status for the Chechen and Ingush
because they were coded as one group.*/

replace current=1 if country=="Russia" & group=="Tatars" & year<1992

keep if current==1

/*Note on current. Some of the cases designated not current were so designated because
they were "no longer at risk." While we do not include them here because no data was
collected on them from Phase II a few of these groups are enumerated in the AMAR data
and may be reintegrated once that data is fully populated.

The group Croats A was not coded as current and was dropped but there was a name
change from Croats B to simply Croats in 2001.  Thus we change the name here.*/

replace group = "Croats" if group == "Croats B" & country=="Yugoslavia"

/*Two additional groups are no longer coded separately due to changes in country
borders and are therefore dropped.  These are Roma in Croatia and East Timorese in
Indonesia.  For further information on groups moving in and out of MAR see paper.*/

drop if group=="Roma" & country=="Croatia"

drop if group=="East Timorese" & country=="Indonesia"

/*6 - changing numcode error - Montagnards coded 81605 (which is a numcode that is not
assigned to a group) instead of 81604 (which
is the Montagnards numcode) for 1945 and 1994*/
list numcode group year if country =="Vietnam"
replace numcode=81604 if group=="Montagnards" & year==1945
replace numcode=81604 if group=="Montagnards" & year==1994
list group year if numcode ==81605
```

```
*7 - changing numcode error - Oromo in Ethiopia coded 53001 (which is numcode for
Afars) instead of 53004 (Oromo numcode) for 1990
list numcode group year if country=="Ethiopia"
list numcode group country year  belief if numcode==53001 & year>1985
list numcode group country year  belief if numcode==53004 & year>1985
replace numcode=53004 if group=="Oromo" & year==1990
list group year if numcode==53004

/*Note: There are a few cases of groups that were coded separately for only a limited
amount of time but not dropped because they are coded current in Phase V perhaps
because they were folded into other cases or because they were dropped after Phase V.
These are the Chinese and Montagnards in South Vietnam (cow code 817, coded 1955-1975
only) and Vietnam (cow code 816, coded 1940-2003) and the Turks in East Germany (cow
code 265, coded 1950-1989), West Germany (cow code 260, coded 1950-1989), and Germany
(cow code 255, coded 1940-1945 and 1990-2003).
Since all of these groups are current and have distinct numcodes that correspond to
their differing countries, we are leaving them as is.*/

/*Note: One additional case has been dropped as of the update.  These are the
Honamese.
However, they were only dropped after the update was completed and the full suite of
data exists on them.  Consequently, we keep them in this data.*/

count if n1==1

/*Note: Remaining are the 283 cases (+ the Honamese) that are continuously coded from
the time they enter into the data until the latest update.*/

save "C:\Users\Aga\Desktop\mardata02.2005.2.dta", replace

clear

*********************************************************************************

*7 - Keeping only the variables that are being reconciled and that are current in the
latest version of the codebook.

use "C:\Users\Aga\Desktop\mardata02.2005.2.dta", clear

keep group country year autgain autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 autlost
auton auton2 autonend autpro belief cc1x cc2x cc3group cc3x cc3xa ccgroup1 ccgroup2
ccode ccx1 ccx2 ccx3 cpop culgr1 culgr2 culgr3 culgr4 culgr5 culgr6 culpo1 culpo2
culpo3 current custom dmemec dmempo ecdis econgr1 econgr2 econgr3 econgr4 econgr5
econgr6 factcc1 factcc2 factcc3 factcc4 fccs1 fccs2 fccs3 fccs4 fccx1 fccx2 fccx3
fccx4 gc2 gc6b gc7 gc10 gc11 gc119 gcc1 gcc2 gcc3 gojpa gpop gpro groupcon id intracon
intercon lang marstat name numcode poldis polgr1 polgr2 polgr3 polgr4 polgr5 polgr6
prot proti race reb rebel religs1 sepkin sepx transyr type n1 region version

save "C:\Users\Aga\Desktop\mardata02.2005.3.dta", replace

clear

*********************************************************************************

/*II. The following code reconciles certain variables from Phase IV codebook (as coded
for the MAR 1940-2003 data) to match the coding conventions in the Phase V codebook
(as coded for MAR 2004-2006 and selection bias data).*/

use "C:\Users\Aga\Desktop\mardata02.2005.3.dta"

*1 - STATA code for reconciling BELIEF variable.
```

```
/*Note: Since there is no direct match for Phase IV (MAR 1940-2003) belief = 0
"unknown" and belief = 2 "Multiple Sects; some different from dominant group", these
values are being recoded as -88 and -77, respectively.
Although the belief variable is relatively static and it is likely that the Phase V
BELIEF coding for groups that have Phase IV belief values of 0 and 2 can be used to
update these groups, we are leaving this up to the discretion of researchers and are
conservatively recoding these as -88 and -77.*/

*generating BELIEF variable
gen BELIEF = .

*reconciling BELIEF
replace BELIEF=1 if belief ==1
replace BELIEF=2 if belief ==3
replace BELIEF=-99 if belief ==-99
replace BELIEF=-88 if belief==2
replace BELIEF=-77 if belief==0
tab belief
tab BELIEF

*examples of reconciled groups

list group country year belief BELIEF if belief==3 & n1==1
/*      +-----------------------------------------------------------------+
        |                   group        country   year   belief   BELIEF |
        |-----------------------------------------------------------------|
  113.  |                  Greeks        Albania   1945        3        2 |
  282.  |              Aborigines      Australia   1940        3        2 |
  310.  |               Armenians     Azerbaijan   1991        3        2 |
  391.  |  Chittagong Hill Tribes     Bangladesh   1975        3        2 |
  412.  |                  Hindus     Bangladesh   1975        3        2 |
        |-----------------------------------------------------------------|
*/

list group country year belief BELIEF if belief==2 & n1==1
/*      +------------------------------------------------------------------------+
        |                      group          country   year   belief   BELIEF |
        |------------------------------------------------------------------------|
  226.  |         Indigenous Peoples        Argentina   1940        2      -88 |
  480.  | Indigenous Highland Peoples         Bolivia   1940        2      -88 |
  508.  |   Lowland Indigenous Peoples         Bolivia   1940        2      -88 |
  594.  |               Afro-Brazilians         Brazil   1940        2      -88 |
  622.  |             Amazonian Indians         Brazil   1940        2      -88 |
        |------------------------------------------------------------------------|
*/

list group country year belief BELIEF if belief==0 & n1==1
/*      +----------------------------------------------------------------------------+
        |                  group           country   year   belief   BELIEF |
        |----------------------------------------------------------------------------|
   29.  |                 Pashtuns       Afghanistan   1940        0      -77 |
   57.  |                   Tajiks       Afghanistan   1940        0      -77 |
   85.  |                   Uzbeks       Afghanistan   1940        0      -77 |
  140.  |                  Berbers           Algeria   1965        0      -77 |
  163.  |                  Bakongo            Angola   1975        0      -77 |
        |----------------------------------------------------------------------------|
*/

*2 - STATA code for reconciling RELIGS1 variable.

/*Note: Since the MAR 2004-2006 codebook is the only one with a value of 2 = Orthodox,
a value of 2 is only coded for the 2004-2006 data.
Also, although there is no 0 value for Religs1 in the Phase IV data, there were 112
observations that were coded 0 for 1940-2003.
Since religion is relatively static from year to year, it is likely that these groups
can be recoded based on their values for 2004-2006.
```

However, we are erring on the side of caution and recoding these cases as -88.*/

```
*generating Religs1 variable
gen RELIGS1 = .

*reconciling RELIGS1
replace RELIGS1=1 if religs1==1
replace RELIGS1=3 if religs1== 2
replace RELIGS1=4 if religs1== 3
replace RELIGS1=5 if religs1== 4
replace RELIGS1=6 if religs1== 5
replace RELIGS1=7 if religs1== 6
replace RELIGS1=8 if religs1== 7
replace RELIGS1=9 if religs1== 8
replace RELIGS1=10 if religs1== 9
replace RELIGS1=-99 if religs1==-99
replace RELIGS1=-88 if religs1==0
tab religs1
tab RELIGS1


*examples of reconciled groups

list group country year religs1 RELIGS1 if religs1==0 & n1==1
/*     +----------------------------------------------------------------------+
       |                        group            country   year   religs1   RELIGS1 |
       |----------------------------------------------------------------------|
3919.  | Other Indigenous Peoples              Mexico   1940         0       -88 |
3947.  |                 Zapotecs              Mexico   1940         0       -88 |
4707.  |                    Lari   Republic of Congo   1960         0       -88 |
4731.  |                 M'boshi   Republic of Congo   1960         0       -88 |
5773.  |                  Uzbeks          Tajikistan   1991         0       -88 |
       |----------------------------------------------------------------------|
6304.  |                  Tajiks          Uzbekistan   1991         0       -88 |
       +----------------------------------------------------------------------+
*/

list group country year religs1 RELIGS1 if religs1==2 & n1==1
/*     +----------------------------------------------------------------------+
       |                 group            country   year   religs1   RELIGS1 |
       |----------------------------------------------------------------------|
1281.  |     Antillean Blacks         Costa Rica   1940         2         3 |
2298.  |              Africans             Guyana   1970         2         3 |
2504.  |                 Mizos              India   1950         2         3 |
2556.  |                 Nagas              India   1950         2         3 |
3282.  |                Kikuyu              Kenya   1965         2         3 |
       |----------------------------------------------------------------------|
*/

list group country year religs1 RELIGS1 if religs1==3 & n1==1
/*     +------------------------------------------------------------------------+
       |                 group            country   year   religs1   RELIGS1 |
       |------------------------------------------------------------------------|
 113.  |                Greeks             Albania   1945         3         4 |
 282.  |             Aborigines          Australia   1940         3         4 |
 310.  |              Armenians         Azerbaijan   1991         3         4 |
 336.  |               Russians         Azerbaijan   1991         3         4 |
 560.  |                  Serbs             Bosnia   1992         3         4 |
       |------------------------------------------------------------------------|
*/

list group country year religs1 RELIGS1 if religs1==4 & n1==1
/*     +------------------------------------------------------+
       |            group      country   year   religs1   RELIGS1 |
       |------------------------------------------------------|
  85.  |           Uzbeks  Afghanistan   1940         4         5 |
 140.  |          Berbers      Algeria   1965         4         5 |
 323.  |          Lezgins   Azerbaijan   1991         4         5 |
 370.  |           Biharis   Bangladesh   1975         4         5 |
 678.  |            Turks     Bulgaria   1940         4         5 |
```

```
        |------------------------------------------------------------|
*/

list group country year religs1 RELIGS1 if religs1==5 & n1==1
/*    +--------------------------------------------------------+
      |       group       country   year   religs1   RELIGS1 |
      |--------------------------------------------------------|
   1. |     Hazaras    Afghanistan   1940         5         6 |
 349. |       Shi'is        Bahrain   1975         5         6 |
2766. | Azerbaijanis           Iran   1940         5         6 |
2822. |    Bakhtiari           Iran   1940         5         6 |
2990. |       Shi'is           Iraq   1940         5         6 |
      |--------------------------------------------------------|
*/

list group country year religs1 RELIGS1 if religs1==6 & n1==1
/*    +-------------------------------------------+
      |  group    country   year   religs1   RELIGS1 |
      |-------------------------------------------|
 548. | Muslims    Bosnia   1992         6         7 |
2794. | Baha'is      Iran   1940         6         7 |
3510. |   Druze   Lebanon   1950         6         7 |
4281. | Ahmadis  Pakistan   1950         6         7 |
5632. |   Alawi     Syria   1950         6         7 |
      +-------------------------------------------+
*/

list group country year religs1 RELIGS1 if religs1==7 & n1==1
/*    +------------------------------------------------------------------+
      |                 group       country   year   religs1   RELIGS1 |
      |------------------------------------------------------------------|
 391. | Chittagong Hill Tribes    Bangladesh   1975         7         8 |
 758. |                   Mons         Burma   1950         7         8 |
 810. |                  Shans         Burma   1950         7         8 |
 908. |             Vietnamese      Cambodia   1955         7         8 |
1169. |               Tibetans         China   1940         7         8 |
      |------------------------------------------------------------------|
*/

list group country year religs1 RELIGS1 if religs1==8 & n1==1
/*    +------------------------------------------------------------------------+
      |                       group       country   year   religs1   RELIGS1 |
      |------------------------------------------------------------------------|
 226. |          Indigenous Peoples     Argentina   1940         8         9 |
 480. | Indigenous Highland Peoples       Bolivia   1940         8         9 |
 508. |   Lowland Indigenous Peoples       Bolivia   1940         8         9 |
 572. |             San Bushmen        Botswana   1970         8         9 |
 622. |         Amazonian Indians         Brazil   1940         8         9 |
      |------------------------------------------------------------------------|
*/

list group country year religs1 RELIGS1 if religs1==9 & n1==1
/*    +--------------------------------------------------------------------------+
      |               group                 country   year   religs1   RELIGS1 |
      |--------------------------------------------------------------------------|
 163. |             Bakongo                  Angola   1975         9        10 |
 205. |           Ovimbundu                  Angola   1975         9        10 |
 254. |                Jews               Argentina   1940         9        10 |
 412. |              Hindus              Bangladesh   1975         9        10 |
 459. |          Lhotshampas                 Bhutan   1975         9        10 |
      |--------------------------------------------------------------------------|
*/
```

*3 - STATA code for reconciling LANG variable.

```
/*Note 1: According to the Phase IV codebook, a value of 0 for lang is defined as
unknown.
However, after reviewing the groups that are coded 0, it appears that this is not
entirely correct, as the language of the group can be easily determined.  For example,
the Catholics in Northern Ireland are coded as 0 for lang for 1940-2003, which
```

according to the Phase IV codebook, indicates that their language relative to the ethnic plurality is unknown. However, a quick search provides information indicating that the group has the same language as the plurality. Since language is relatively static, it is likely that the values coded for these groups for 2004-2006 can be used to update their language codes for 1940-2003. However, we are leaving this up to the discretion of researchers and are conservatively recoding these cases as -88.

Note 2: According to the Phase IV codebook, a value of 1 for lang is defined as "groups speaks the same language", which is matched to Phase V value of 2, "group speaks primarily one language, different from plurality group".
To verify that this recoding was correct, we checked whether groups that were lang=1 for 1940-2003 and should therefore be recoded as LANG=2, were also coded as LANG=2 for 2004-2006. We found that there are 38 groups that are coded lang=1 for 1940-2003 and LANG=2 for 2004-2006, as expected.
However, we also found that there are 91 groups that are coded lang=1 for 1940-2003 but LANG=1 for 2004-2006, which is contrary to coding conventions.
It is possible that these ethnic groups changed their language relative to the plurality over time. These groups are noted in the codebook.
It is also possible that the Phase IV value of lang=1 doesn't capture information about the proportion of the group that speaks the same language, whereas the Phase V value of LANG=2 does. While the Phase IV value of lang=1 codes groups that speak the same language (which is distinct from the plurality), it does not specify whether a plurality of the group itself speaks the language, whereas Phase V LANG=2 coding does. Since MAR 1940-2003 lang 1 doesn't distinguish whether a plurality of the group speaks the language, some of the groups that are recoded as MAR 1940-2003 LANG=1 for 1940-2003 are coded MAR 2004-2006 LANG=1, and not LANG=2.
Therefore, it is possible that for some groups, the 2004-2006 codes more precisely specify the extent to which the ethnic group speaks their own language and can be used to update the 1940-2003 codes.
In other cases, it is possible the 2004-2006 codes reflect changes in the group's language that occurred over time.
While it is likely that with additional research researchers can use the 2004-2006 codes to update the 1940-2003 values, we are recoding Phase IV lang=1 as LANG=2 for 1940-2003 and leaving it up to researchers to update the 1940-2003 codes with the 2004-2006 values.*/

```
*generating LANG variable*
gen LANG =.

*reconciling LANG
replace LANG=0 if lang == 3
replace LANG=1 if lang ==2
replace LANG=2 if lang ==1
replace LANG=-99 if lang ==-99
replace LANG=-88 if lang==0
tab lang
tab LANG

*examples of reconciled groups

list group country year lang LANG if lang==0 & n1==1
/*    +-----------------------------------------------------------------------------+
      |                      group                 country    year  lang   LANG |
      |-----------------------------------------------------------------------------|
 254. |                      Jews               Argentina    1940      0    -88 |
 349. |                    Shi'is                 Bahrain    1975      0    -88 |
 412. |                    Hindus              Bangladesh    1975      0    -88 |
 758. |                      Mons                   Burma    1950      0    -88 |
 784. |        Rohingya (Arakanese)                 Burma    1950      0    -88 |
      |-----------------------------------------------------------------------------|
*/

list group country year lang LANG if lang==1 & n1==1
/*    +------------------------------------------------------------------------------+
      |                       group                    country    year    lang   LANG |
```

```
     |-----------------------------------------------------------------------------|
   1. |                   Hazaras                     Afghanistan  1940     1       2 |
  29. |                   Pashtuns                    Afghanistan  1940     1       2 |
  57. |                   Tajiks                      Afghanistan  1940     1       2 |
  85. |                   Uzbeks                      Afghanistan  1940     1       2 |
 113. |                   Greeks                          Albania  1945     1       2 |
     |-----------------------------------------------------------------------------|
*/

list group country year lang LANG if lang==2 & n1==1
/*    +-----------------------------------------------------------------------------+
      |                          group              country   year   lang   LANG |
      |-----------------------------------------------------------------------------|
 163. |                        Bakongo              Angola    1975      2      1 |
 184. |                        Cabinda              Angola    1975      2      1 |
 205. |                      Ovimbundu              Angola    1975      2      1 |
 226. |            Indigenous Peoples             Argentina    1940      2      1 |
 282. |                     Aborigines             Australia    1940      2      1 |
      |-----------------------------------------------------------------------------|
*/

list group country year lang LANG if lang==3 & n1==1
/*    +----------------------------------------------------------------+
      |            group            country   year   lang   LANG |
      |----------------------------------------------------------------|
 140. |            Berbers           Algeria   1965      3      0 |
 908. |         Vietnamese          Cambodia   1955      3      0 |
1890. |            Basques            France   1940      3      0 |
2298. |           Africans            Guyana   1970      3      0 |
2320. |        East Indians          Guyana   1970      3      0 |
      |----------------------------------------------------------------|
*/
*4 - STATA code for reconciling RACE variable.


/*NOTE: The Phase IV (1940-2003 data) codebook notes that race=0 is "unknown", even
though there is already a -99 "missing" value.
Further, the Phase V (2004-2006 and selection bias data) codebook has a "no physical
differences in appearance" (RACE=0) value, but the Phase IV codebook does not.
Since race is a static variable, it is likely that the Phase IV race=0 value should be
recoded as something other than -99 and that the 2004-2006 values for these groups can
be used to replace the coding for 1940-2003.
However, we are conservatively coding these cases -88 and leaving it up to the
discretion of researchers to impute the 2004-2006 values for 1940-2003.*/

*generating RACE variable
gen RACE = .

*reconciling RACE
replace RACE=1 if race== 1
replace RACE=2 if race == 2
replace RACE=3 if race ==3
replace RACE=-99 if race ==-99
replace RACE=-88 if race==0
tab race
tab RACE

*example of reconciled groups

list group country year race RACE if race==0 & n1==1
/*    +-----------------------------------------------------------------------------------+
      |                             group              country   year   race   RACE |
      |-----------------------------------------------------------------------------------|
   1. |                   Hazaras                     Afghanistan  1940      0    -88 |
  29. |                   Pashtuns                    Afghanistan  1940      0    -88 |
  57. |                   Tajiks                      Afghanistan  1940      0    -88 |
  85. |                   Uzbeks                      Afghanistan  1940      0    -88 |
 113. |                   Greeks                          Albania  1945      0    -88 |
      |-----------------------------------------------------------------------------------|*/
```

```
*5 - STATA code for reconciling GC2 variable.

*generating GC2 variable
gen GC2 = .

*reconciling GC2 variable
replace GC2=1 if gc2== 1
replace GC2=0 if gc2== 2
replace GC2=-99 if gc2 ==-99
tab gc2
tab GC2

*example of reconciled groups

list group country year gc2 GC2 if gc2==2 & n1==1
/*     +---------------------------------------------------------------------+
       |                    group            country   year    gc2    GC2 |
       |---------------------------------------------------------------------|
  254. |                     Jews           Argentina   1940      2      0 |
  282. |                Aborigines           Australia   1940      2      0 |
  336. |                  Russians          Azerbaijan   1991      2      0 |
  349. |                    Shi'is             Bahrain   1975      2      0 |
  370. |                   Biharis          Bangladesh   1975      2      0 |
       |---------------------------------------------------------------------|
*/

*6 - STATA code for reconciling GC7 variable.

*generating GC7 variable
gen GC7 = .

*reconciling GC7 variable
replace GC7=1 if gc7== 3
replace GC7=2 if gc7== 2
replace GC7=3 if gc7 ==1
replace GC7=-99 if gc7 ==-99
tab gc7
tab GC7

*examples of reconciled groups

list group country year gc7 GC7 if gc7==1 & n1==1
/*     +----------------------------------------------------------------------+
       |                     group            country   year    gc7    GC7 |
       |----------------------------------------------------------------------|
   85. |                    Uzbeks          Afghanistan   1940      1      3 |
  113. |                    Greeks              Albania   1945      1      3 |
  184. |                   Cabinda               Angola   1975      1      3 |
  310. |                  Armenians          Azerbaijan   1991      1      3 |
  323. |                   Lezgins          Azerbaijan   1991      1      3 |
       |----------------------------------------------------------------------|
*/

list group country year gc7 GC7 if gc7==3 & n1==1
/*     +-------------------------------------------------------------------+
       |                 group            country   year    gc7    GC7 |
       |-------------------------------------------------------------------|
   57. |                Tajiks        Afghanistan   1940      3      1 |
  560. |                 Serbs             Bosnia   1992      3      1 |
  572. |           San Bushmen           Botswana   1970      3      1 |
  678. |                 Turks           Bulgaria   1940      3      1 |
 1698. |              Russians            Estonia   1991      3      1 |
       |-------------------------------------------------------------------|
*/
```

```
*7 - STATA code for reconciling GC10 variable.

*generating GC10 variable
gen GC10 =.

*reconciling GC10
replace GC10=0 if gc10==1 | gc10==3
replace GC10=1 if gc10==2
replace GC10=2 if gc10==4
replace GC10=3 if gc10==5
replace GC10=-99 if gc10==-99
tab gc10
tab GC10

*examples of reconciled groups

list group country year gc10 GC10 if gc10==1 & n1==1
/*     +-------------------------------------------------------------------+
       |             group                      country    year    gc10   GC10 |
       |-------------------------------------------------------------------|
    1. |           Hazaras              Afghanistan   1940       1      0 |
  282. |         Aborigines                Australia   1940       1      0 |
  933. |           Bamileke                 Cameroon   1960       1      0 |
  957. |             Kirdi                  Cameroon   1960       1      0 |
  981. |         Westerners                 Cameroon   1960       1      0 |
       |-------------------------------------------------------------------|
*/

list group country year gc10 GC10 if gc10==3 & n1==1
/*     +-------------------------------------------+
       |   group         country    year    gc10   GC10 |
       |-------------------------------------------|
2660. | Acehnese      Indonesia   1950       3      0 |
5330. |    Zulus   South Africa   1940       3      0 |
5632. |    Alawi          Syria   1950       3      0 |
6005. | Baganda         Uganda   1965       3      0 |
       +-------------------------------------------+
*/

list group country year gc10 GC10 if gc10==2 & n1==1
/*     +--------------------------------------------------------------+
       |                group           country   year    gc10   GC10 |
       |--------------------------------------------------------------|
  140. |              Berbers           Algeria   1965       2      1 |
  254. |                 Jews         Argentina   1940       2      1 |
  310. |             Armenians        Azerbaijan   1991       2      1 |
  336. |             Russians        Azerbaijan   1991       2      1 |
  349. |                Shi'is           Bahrain   1975       2      1 |
       |--------------------------------------------------------------|
*/

list group country year gc10 GC10 if gc10==4 & n1==1
/*     +-----------------------------------------------------------------+
       |                group              country   year    gc10   GC10 |
       |-----------------------------------------------------------------|
   29. |              Pashtuns          Afghanistan   1940       4      2 |
  113. |                Greeks             Albania   1945       4      2 |
  323. |               Lezgins          Azerbaijan   1991       4      2 |
  536. |                Croats              Bosnia   1992       4      2 |
  548. |               Muslims              Bosnia   1992       4      2 |
       |-----------------------------------------------------------------|
*/

list group country year gc10 GC10 if gc10==5 & n1==1
/*     +-----------------------------------------------------------------+
       |                group              country   year    gc10   GC10 |
       |-----------------------------------------------------------------|
   57. |                Tajiks          Afghanistan   1940       5      3 |
   85. |                Uzbeks          Afghanistan   1940       5      3 |
```

```
163. |              Bakongo                    Angola 1975    5      3 |
184. |              Cabinda                    Angola 1975    5      3 |
226. |   Indigenous Peoples                 Argentina 1940    5      3 |
     |--------------------------------------------------------------------|
*/


*8 - STATA code for reconciling GC11 variable.


*generating GC11 variable
gen GC11 =.

*reconciling GC11
replace GC11=0 if gc11==1
replace GC11=1 if gc11==2
replace GC11=2 if gc11==3
replace GC11=3 if gc11==4
replace GC11=-99 if gc11==-99
tab gc11
tab GC11


*examples of reconciled groups

list group country year gc11 GC11 if gc11==1 & n1==1
/*    +----------------------------------------------------------------+
      |                group            country  year  gc11  GC11 |
      |----------------------------------------------------------------|
 391. | Chittagong Hill Tribes        Bangladesh  1975     1     0 |
1507. |         Haitian Blacks  Dominican Republic 1940     1     0 |
1687. |                  Afars           Eritrea  1993     1     0 |
1711. |                  Afars          Ethiopia  1945     1     0 |
1946. |    Muslim (Noncitizens)            France  1940     1     0 |
      |----------------------------------------------------------------|
*/


list group country year gc11 GC11 if gc11==2 & n1==1
/*    +--------------------------------------------------------------------+
      |                     group            country  year  gc11  GC11 |
      |--------------------------------------------------------------------|
 140. |                   Berbers            Algeria  1965     2     1 |
 163. |                   Bakongo             Angola  1975     2     1 |
 184. |                   Cabinda             Angola  1975     2     1 |
 254. |                      Jews          Argentina  1940     2     1 |
 323. |                   Lezgins         Azerbaijan  1991     2     1 |
      |--------------------------------------------------------------------|
*/


list group country year gc11 GC11 if gc11==3 & n1==1
/*    +--------------------------------------------------------------------+
      |                     group            country  year  gc11  GC11 |
      |--------------------------------------------------------------------|
  29. |                  Pashtuns        Afghanistan  1940     3     2 |
 226. |         Indigenous Peoples          Argentina  1940     3     2 |
 370. |                   Biharis         Bangladesh  1975     3     2 |
 594. |           Afro-Brazilians             Brazil  1940     3     2 |
 836. |             Zomis (Chins)              Burma  1950     3     2 |
      |--------------------------------------------------------------------|
*/


list group country year gc11 GC11 if gc11==4 & n1==1
/*    +----------------------------------------------------------------------+
      |                       group            country  year  gc11  GC11 |
      |----------------------------------------------------------------------|
  57. |                     Tajiks        Afghanistan  1940     4     3 |
  85. |                     Uzbeks        Afghanistan  1940     4     3 |
 113. |                     Greeks            Albania  1945     4     3 |
 310. |                   Armenians         Azerbaijan  1991     4     3 |
 336. |                   Russians         Azerbaijan  1991     4     3 |
      |----------------------------------------------------------------------|
*/
```

*9 – <u>STATA code for reconciling EMIG variable</u>.

```
/*NOTE: There are cases in which dmemec =-99 and demempo==0 and vise versa.
For these cases, we are recoding based on the value of 0, rather than -99; therefore,
groups are recoded EMIG =0 if at least one of the constituent Phase IV variables is
coded 0. */

*generating EMIG variable
gen EMIG =.

*reconciling EMIG
replace EMIG=0 if dmemec==0 & dmempo==0
replace EMIG=1 if dmemec==1 | dmempo==1
replace EMIG=2 if dmemec==2 | dmempo==2
replace EMIG=3 if dmemec==3 | dmempo==3
replace EMIG=-99 if dmemec==-99 & dmempo==-99
replace EMIG=0 if dmemec==-99 & dmempo==0
replace EMIG=0 if dmemec==0 & dmempo==-99
tab dmemec
tab dmempo
tab EMIG

*examples of reconciled groups

list group country year dmemec dmempo EMIG if dmemec==0 & n1==1
/*     +------------------------------------------------------------+
       |          group    country   year   dmemec   dmempo   EMIG |
       |------------------------------------------------------------|
  536. |         Croats     Bosnia   1992        0        2      2 |
  548. |        Muslims     Bosnia   1992        0        3      3 |
  560. |          Serbs     Bosnia   1992        0        0      0 |
 1309. |          Serbs    Croatia   1992        0        3      3 |
 4049. |        Basters    Namibia   1990        0        0      0 |
       |------------------------------------------------------------|
*/

list group country year dmemec dmempo EMIG if dmempo==0 & n1==1
/*     +------------------------------------------------------------+
       |          group    country   year   dmemec   dmempo   EMIG |
       |------------------------------------------------------------|
  560. |          Serbs     Bosnia   1992        0        0      0 |
 4049. |        Basters    Namibia   1990        0        0      0 |
 4063. | East Caprivians   Namibia   1990        0        0      0 |
 4077. |      Europeans    Namibia   1990        0        0      0 |
 4091. |    San Bushmen    Namibia   1990        0        0      0 |
       |------------------------------------------------------------|
*/

list group country year dmemec dmempo EMIG if dmemec==1
/*     +-------------------------------------------------------------------------+
       |                     group     country   year   dmemec   dmempo   EMIG |
       |-------------------------------------------------------------------------|
  137. |                    Greeks     Albania   2001        1        0      1 |
  138. |                    Greeks     Albania   2002        1        0      1 |
  139. |                    Greeks     Albania   2003        1        0      1 |
  149. |                   Berbers     Algeria   1990        1        0      1 |
  151. |                   Berbers     Algeria   1992        1        0      1 |
       |-------------------------------------------------------------------------|
*/

list group country year dmemec dmempo EMIG if dmempo==1
/*     +---------------------------------------------------------------------------+
       |                     group      country   year   dmemec   dmempo   EMIG |
       |---------------------------------------------------------------------------|
   52. |                  Pashtuns   Afghanistan   1999        0        1      1 |
   53. |                  Pashtuns   Afghanistan   2000        0        1      1 |
   71. |                    Tajiks   Afghanistan   1990        0        1      1 |
   73. |                    Tajiks   Afghanistan   1992        0        1      1 |
```

```
  75. |                         Tajiks    Afghanistan  1994       0       1      1 |
      |-------------------------------------------------------------------------------|
*/

list group country year dmemec dmempo EMIG if dmemec==2
/*    +-------------------------------------------------------------------------+
      |                     group        country   year   dmemec   dmempo   EMIG |
      |-------------------------------------------------------------------------|
 135. |                     Greeks        Albania   1999        2        0      2 |
 136. |                     Greeks        Albania   2000        2        0      2 |
 279. |                       Jews      Argentina   2001        2        1      2 |
 280. |                       Jews      Argentina   2002        2        1      2 |
 281. |                       Jews      Argentina   2003        2        1      2 |
      |-------------------------------------------------------------------------|
*/

list group country year dmemec dmempo EMIG if dmempo==2
/*    +-----------------------------------------------------------------------------------+
      |                     group            country   year   dmemec   dmempo   EMIG |
      |-----------------------------------------------------------------------------------|
  24. |                   Hazaras        Afghanistan   1999        0        2      2 |
  25. |                   Hazaras        Afghanistan   2000        0        2      2 |
  27. |                   Hazaras        Afghanistan   2002        0        2      2 |
  28. |                   Hazaras        Afghanistan   2003        0        2      2 |
  99. |                    Uzbeks        Afghanistan   1990        0        2      2 |
      |-----------------------------------------------------------------------------------|
*/

list group country year dmemec dmempo EMIG if dmemec==3
/*    +-----------------------------------------------------------------+
      |                 group     country   year   dmemec   dmempo   EMIG |
      |-----------------------------------------------------------------|
 126. |                 Greeks     Albania   1990        3        1      3 |
 128. |                 Greeks     Albania   1992        3        1      3 |
 130. |                 Greeks     Albania   1994        3        1      3 |
 692. |                  Turks    Bulgaria   1990        3        2      3 |
 694. |                  Turks    Bulgaria   1992        3        0      3 |
      |-----------------------------------------------------------------|
*/
list group country year dmemec dmempo EMIG if dmempo==3
/*    +-------------------------------------------------------------------------------+
      |                     group         country   year   dmemec   dmempo   EMIG |
      |-------------------------------------------------------------------------------|
  15. |                   Hazaras     Afghanistan   1990        0        3      3 |
  17. |                   Hazaras     Afghanistan   1992        0        3      3 |
  19. |                   Hazaras     Afghanistan   1994        0        3      3 |
  26. |                   Hazaras     Afghanistan   2001        0        3      3 |
  43. |                   Pashtuns    Afghanistan   1990        0        3      3 |
      |-------------------------------------------------------------------------------|
*/
```

*10 - STATA code for reconciling CULPO2 variable.

/*NOTE: There are cases in which culpo2 =-99 and culpo3==0 and vise versa.
For these cases, we are recoding based on the value of 0, rather than -99; therefore,
groups are recoded CULPO2 =0 if at least one of the constituent Phase IV variables is
coded 0. */

*generating CULPO2 variable
gen CULPO2=.

*reconciling CULPO2
replace CULPO2=0 if culpo2==0 & culpo3==0
replace CULPO2=1 if culpo2==1 | culpo3==1
replace CULPO2=2 if culpo2==2 | culpo3==2
replace CULPO2=3 if culpo2==3 | culpo3==3
replace CULPO2=-99 if culpo2==-99 & culpo3==-99
replace CULPO2=0 if culpo2==0 & culpo3==-99
replace CULPO2=0 if culpo2==-99 & culpo3==0

```
tab culpo2
tab culpo3
tab CULPO2

*examples of reconciled groups

list group country year culpo2 culpo3 CULPO2 if culpo2==0 & n1==1
/*    |----------------------------------------------------------------|
4063. | East Caprivians  Namibia  1990       0         0        0 |
4077. |       Europeans  Namibia  1990       0         2        2 |
4091. |    San Bushmen   Namibia  1990       0         0        0 |
4811. |          Avars    Russia  1992       0         0        0 |
4823. |          Buryat   Russia  1992       0         1        1 |
      |----------------------------------------------------------------|
*/

list group country year culpo2 culpo3 CULPO2 if culpo3==0 & n1==1
/*    +----------------------------------------------------------------+
      |          group   country   year   culpo2   culpo3   CULPO2 |
      |----------------------------------------------------------------|
 536. |         Croats    Bosnia   1992        0        0        0 |
 548. |        Muslims    Bosnia   1992        0        0        0 |
 560. |          Serbs    Bosnia   1992        0        0        0 |
1309. |          Serbs   Croatia   1992        0        0        0 |
4049. |        Basters   Namibia   1990        0        0        0 |
      |----------------------------------------------------------------|
*/


list group country year culpo2 culpo3 CULPO2 if culpo2==1
/*    +--------------------------------------------------------------------------------+
      |                     group          country   year   culpo2   culpo3   CULPO2 |
      |--------------------------------------------------------------------------------|
 155. |                   Berbers          Algeria   1996        1        3        3 |
 240. |        Indigenous Peoples        Argentina   1990        1        2        2 |
 242. |        Indigenous Peoples        Argentina   1992        1        2        2 |
 244. |        Indigenous Peoples        Argentina   1994        1        2        2 |
 246. |        Indigenous Peoples        Argentina   1996        1        2        2 |
      |--------------------------------------------------------------------------------|
*/

list group country year culpo2 culpo3 CULPO2 if culpo3==1
/*    +--------------------------------------------------------------------------------+
      |                     group          country   year   culpo2   culpo3   CULPO2 |
      |--------------------------------------------------------------------------------|
 251. |        Indigenous Peoples        Argentina   2001        0        1        1 |
 252. |        Indigenous Peoples        Argentina   2002        0        1        1 |
 253. |        Indigenous Peoples        Argentina   2003        0        1        1 |
 440. |                    Poles          Belarus   1998        0        1        1 |
 500. | Indigenous Highland Peoples       Bolivia   1996        1        1        1 |
      |--------------------------------------------------------------------------------|
*/

list group country year culpo2 culpo3 CULPO2 if culpo2==2
/*    |----------------------------------------------------------------|
 159. |        Berbers        Algeria   2000        2        3        3 |
 160. |        Berbers        Algeria   2001        2        3        3 |
 161. |        Berbers        Algeria   2002        2        0        2 |
 324. |        Lezgins     Azerbaijan   1992        2        2        2 |
 326. |        Lezgins     Azerbaijan   1994        2        2        2 |
      |----------------------------------------------------------------|
*/

list group country year culpo2 culpo3 CULPO2 if culpo3==2
/*    +----------------------------------------------------------------------------+
      |                       group      country   year   culpo2   culpo3   CULPO2 |
      |----------------------------------------------------------------------------|
 156. |                     Berbers      Algeria   1997        2        2        2 |
```

```
240. |           Indigenous Peoples     Argentina    1990        1          2          2 |
242. |           Indigenous Peoples     Argentina    1992        1          2          2 |
244. |           Indigenous Peoples     Argentina    1994        1          2          2 |
246. |           Indigenous Peoples     Argentina    1996        1          2          2 |
     |---------------------------------------------------------------------------------|
*/

list group country year culpo2 culpo3 CULPO2 if culpo2==3
/*   +---------------------------------------------------------------------------------+
     |                         group     country    year   culpo2   culpo3   CULPO2 |
     |---------------------------------------------------------------------------------|
157. |                        Berbers     Algeria    1998        3        3        3 |
802. |          Rohingya (Arakanese)        Burma    1996        3        2        3 |
803. |          Rohingya (Arakanese)        Burma    1997        3        2        3 |
804. |          Rohingya (Arakanese)        Burma    1998        3        2        3 |
805. |          Rohingya (Arakanese)        Burma    1999        3        2        3 |
     |---------------------------------------------------------------------------------|
*/

list group country year culpo2 culpo3 CULPO2 if culpo3==3
/*   +---------------------------------------------------------------------------------+
     |                         group     country    year   culpo2   culpo3   CULPO2 |
     |---------------------------------------------------------------------------------|
149. |                        Berbers     Algeria    1990        2        3        3 |
151. |                        Berbers     Algeria    1992        2        3        3 |
153. |                        Berbers     Algeria    1994        2        3        3 |
155. |                        Berbers     Algeria    1996        1        3        3 |
157. |                        Berbers     Algeria    1998        3        3        3 |
     |---------------------------------------------------------------------------------|
*/
```

*11 - STATA code for reconciling POLGR variable.

```
*generating POLGR variable
gen POLGR =.

*reconciling POLGR variable
replace POLGR=-99 if autgr1==99 | autgr2==-99 | autgr3==-99 | autgr4==-99 | autgr5==-
99 | autgr6==-99 | polgr1==-99 | polgr2==-99 | polgr3==-99 | polgr4==-99 | polgr5==-
99| polgr6==-99
replace POLGR=0 if autgr3==0 | autgr2==0 | autgr1==0 | autgr4==0 | autgr5==0 |
autgr6==0 | polgr1==0 | polgr2==0 | polgr3==0 | polgr4==0 | polgr5==0 | polgr6==0
replace POLGR=1 if polgr1>0 | polgr2>0 | polgr3>0 | polgr4>0 | polgr5>0 | polgr6>0
replace POLGR=3 if autgr1>0 | autgr4>0 | autgr5>0 | autgr6>0
replace POLGR=4 if autgr3>0 | autgr2>0

tab POLGR
list autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5
polgr6 POLGR
```

*examples of reconciled groups

```
list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if autgr1>0
/*   +----------------------------------------------------------------------------------------------------------------------------+
     |                       group   autgr1   autgr2   autgr3   autgr4   autgr5   autgr6   polgr1   polgr2   polgr3   polgr4   polgr5   polgr6   POLGR |
     |----------------------------------------------------------------------------------------------------------------------------|
149. |                      Berbers        2        0        0        0        0      -99        0        2        3        0        3        0       3 |
414. |                       Hindus        1        0        0        0      -99      -99        0        0        0        0        0        0       3 |
494. | Indigenous Highland Peoples        2        0        0        0        0      -99        0        2        1        1        3        0       3 |
580. |               San Bushmen        1        0        0        0        0      -99        0        0        0        1        0        0       3 |
986. |               Westerners        1        0        0        0      -99      -99        0        1        1        0        0        0       3 |
     |----------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if autgr2>0
/*   +----------------------------------------------------------------------------------------------------------------------------+
     |                       group   autgr1   autgr2   autgr3   autgr4   autgr5   autgr6   polgr1   polgr2   polgr3   polgr4   polgr5   polgr6   POLGR |
     |----------------------------------------------------------------------------------------------------------------------------|
126. |                       Greeks        0        2        0        0        0      -99        0        2        2        0        0        0       4 |
131. |                       Greeks      -99        3        0        0        0        0      -99        2        3        0        0        0       4 |
165. |                      Bakongo        0        1        0        1      -99      -99        0        1        1        1        0        0       4 |
240. |           Indigenous Peoples        0        3        0        3        0      -99        0        2        2        1        0        0       4 |
245. |           Indigenous Peoples      -99        3        0        3        0        0      -99        2        2        1        0        0       4 |
     |----------------------------------------------------------------------------------------------------------------------------|
*/
```

```
list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if autgr3>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
 393. | Chittagong Hill Tribes      0       1       3       0     -99     -99       0       3       0       3       3       0      4 |
      |---------------------------------------------------------------------------------------------------------------------------|
 424. |                 Hindus    -99       0       3       0       0       0     -99       0       1       1       0       0      4 |
 539. |                 Croats    -99       2       2       1       0       0     -99       0       0       0       0       0      4 |
 545. |                 Croats    -99       2       2       1       0       0     -99       2       2       3       2       0      4 |
 546. |                 Croats    -99       2       2       1       0       0     -99       2       2       3       2       0      4 |
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if autgr4>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
 478. |             Lhotshampas    -99       0       0       3       0       0     -99       3       1       1       1       0      3 |
 479. |             Lhotshampas    -99       0       0       3       0       0     -99       3       1       1       1       0      3 |
 539. |                 Croats    -99       2       2       1       0       0     -99       0       0       0       0       0      4 |
 545. |                 Croats    -99       2       2       1       0       0     -99       2       2       3       2       0      4 |
 546. |                 Croats    -99       2       2       1       0       0     -99       2       2       3       2       0      4 |
      |---------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if autgr5>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
 154. |                Berbers    -99       0       0       3       2       0     -99       1       2       0       2       0      3 |
 160. |                Berbers    -99       0       0       2       2       0     -99       0       0       0       2       0      3 |
 161. |                Berbers    -99       0       0       2       2       0     -99       0       0       0       2       0      3 |
 162. |                Berbers    -99       0       0       2       2       0     -99       0       0       0       2       0      3 |
 296. |             Aborigines      0       0       3       1       2     -99       0       1       2       2       0       0      4 |
      |---------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if autgr6>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
 217. |              Ovimbundu    -99       0       0       0       0       3     -99       0       1       0       0       0      3 |
 641. |      Amazonian Indians    -99       0       0       0       1       2     -99       1       2       1       2       0      3 |
1999. |                   Roma    -99       0       0       0       0       1     -99       0       0       0       2       1      3 |
2000. |                   Roma    -99       0       0       0       0       1     -99       0       0       0       2       1      3 |
2001. |                   Roma    -99       0       0       0       0       1     -99       0       0       0       2       1      3 |
      |---------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if polgr1>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
 575. |            San Bushmen      0       0       0       0     -99     -99       1       0       0       0       0       0      1 |
 687. |                  Turks      0       0       0       0     -99     -99       3       0       0       0       0       3      1 |
 718. |                Kachins      0       0       0       1       0     -99       2       1       1       1       0       0      3 |
 739. |                 Karens      0       0       0       3     -99     -99       3       0       0       0       0       0      3 |
 843. |          Zomis (Chins)      0       0       0       3     -99     -99       3       0       0       0       0       0      3 |
      |---------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if polgr2>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
 131. |                 Greeks    -99       3       0       0       0       0     -99       2       3       0       0       0      4 |
 137. |                 Greeks    -99       0       0       0       0       0     -99       1       2       3       2       0      1 |
 138. |                 Greeks    -99       0       0       0       0       0     -99       1       2       3       2       0      1 |
 139. |                 Greeks    -99       0       0       0       0       0     -99       1       2       3       2       0      1 |
 149. |                Berbers      2       0       0       0       0     -99       0       2       3       0       3       0      3 |
      |---------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if polgr3>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
 144. |                Berbers      0       0       0       0     -99     -99       0       0       3       3       0       0      1 |
 149. |                Berbers      2       0       0       0       0     -99       0       2       3       0       3       0      3 |
 154. |                Berbers    -99       0       0       3       2       0     -99       1       2       0       2       0      3 |
 165. |                Bakongo      0       1       0       1     -99     -99       0       1       1       1       0       0      4 |
 170. |                Bakongo      0       0       0       0       0     -99       0       0       1       0       0       0      1 |
      |---------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if polgr4>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
 207. |              Ovimbundu      0       0       0       0     -99     -99       0       0       3       1       0       0      1 |
 235. |     Indigenous Peoples      0       0       0       0     -99     -99       0       1       0       1       0       0      1 |
 240. |     Indigenous Peoples      0       3       0       3       0     -99       0       2       2       1       0       0      4 |
 245. |     Indigenous Peoples    -99       3       0       3       0       0     -99       2       2       1       0       0      4 |
 251. |     Indigenous Peoples    -99       2       0       2       0       0     -99       2       2       1       0       0      4 |
      |---------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if polgr5>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
  15. |                Hazaras      0       0       0       0       0     -99       0       0       1       0       1       1      1 |
  55. |               Pashtuns    -99       0       0       0       0       0     -99       1       1       2       1       0      1 |
  56. |               Pashtuns    -99       0       0       0       0       0     -99       1       1       2       1       0      1 |
  71. |                 Tajiks      0       0       0       1       0     -99       0       0       1       0       1       0      3 |
  99. |                 Uzbeks      0       0       0       1       0     -99       0       0       1       0       1       0      3 |
      |---------------------------------------------------------------------------------------------------------------------------|
*/

list group autgr1 autgr2 autgr3 autgr4 autgr5 autgr6 polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 POLGR if polgr6>0
/*    +---------------------------------------------------------------------------------------------------------------------------+
      |                  group  autgr1  autgr2  autgr3  autgr4  autgr5  autgr6  polgr1  polgr2  polgr3  polgr4  polgr5  polgr6  POLGR |
      |---------------------------------------------------------------------------------------------------------------------------|
1262. |     Indigenous Peoples      0       0       0       3     -99     -99       1       0       0       1       0       3      3 |
```

```
1645. |              Copts          0       0       0       0       0     -99       0       0       1       2       1       1       1 |
1678. |  Indigenous Peoples       -99       0       0       0       0       0     -99       0       0       2       0       2       1 |
1751. |             Amhara          0       0       3       1       0     -99       0       0       2       0       0       1       4 |
1865. |       East Indians        -99       0       0       0       0       0     -99       0       1       1       1       2       1 |
      |------------------------------------------------------------------------------------------------------------------------------|
*/
```

*12- <u>STATA code for reconciling ECGR variable</u>.

```
*generating ECGR variable
gen ECGR =.

*reconciling ECGR variable
replace ECGR=-99 if econgr1==-99 | econgr2==-99 | econgr3==-99 | econgr4==-99 |
econgr5==-99 | econgr6==-99
replace ECGR=0 if econgr1==0 | econgr2==0 | econgr3==0 | econgr4==0 | econgr5==0 |
econgr6==0
replace ECGR=1 if econgr1>0
replace ECGR=2 if econgr2>0 | econgr3>0 | econgr4>0 | econgr5>0

*MAR 1940-2003 ECONGR6 - recoding 18 observations (see notes in codebook for detailed
explanation).

*1. Hill Tribes in Burma in 1985
replace ECGR=2 if numcode==77508 & year==1985

*2. Rohingya (Arakanese) in Burma 1985
replace ECGR=1 if numcode==77501 & year==1985

*3. Rohingya (Arakanese) in Burma 1995
replace ECGR=1 if numcode==77501 & year==1995

*4. Turkmen in China in 1985
replace ECGR=2 if numcode==71003 & year==1985

*5. Eritreans in Ethiopia in 1985
replace ECGR=2 if numcode==53002 & year==1985

*6. Palestinians in Lebanon in 1985
replace ECGR=1 if numcode==66003 & year==1985

*7. Merina in Madagascar in 1990
replace ECGR=2 if numcode==58002 & year==1990

*8. Kewri in Mauritania in 1985
replace ECGR=1 if numcode==43501 & year==1985

*9. Pashtuns (Pushtuns) in Pakistan in 1985
replace ECGR=1 if numcode==77005 & year==1985

*10. Chinese in Panama in 1990
replace ECGR=1 if numcode==9503 & year==1990

*11. Basques in Spain in 1985
replace ECGR=2 if numcode==23001 & year==1985

*12. Northern Hill Tribes in Thailand in 1990
replace ECGR=1 if numcode==80003 & year==1990

*13. Europeans in Zimbabwe in 1985
replace ECGR=2 if numcode==55201 & year==1985

*14. Avars in Russia in 1995
replace ECGR=2 if numcode==36524 & year==1995

*15. Roma in Greece in 1995
```

```
replace ECGR=2 if numcode==35002 & year==1995

*16. Roma in Greece in 2001
replace ECGR=2 if numcode==35002 & year==2001

*17. Roma in Greece in 2002
replace ECGR=2 if numcode==35002 & year==2002

*18. Roma in Greece in 2003
replace ECGR=2 if numcode==35002 & year==2003

tab ECGR

*examples of reconciled groups

list group econgr1 econgr2 econgr3 econgr4 econgr5 econgr6 ECGR if econgr1>0
/*    +----------------------------------------------------------------------------------------+
      |                group   econgr1   econgr2   econgr3   econgr4   econgr5   econgr6   ECGR |
      |----------------------------------------------------------------------------------------|
744.  |                Karens         2         0         1         0         0         0      2 |
770.  |                  Mons         2         1         1         0         0         0      2 |
796.  |   Rohingya (Arakanese)        2         0         0         0         0         0      1 |
848.  |          Zomis (Chins)        1         0         0         0         0         0      1 |
919.  |            Vietnamese         2         0         0         0         1         0      2 |
      |----------------------------------------------------------------------------------------|
*/

list group econgr1 econgr2 econgr3 econgr4 econgr5 econgr6 ECGR if econgr2>0
/*    +----------------------------------------------------------------------------------------+
      |                group   econgr1   econgr2   econgr3   econgr4   econgr5   econgr6   ECGR |
      |----------------------------------------------------------------------------------------|
137.  |                Greeks       -99         2         2         0         2         0      2 |
138.  |                Greeks       -99         2         2         0         2         0      2 |
139.  |                Greeks       -99         2         2         0         2         0      2 |
144.  |               Berbers         0         3         3         0         0         0      2 |
182.  |               Bakongo       -99         1         1         0         0         0      2 |
      |----------------------------------------------------------------------------------------|
*/

list group econgr1 econgr2 econgr3 econgr4 econgr5 econgr6 ECGR if econgr3>0
/*    +----------------------------------------------------------------------------------------+
      |                group   econgr1   econgr2   econgr3   econgr4   econgr5   econgr6   ECGR |
      |----------------------------------------------------------------------------------------|
 26.  |               Hazaras       -99         0         2         0         0         0      2 |
 27.  |               Hazaras       -99         0         2         0         0         0      2 |
 28.  |               Hazaras       -99         0         2         0         0         0      2 |
 55.  |              Pashtuns       -99         0         2         0         3         0      2 |
 56.  |              Pashtuns       -99         0         2         0         3         0      2 |
      |----------------------------------------------------------------------------------------|
*/

list group econgr1 econgr2 econgr3 econgr4 econgr5 econgr6 ECGR if econgr4>0
/*    +----------------------------------------------------------------------------------------+
      |                group   econgr1   econgr2   econgr3   econgr4   econgr5   econgr6   ECGR |
      |----------------------------------------------------------------------------------------|
160.  |               Berbers       -99         0         0         2         0         0      2 |
161.  |               Berbers       -99         0         0         2         0         0      2 |
162.  |               Berbers       -99         0         0         2         0         0      2 |
235.  |     Indigenous Peoples        0         1         0         1         3         3      2 |
240.  |     Indigenous Peoples        0         2         2         3         3         0      2 |
      |----------------------------------------------------------------------------------------|
*/

list group econgr1 econgr2 econgr3 econgr4 econgr5 econgr6 ECGR if econgr5>0
/*    +----------------------------------------------------------------------------------------+
      |                group   econgr1   econgr2   econgr3   econgr4   econgr5   econgr6   ECGR |
      |----------------------------------------------------------------------------------------|
 55.  |              Pashtuns       -99         0         2         0         3         0      2 |
 56.  |              Pashtuns       -99         0         2         0         3         0      2 |
137.  |                Greeks       -99         2         2         0         2         0      2 |
138.  |                Greeks       -99         2         2         0         2         0      2 |
139.  |                Greeks       -99         2         2         0         2         0      2 |
      |----------------------------------------------------------------------------------------|
*/

list group econgr1 econgr2 econgr3 econgr4 econgr5 econgr6 ECGR if econgr6>0
/*    +----------------------------------------------------------------------------------------+
```

```
     |                            group   econgr1   econgr2   econgr3   econgr4   econgr5   econgr6   ECGR |
     |----------------------------------------------------------------------------------------------------|
613. |            Afro-Brazilians       -99         2         1         2         3         1      2 |
631. |           Amazonian Indians        0         3         0         0         3         3      2 |
723. |                    Kachins       -99         1         1         0         1         1      2 |
749. |                     Karens       -99         3         1         0         1         1      2 |
791. |        Rohingya (Arakanese)        0         0         0         0         0         3      1 |
     |----------------------------------------------------------------------------------------------------|
*/
```

*13 - STATA code for reconciling CULGR variable.

*generating CULGR variable
gen CULGR =.

*reconciling CULGR variable
replace CULGR=-99 if culgr1==-99 | culgr2==-99 | culgr3==-99 | culgr4==-99 | culgr5==-99 | culgr6==-99
replace CULGR=0 if culgr1==0 | culgr2==0 | culgr3==0 | culgr4==0 | culgr5==0 | culgr6==0
replace CULGR=1 if culgr1>0
replace CULGR=2 if culgr2>0 | culgr3>0 | culgr4>0 | culgr5>0
tab CULGR

*MAR 1940-2003 CULGR6 - recoding 8 observations (see notes in codebook for detailed explanation).

*1. recoding Haitian Blacks in Dominican Republic in 1995
replace CULGR=2 if numcode==4201 & year==1995

*2. recoding Afro-Guyanas in Guyana in 1995
replace CULGR=1 if numcode==11001 & year==1995

*3. recoding Blacks in Panama in 1995
replace CULGR=1 if numcode==9501 & year==1995

*4. recoding Roma in Russia in 1995
replace CULGR=2 if numcode==36517 & year==1995

*5. recoding Nagas in India in 1985
replace CULGR=1 if numcode==75009 & year==1985

*6. recoding Asians in South Africa in 1985
replace CULGR=1 if numcode==56001 & year==1985

*7. recoding Coloreds in South Africa in 1985
replace CULGR=1 if numcode==56003 & year==1985

*8. replace Antillean Blacks in Costa Rica in 1985
replace CULGR=1 if numcode==9401 & year==1985

tab CULGR

*examples of reconciled groups

list group culgr1 culgr2 culgr3 culgr4 culgr5 culgr6 CULGR if culgr1>0
```
/*   +---------------------------------------------------------------------------------------+
     |               group   culgr1   culgr2   culgr3   culgr4   culgr5   culgr6   CULGR |
     |---------------------------------------------------------------------------------------|
263. |                Jews        1        0      -99      -99        0        2       1 |
273. |                Jews        2        1        2        0        1        0       2 |
279. |                Jews        2        1        2        0        1        0       2 |
280. |                Jews        2        1        2        0        1        0       2 |
281. |                Jews        2        1        2        0        1        0       2 |
     |---------------------------------------------------------------------------------------|
*/
```

list group culgr1 culgr2 culgr3 culgr4 culgr5 culgr6 CULGR if culgr2>0

```
/*   +------------------------------------------------------------------------------------------------+
     |                    group   culgr1   culgr2   culgr3   culgr4   culgr5   culgr6   CULGR |
     |------------------------------------------------------------------------------------------------|
 27. |                  Hazaras        2        2        0        0        1        0       2 |
 28. |                  Hazaras        2        2        0        0        1        0       2 |
 84. |                   Tajiks        0        2        0        0        3        0       2 |
112. |                   Uzbeks        0        2        0        2        1        0       2 |
126. |                   Greeks        3        2        2        2        0      -99       2 |
     |------------------------------------------------------------------------------------------------|
*/

list group culgr1 culgr2 culgr3 culgr4 culgr5 culgr6 CULGR if culgr3>0
/*   +------------------------------------------------------------------------------------------------+
     |                    group   culgr1   culgr2   culgr3   culgr4   culgr5   culgr6   CULGR |
     |   |-------------------------------------------------------------------------------------------- |
320. |                Armenians        1        1        2        0        1        0       2 |
321. |                Armenians        1        1        2        0        1        0       2 |
322. |                Armenians        1        1        2        0        1        0       2 |
327. |                  Lezgins        0        0        2        0        0        0       2 |
340. |                 Russians        0        0        3        0        2        0       2 |
     |------------------------------------------------------------------------------------------------|
*/

list group culgr1 culgr2 culgr3 culgr4 culgr5 culgr6 CULGR if culgr4>0
/*   +------------------------------------------------------------------------------------------------+
     |                    group   culgr1   culgr2   culgr3   culgr4   culgr5   culgr6   CULGR |
     |------------------------------------------------------------------------------------------------|
112. |                   Uzbeks        0        2        0        2        1        0       2 |
126. |                   Greeks        3        2        2        2        0      -99       2 |
131. |                   Greeks        3        2        2        2        0        0       2 |
137. |                   Greeks        2        2        2        2        0        0       2 |
138. |                   Greeks        2        2        2        2        0        0       2 |
     |------------------------------------------------------------------------------------------------|
*/

list group culgr1 culgr2 culgr3 culgr4 culgr5 culgr6 CULGR if culgr5>0
/*   +------------------------------------------------------------------------------------------------+
     |                    group   culgr1   culgr2   culgr3   culgr4   culgr5   culgr6   CULGR |
     |------------------------------------------------------------------------------------------------|
 15. |                  Hazaras        0        0        0        0        1      -99       2 |
 20. |                  Hazaras        2        0        0        0        1        0       2 |
 26. |                  Hazaras        2        0        0        0        1        0       2 |
 27. |                  Hazaras        2        2        0        0        1        0       2 |
 28. |                  Hazaras        2        2        0        0        1        0       2 |
     |------------------------------------------------------------------------------------------------|
*/

list group culgr1 culgr2 culgr3 culgr4 culgr5 culgr6 CULGR if culgr6>0
/*   +----------------------------------------------------------------------------------+
     |                group   culgr1   culgr2   culgr3   culgr4   culgr5   culgr6   CULGR |
     |----------------------------------------------------------------------------------|
263. |                 Jews        1        0      -99      -99        0        2       1 |
613. |       Afro-Brazilians        2        2        0        0        0        1       2 |
641. |      Amazonian Indians        2        2        0        0        1        2       2 |
687. |                Turks        3        3      -99      -99        0        3       2 |
723. |               Kachins        1        1        3        3        2        1       2 |
     |----------------------------------------------------------------------------------|
*/
```

*14 - <u>STATA code for reconciling FACTSEV1-FACTSEV3 variables</u>.

```
*reconciling FACSTSEV1-3
tab fccs1
rename fccs1 FACTSEV1
tab FACTSEV1

tab fccs2
rename fccs2 FACTSEV2
tab FACTSEV2

tab fccs3
rename fccs3 FACTSEV3
tab FACTSEV3
```

*15 - <u>STATA code for reconciling CCGROUPSEV1-CCGROUPSEV3 variables</u>.

```
/*Note 1: It is important to note that in the 1940-2003 data, the first intercommunal
conflict severity variable, cc1x-cc3x, is coded every 10 years from 1940-1990 and the
second intercommunal conflict severity variable, gcc1-gcc3, is coded annually from
1990-2000 (although there is 1 group that is coded from 1981-1989, Darfur Muslims in
Sudan).
As such, there is one year of coding overlap between these two variables.
Further, cc1x-cc3x for 1990 records the highest level of conflict for the 1990s
and gcc1-gcc3 for 1990 records the highest level of conflict for 1990 specifically.
Since both variables are coded in the same way (i.e., the exact same coding levels)
and gcc1-gcc3 codes annual data beginning in 1990, CCGROUPSEV1-CCGROUPSEV3was
reconciled with decennial data from cc1x-cc3x for 1940-1980 and annual data from gcc1-
gcc3 for 1990-2003.

Note 2: It is also important to note that the Phase IV intercommunal conflict
variables for 1940-2003, cc1x-cc3x and gcc1-gcc3, are not ordered in descending order
of severity, whereas the Phase V intercommunal conflict variables for 2004-2006,
CCGROUPSEV1-3, are.
This means that for 2004-2006, CCGROUPSEV1 lists the group with the highest severity
of intercommunal conflict, CCGROUPSEV2 lists the group with the second-highest level
of intercommunal conflict, and CCGROUPSEV3 lists the group with the third-highest
level of inter-communal conflict, whereas cc1x-cc3x and gcc1-gcc3 do not.
We are noting these cases, but leaving it up to researchers to re-order the coding for
these groups.*/

*generating CCGROUPSEV1
gen CCGROUPSEV1=.

*reconciling CCGROUPSEV1
replace CCGROUPSEV1=-99 if cc1x==-99  & year < 1990 | gcc1==-99
replace CCGROUPSEV1=0 if cc1x==0  & year < 1990 | gcc1==0
replace CCGROUPSEV1=1 if cc1x==1  & year < 1990 | gcc1==1
replace CCGROUPSEV1=2 if cc1x==2  & year < 1990 | gcc1==2
replace CCGROUPSEV1=3 if cc1x==3  & year < 1990 | gcc1==3
replace CCGROUPSEV1=4 if cc1x==4  & year < 1990 |gcc1==4
replace CCGROUPSEV1=5 if cc1x==5  & year < 1990 |gcc1==5
replace CCGROUPSEV1=6 if cc1x==6  & year < 1990 |gcc1==6
tab CCGROUPSEV1

*checking CCGROUPSEV1
list year cc1x gcc1 CCGROUPSEV1 if cc1x>0 & year<1990
list year cc1x gcc1 CCGROUPSEV1 if year>1989

*generating CCGROUPSEV2
gen CCGROUPSEV2=.

*reconciling CCGROUPSEV2
replace CCGROUPSEV2=-99 if cc2x==-99  & year < 1990 | gcc2==-99
replace CCGROUPSEV2=0 if cc2x==0  & year < 1990 | gcc2==0
replace CCGROUPSEV2=1 if cc2x==1  & year < 1990 | gcc2==1
replace CCGROUPSEV2=2 if cc2x==2  & year < 1990 | gcc2==2
replace CCGROUPSEV2=3 if cc2x==3  & year < 1990 | gcc2==3
replace CCGROUPSEV2=4 if cc2x==4  & year < 1990 |gcc2==4
replace CCGROUPSEV2=5 if cc2x==5  & year < 1990 |gcc2==5
replace CCGROUPSEV2=6 if cc2x==6  & year < 1990 |gcc2==6
tab CCGROUPSEV2

*checking CCGROUPSEV2
list year cc2x gcc2 CCGROUPSEV2 if cc2x>0 & year<1990
list year cc2x gcc2 CCGROUPSEV2 if year>1989

*generating CCGROUPSEV3
gen CCGROUPSEV3=.
```

```
*reconciling CCGROUPSEV3
replace CCGROUPSEV3=-99 if cc3x==-99  & year < 1990 | gcc3==-99
replace CCGROUPSEV3=0 if cc3x==0  & year < 1990 | gcc3==0
replace CCGROUPSEV3=1 if cc3x==1  & year < 1990 | gcc3==1
replace CCGROUPSEV3=2 if cc3x==2  & year < 1990 | gcc3==2
replace CCGROUPSEV3=3 if cc3x==3  & year < 1990 | gcc3==3
replace CCGROUPSEV3=4 if cc3x==4  & year < 1990 |gcc3==4
replace CCGROUPSEV3=5 if cc3x==5  & year < 1990 |gcc3==5
replace CCGROUPSEV3=6 if cc3x==6  & year < 1990 |gcc3==6
tab CCGROUPSEV3

*checking CCGROUPSEV3
list year cc3x gcc3 CCGROUPSEV3 if cc3x>0
list year cc3x gcc3 CCGROUPSEV3 if year>1989

*examples of reconciled groups

list numcode year cc1x cc2x cc3x gcc1 gcc2 gcc3 CCGROUPSEV1 CCGROUPSEV2 CCGROUPSEV3
/*      +-----------------------------------------------------------------------------------------+
     | numcode   year   cc1x   cc2x   cc3x   gcc1   gcc2   gcc3   CCGROU~1   CCGROU~2   CCGROU~3 |
     |-----------------------------------------------------------------------------------------|
 16. |   70001   1991    -99    -99    -99      0      0      0          0          0          0 |
 17. |   70001   1992    -99    -99    -99      6      0      6          6          0          6 |
 18. |   70001   1993    -99    -99    -99      0      6      0          0          6          0 |
 19. |   70001   1994    -99    -99    -99      0      6      0          0          6          0 |
 20. |   70001   1995    -99    -99    -99      0      6      0          0          6          0 |
     |-----------------------------------------------------------------------------------------|
*/
```

*16 - STATA code for merging 1940-2003 proti and prot variables to create PROT variable

```
/*Note: proti codes data every five years from 1940-1999, whereas prot codes data
annually from 1985.
Therefore, using proti variable from 1940-1984 and prot variable from 1985-2003.*/

*generating PROT variable
gen PROT=.

*recoding proti and prot variables to create PROT variable
replace PROT=-99 if proti==-99 & year <1985 | prot==-99
replace PROT=0 if proti==0 & year<1985 | prot==0
replace PROT=1 if proti==1 & year <1985 | prot==1
replace PROT=2 if proti==2 & year<1985 | prot==2
replace PROT=3 if proti==3 & year<1985 | prot==3
replace PROT=4 if proti==4 & year<1985 | prot==4
replace PROT=5 if proti==5 & year <1985 | prot==5

list numcode year proti prot PROT
list numcode year proti prot PROT if year <1985
list numcode year proti prot PROT if year ==1985
list numcode year proti prot PROT if year >1985

/*Note: there was one observation - the Chinese in Thailand for 2003 - that was coded
a 6 for prot.  Since a value of 6 does not exist, we reviewed the notes and sources to
determine how to recode. The notes from the MAR 2003 update Access database are as
follows: "Ethnic Chinese repeatedly criticize the government's policy of affirmative
action towards the Malays, in both education and business (May 31, 2001; CNN
"Malaysian press deal a 'freedom threat'. In 2003, the loudest protests came from the
Chinese, whose schools have had to fend off many assaults on their independence, and
the government had to give ground to them. Maths and science classes in Chinese
primary schools will continue to be taught mainly in Chinese, though there will be
extra tuition in English in those subjects. (April 3, 2003; The Economist "The
Slaughter of scared cows") Since note indicates Chinese are criticizing the
government's policy but not staging any public demonstrations to do so, recoding as
PROT=1.*/
```

```
replace PROT=1 if numcode==80001 & year==2003

*examples of reconciled groups

list country group year proti prot PROT if proti==0 & n1==1
/*    +-------------------------------------------------------------------------------+
      |                       country                      group   year   proti   prot   PROT |
      |-------------------------------------------------------------------------------|
   2. |                    Afghanistan                    Hazaras   1945       0    -99      0 |
   3. |                    Afghanistan                    Hazaras   1950       0    -99      0 |
   4. |                    Afghanistan                    Hazaras   1955       0    -99      0 |
   5. |                    Afghanistan                    Hazaras   1960       0    -99      0 |
   6. |                    Afghanistan                    Hazaras   1965       0    -99      0 |
      |-------------------------------------------------------------------------------|
   7. |                    Afghanistan                    Hazaras   1970       0    -99      0 |
   8. |                    Afghanistan                    Hazaras   1975       0    -99      0 |
   9. |                    Afghanistan                    Hazaras   1980       0    -99      0 |
  10. |                    Afghanistan                    Hazaras   1985       0      0      0 |
  15. |                    Afghanistan                    Hazaras   1990       0      0      0 |
      |-------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if prot==0 & n1==1
/*    +--------------------------------------------------------+
      |     country        group   year   proti   prot   PROT |
      |--------------------------------------------------------|
 433. |     Belarus        Poles   1991     -99      0      0 |
 446. |     Belarus     Russians   1991     -99      0      0 |
 536. |      Bosnia       Croats   1992     -99      0      0 |
 548. |      Bosnia      Muslims   1992     -99      0      0 |
 560. |      Bosnia        Serbs   1992     -99      0      0 |
      |--------------------------------------------------------|
*/

list country group year proti prot PROT if proti==1
/*    +-------------------------------------------------------------------------------+
      |                       country                       group   year   proti   prot   PROT |
      |-------------------------------------------------------------------------------|
  76. |                    Afghanistan                      Tajiks   1995       1      0      0 |
 231. |                      Argentina          Indigenous Peoples   1965       1    -99      1 |
 232. |                      Argentina          Indigenous Peoples   1970       1    -99      1 |
 235. |                      Argentina          Indigenous Peoples   1985       1      1      1 |
 340. |                     Azerbaijan                    Russians   1995       1      1      1 |
      |-------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if prot==1
/*    +-------------------------------------------------------------------------------+
      |                       country                       group   year   proti   prot   PROT |
      |-------------------------------------------------------------------------------|
 239. |                      Argentina          Indigenous Peoples   1989     -99      1      1 |
 251. |                      Argentina          Indigenous Peoples   2001     -99      1      1 |
 291. |                      Australia                   Aborigines   1985       4      1      1 |
 292. |                      Australia                   Aborigines   1986     -99      1      1 |
 293. |                      Australia                   Aborigines   1987     -99      1      1 |
      |-------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if proti==2
/*    +-------------------------------------------------------------------------------+
      |                       country                       group   year   proti   prot   PROT |
      |-------------------------------------------------------------------------------|
 131. |                        Albania                       Greeks   1995       2      2      2 |
 170. |                         Angola                      Bakongo   1990       2      0      0 |
 175. |                         Angola                      Bakongo   1995       2      2      2 |
 207. |                         Angola                     Ovimbundu   1985       2      2      2 |
 212. |                         Angola                     Ovimbundu   1990       2      2      2 |
      |-------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if prot==2
/*    +-------------------------------------------------------------------------------+
      |                       country                       group   year   proti   prot   PROT |
      |-------------------------------------------------------------------------------|
 127. |                        Albania                       Greeks   1991     -99      2      2 |
 128. |                        Albania                       Greeks   1992     -99      2      2 |
 131. |                        Albania                       Greeks   1995       2      2      2 |
 132. |                        Albania                       Greeks   1996     -99      2      2 |
 138. |                        Albania                       Greeks   2002     -99      2      2 |
```

80

```
      |---------------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if proti==3
/*    +------------------------------------------------------------------------------------+
      |                       country                group   year   proti   prot    PROT |
      |------------------------------------------------------------------------------------|
  290. |                     Australia             Aborigines   1980      3    -99       3 |
  296. |                     Australia             Aborigines   1990      3      3       3 |
  314. |                    Azerbaijan               Armenians   1995      3      2       2 |
  327. |                    Azerbaijan                 Lezgins   1995      3      2       2 |
  350. |                       Bahrain                  Shi'is   1980      3    -99       3 |
      |------------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if prot==3
/*    +------------------------------------------------------------------------------------+
      |                       country                group   year   proti   prot    PROT |
      |------------------------------------------------------------------------------------|
  301. |                     Australia             Aborigines   1995      4      3       3 |
  303. |                     Australia             Aborigines   1997    -99      3       3 |
  306. |                     Australia             Aborigines   2000    -99      3       3 |
  307. |                     Australia             Aborigines   2001    -99      3       3 |
  308. |                     Australia             Aborigines   2002    -99      3       3 |
      |------------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if proti==4
/*    +--------------------------------------------------------------------------------------+
      |                       country                 group   year   proti   prot    PROT |
      |--------------------------------------------------------------------------------------|
  143. |                       Algeria                Berbers   1980      4    -99       4 |
  291. |                     Australia             Aborigines   1985      4      1       1 |
  301. |                     Australia             Aborigines   1995      4      3       3 |
  377. |                    Bangladesh                 Biharis   1990      4      3       3 |
  466. |                        Bhutan             Lhotshampas   1990      4      4       4 |
      |--------------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if prot==4
/*    +--------------------------------------------------------------------------------------+
      |                       country                 group   year   proti   prot    PROT |
      |--------------------------------------------------------------------------------------|
  158. |                       Algeria                Berbers   1999    -99      4       4 |
  159. |                       Algeria                Berbers   2000    -99      4       4 |
  160. |                       Algeria                Berbers   2001    -99      4       4 |
  161. |                       Algeria                Berbers   2002    -99      4       4 |
  294. |                     Australia             Aborigines   1988    -99      4       4 |
      |--------------------------------------------------------------------------------------|
*/

list country group year proti prot PROT if proti==5
/*    +-----------------------------------------------------------------------+
      |               country            group   year   proti   prot   PROT |
      |-----------------------------------------------------------------------|
 2443. |                 India         Assamese   1995      5      2      2 |
 2490. |                 India        Kashmiris   1990      5      4      4 |
 2495. |                 India        Kashmiris   1995      5      5      5 |
 2537. |                 India          Muslims   1985      5      2      2 |
 2542. |                 India          Muslims   1990      5      5      5 |
      |-----------------------------------------------------------------------|
*/

list country group year proti prot PROT if prot==5
/*    +-----------------------------------------------------------------------+
      |               country            group   year   proti   prot   PROT |
      |-----------------------------------------------------------------------|
  149. |               Algeria          Berbers   1990      5      5      5 |
  151. |               Algeria          Berbers   1992    -99      5      5 |
  152. |               Algeria          Berbers   1993    -99      5      5 |
  153. |               Algeria          Berbers   1994    -99      5      5 |
  154. |               Algeria          Berbers   1995      5      5      5 |
      |-----------------------------------------------------------------------|
*/
```

/*17 – <u>STATA code for merging 1940-2003 rebel and reb variables to create REB variable</u>

Note: rebel is coded every five years from 1945-1999 and reb is coded annually from 1985. Therefore, using rebel variable from 1940-1984 and reb variable from 1985-2003.*/

```
*generating REB variable
gen REB=.

*recoding rebel and reb variables to create REB variable
replace REB=-99 if rebel==-99 & year<1985| reb==-99
replace REB=0 if rebel==0 & year<1985| reb==0
replace REB=1 if rebel==1 & year<1985 | reb==1
replace REB=2 if rebel==2 & year<1985 | reb==2
replace REB=3 if rebel==3 & year<1985 | reb==3
replace REB=4 if rebel==4 & year<1985 | reb==4
replace REB=5 if rebel==5 & year<1985 | reb==5
replace REB=6 if rebel==6 & year<1985 | reb==6
replace REB=7 if rebel==7 & year<1985 | reb==7

*examples of reconciled groups

list group country year rebel reb REB if rebel==0
/*     +------------------------------------------------------------------------------+
       |                    group          country  year  rebel   reb   REB |
       |------------------------------------------------------------------------------|
  30. |                 Pashtuns       Afghanistan  1945      0   -99     0 |
  31. |                 Pashtuns       Afghanistan  1950      0   -99     0 |
  32. |                 Pashtuns       Afghanistan  1955      0   -99     0 |
       |------------------------------------------------------------------------------|
  33. |                 Pashtuns       Afghanistan  1960      0   -99     0 |
  34. |                 Pashtuns       Afghanistan  1965      0   -99     0 |
  35. |                 Pashtuns       Afghanistan  1970      0   -99     0 |
  36. |                 Pashtuns       Afghanistan  1975      0   -99     0 |
  37. |                 Pashtuns       Afghanistan  1980      0   -99     0 |
       |------------------------------------------------------------------------------|
  38. |                 Pashtuns       Afghanistan  1985      0     0     0 |
  43. |                 Pashtuns       Afghanistan  1990      0     0     0 |
  48. |                 Pashtuns       Afghanistan  1995      0     0     0 |
*/

list group country year rebel reb REB if reb==0
/*     +------------------------------------------------------------------------------+
       |                    group          country  year  rebel   reb   REB |
       |------------------------------------------------------------------------------|
  10. |                  Hazaras       Afghanistan  1985      7     0     0 |
  11. |                  Hazaras       Afghanistan  1986    -99     0     0 |
  12. |                  Hazaras       Afghanistan  1987    -99     0     0 |
  13. |                  Hazaras       Afghanistan  1988    -99     0     0 |
  14. |                  Hazaras       Afghanistan  1989    -99     0     0 |
       |------------------------------------------------------------------------------|
*/

list group country year rebel reb REB if rebel==1
/*     +------------------------------------------------------------------------------+
       |                    group          country  year  rebel   reb   REB |
       |------------------------------------------------------------------------------|
 126. |                   Greeks           Albania  1990      1     0     0 |
 131. |                   Greeks           Albania  1995      1     1     1 |
 140. |                  Berbers           Algeria  1965      1   -99     1 |
 185. |                  Cabinda            Angola  1980      1   -99     1 |
 186. |                  Cabinda            Angola  1985      1     1     1 |
       |------------------------------------------------------------------------------|
*/

list group country year rebel reb REB if reb==1
/*     +------------------------------------------------------------------------------+
       |                    group          country  year  rebel   reb   REB |
       |------------------------------------------------------------------------------|
 130. |                   Greeks           Albania  1994    -99     1     1 |
 131. |                   Greeks           Albania  1995      1     1     1 |
 160. |                  Berbers           Algeria  2001    -99     1     1 |
 186. |                  Cabinda            Angola  1985      1     1     1 |
 187. |                  Cabinda            Angola  1986    -99     1     1 |
```

```
      |----------------------------------------------------------------------------|
*/

list group country year rebel reb REB if rebel==2
/*    +---------------------------------------------------------------------------+
      |                      group           country   year   rebel   reb   REB |
      |----------------------------------------------------------------------------|
  361. |                     Shi'is           Bahrain   1995       2     2     2 |
  885. |                     Tutsis           Burundi   1965       2   -99     2 |
 1066. |                  Quebecois            Canada   1965       2   -99     2 |
 1067. |                  Quebecois            Canada   1970       2   -99     2 |
 1211. |                    Turkmen             China   1990       2     2     2 |
      |----------------------------------------------------------------------------|
*/

list group country year rebel reb REB if reb==2
/*    +---------------------------------------------------------------------------+
      |                      group           country   year   rebel   reb   REB |
      |----------------------------------------------------------------------------|
   55. |                   Pashtuns       Afghanistan   2002     -99     2     2 |
  191. |                    Cabinda            Angola   1990       4     2     2 |
  361. |                     Shi'is           Bahrain   1995       2     2     2 |
 1211. |                    Turkmen             China   1990       2     2     2 |
 1212. |                    Turkmen             China   1991     -99     2     2 |
      |----------------------------------------------------------------------------|
*/

list group country year rebel reb REB if rebel==3
/*    +---------------------------------------------------------------------------+
      |                      group           country   year   rebel   reb   REB |
      |----------------------------------------------------------------------------|
  466. |                Lhotshampas            Bhutan   1990       3     3     3 |
  630. |           Amazonian Indians           Brazil   1980       3   -99     3 |
  636. |           Amazonian Indians           Brazil   1990       3     1     1 |
  863. |                      Hutus           Burundi   1970       3   -99     3 |
  899. |                     Tutsis           Burundi   1995       3     0     0 |
      |----------------------------------------------------------------------------|
*/

list group country year rebel reb REB if reb==3
/*    +--------------------------------------------------------------------------+
      |              group           country   year   rebel   reb   REB |
      |----------------------------------------------------------------------------|
  217. |           Ovimbundu            Angola   1995       7     3     3 |
  315. |           Armenians        Azerbaijan   1996     -99     3     3 |
  317. |           Armenians        Azerbaijan   1998     -99     3     3 |
  318. |           Armenians        Azerbaijan   1999     -99     3     3 |
  319. |           Armenians        Azerbaijan   2000     -99     3     3 |
      |----------------------------------------------------------------------------|
*/

list group country year rebel reb REB if rebel==4
/*    +----------------------------------------------------------------+
      |              group        country   year   rebel   reb   REB |
      |----------------------------------------------------------------|
  184. |            Cabinda         Angola   1975       4   -99     4 |
  191. |            Cabinda         Angola   1990       4     2     2 |
  196. |            Cabinda         Angola   1995       4     4     4 |
  707. |            Kachins          Burma   1955       4   -99     4 |
  789. | Rohingya (Arakanese)       Burma   1975       4   -99     4 |
      |----------------------------------------------------------------|
*/

list group country year rebel reb REB if reb==4
/*    +----------------------------------------------------------------------+
      |                group           country   year   rebel   reb   REB |
      |----------------------------------------------------------------------|
   56. |              Pashtuns       Afghanistan   2003     -99     4     4 |
  192. |               Cabinda            Angola   1991     -99     4     4 |
  193. |               Cabinda            Angola   1992     -99     4     4 |
  194. |               Cabinda            Angola   1993     -99     4     4 |
```

```
 195. |                    Cabinda                Angola 1994      -99       4       4 |
      |----------------------------------------------------------------------------------|
 */

list group country year rebel reb REB if rebel==5
/*    +---------------------------------------------------------------------------+
      |                 group      country   year   rebel    reb   REB |
      |----------------------------------------------------------------------------|
 843. |         Zomis (Chins)         Burma   1985       5      5     5 |
 848. |         Zomis (Chins)         Burma   1990       5      5     5 |
 871. |                 Hutus       Burundi   1990       5      3     3 |
1773. |                 Oromo      Ethiopia   1985       5      5     5 |
1778. |                 Oromo      Ethiopia   1990       5      5     5 |
      |----------------------------------------------------------------------------|
 */

list group country year rebel reb REB if reb==5
/*    +--------------------------------------------------------------------------------+
      |                  group        country   year   rebel    reb   REB |
      |------------------------------------------------------------------------------------|
  17. |                Hazaras    Afghanistan   1992     -99      5     5 |
  18. |                Hazaras    Afghanistan   1993     -99      5     5 |
  19. |                Hazaras    Afghanistan   1994     -99      5     5 |
 103. |                 Uzbeks    Afghanistan   1994     -99      5     5 |
 393. | Chittagong Hill Tribes     Bangladesh   1985       5      5     5 |
      |------------------------------------------------------------------------------------|
 */

list group country year rebel reb REB if rebel==6
/*    +--------------------------------------------------------------------------------+
      |                  group        country   year   rebel    reb   REB |
      |------------------------------------------------------------------------------------|
 815. |                  Shans          Burma   1975       6    -99     6 |
 816. |                  Shans          Burma   1980       6    -99     6 |
 876. |                  Hutus        Burundi   1995       6      4     4 |
1093. |             Southerners           Chad   1980       6    -99     6 |
1171. |               Tibetans          China   1950       6    -99     6 |
      |------------------------------------------------------------------------------------|
 */

list group country year rebel reb REB if reb==6
/*    +--------------------------------------------------------------------------------+
      |                  group        country   year   rebel    reb   REB |
      |------------------------------------------------------------------------------------|
  24. |                Hazaras    Afghanistan   1999     -99      6     6 |
  25. |                Hazaras    Afghanistan   2000     -99      6     6 |
  26. |                Hazaras    Afghanistan   2001     -99      6     6 |
 178. |                Bakongo         Angola   1998     -99      6     6 |
 203. |                Cabinda         Angola   2002     -99      6     6 |
      |------------------------------------------------------------------------------------|
 */

list group country year rebel reb REB if rebel==7
/*    +--------------------------------------------------------------------------------+
      |                  group        country   year   rebel    reb   REB |
      |------------------------------------------------------------------------------------|
   9. |                Hazaras    Afghanistan   1980       7    -99     7 |
  10. |                Hazaras    Afghanistan   1985       7      0     0 |
  20. |                Hazaras    Afghanistan   1995       7      0     0 |
  65. |                 Tajiks    Afghanistan   1980       7    -99     7 |
  66. |                 Tajiks    Afghanistan   1985       7      7     7 |
      |------------------------------------------------------------------------------------|
 */

list group country year rebel reb REB if reb==7
/*    +--------------------------------------------------------------------------------+
      |                 group       country   year   rebel    reb   REB |
      |----------------------------------------------------------------------------------|
  21. |               Hazaras   Afghanistan   1996     -99      7     7 |
  22. |               Hazaras   Afghanistan   1997     -99      7     7 |
  23. |               Hazaras   Afghanistan   1998     -99      7     7 |
```

```
  66. |          Tajiks                   Afghanistan 1985        7       7      7 |
  67. |          Tajiks                   Afghanistan 1986      -99       7      7 |
      |-------------------------------------------------------------------------|
*/


/*Now renaming variables that were not reconciled but are included in the reconciled
dataset.
Specifically, renaming them from lower-case to upper-case to match spelling in MAR
2004-2006 data.
Note: not renaming group, poldis, and ecdis in order to merge with Asal data (see
below).*/

rename region AMAR_Region
rename gpop GPOP
rename cpop CPOP
rename gpro GPRO
rename custom CUSTOM
rename groupcon GROUPCON
rename gc119 GC119
rename autlost AUTLOST
rename autonend AUTONEND
rename transyr TRANSYR
rename sepkin SEPKIN
rename culpo1 CULPO1
rename gojpa GOJPA
rename auton2 AUTON2
rename autgain AUTGAIN
rename autpro AUTPRO
rename intracon INTRACON
rename factcc1 FACTCC1
rename factcc2 FACTCC2
rename factcc3 FACTCC3
rename intercon INTERCON
rename ccgroup1 CCGROUP1
rename ccgroup2 CCGROUP2
rename cc3group CCGROUP3


*saving this version of the data before dropping variables that were reconciled but
don't appear in Phase V
save "C:\Users\Aga\Desktop\mardata02.2005.3.1.dta", replace


*Now dropping variables from Phase IV that were reconciled but that are not coded in
Phase V onward.*/
drop autgr1 autgr2 autgr3 autgr4 autgr5 autgr6  belief cc1x cc2x cc3x cc3xa ccx1 ccx2
ccx2 ccx3 culgr1 culgr2 culgr3 culgr4 culgr5 culgr6 culpo2 culpo3 dmemec dmempo
econgr1 econgr2 econgr3 econgr4 econgr5 econgr6 factcc4 fccs4 fccx1 fccx2 fccx3 fccx4
gc2 gc6b gc7 gc10 gc11 gcc2 gcc3 lang polgr1 polgr2 polgr3 polgr4 polgr5 polgr6 prot
proti race reb rebel religs1 sepx

save "C:\Users\Aga\Desktop\mardata02.2005.4.dta", replace

clear

********************************************************************************************


/*III. The following code replaces the POLDIS and ECDIS coding for 1980-1989.
These variables were coded quiquennially from 1940-1985 and annually from 1985-2003 in
the Phase IV 1940-2003 data.
However, as part of his dissertation project at UMD College Park, Victor Asal
(currently at SUNY-Albany) recoded and updated these variables to include annual
coding for 1980-1989.
As such, the Asal data replaces the POLDIS and ECDIS data in the Phase IV 1940-2003
dataset for 1980-1989 only.*/
```

```
insheet using "C:\Users\Aga\Desktop\AMAR_AMY for checking March 10
2015\discrimination_data_1980-1989.csv"
save "C:\Users\Aga\Desktop\MAR_AsalPoldisEcdis19801989.dta"
clear


*1. prepping dataset-
use "C:\Users\Aga\Desktop\MAR_AsalPoldisEcdis19801989.dta", clear

*correcting group names
replace group="Catholics in Northern Ireland" if group=="Catholics" & scode=="UKG"
replace group="Diolas in Casamance" if group=="Diola"
replace group="Native Americans" if numcode==203
replace group="Haitian Blacks" if numcode==4201
replace group="Other Indigenous Peoples" if numcode==7004
replace group="Blacks" if numcode==9501
replace group="Blacks" if numcode==10001
replace group="Blacks" if numcode==10101
replace group="Africans" if numcode==11001
replace group="Blacks" if numcode==13001
replace group="Indigenous Highland Peoples" if numcode==13002
replace group="Lowland Indigenous Peoples" if numcode==13003
replace group="Blacks (Afro-Peruvians)" if numcode==13501
replace group="Indigenous Highland Peoples" if numcode==13502
replace group="Lowland Indigenous Peoples" if numcode==13503
replace group="Amazonian Indians" if numcode==14002
replace group="Indigenous Highland Peoples" if numcode==14501
replace group="Lowland Indigenous Peoples" if numcode==14502
replace group="Muslim (Noncitizens)" if numcode==22006
replace group="Jurassians" if numcode==22501
replace group="South Tyrolians" if numcode==32501
replace group="Muslims" if numcode==35001
replace group="Magyars (Hungarians)" if numcode==36002
replace group="Luba" if numcode==49003
replace group="Lunda, Yeke" if numcode==49005
replace group="Maasai" if numcode==50105
replace group="Kalenjin" if numcode==50109
replace group="Tutsis" if numcode==51701
replace group="Hutus" if numcode==51702
replace group="Ovimbundu" if numcode==54002
replace group="Bemba" if numcode==55101
replace group="Zulus" if numcode==56006
replace group="Turkmen" if numcode==63008
replace group="Alawi" if numcode==65201
replace group="Hui Muslims" if numcode==71002
replace group="Turkmen" if numcode==71003
replace group="Lhotshampas" if numcode==76001
replace group="Pashtuns (Pushtuns)" if numcode==77005
replace group="Rohingya (Arakanese)" if numcode==77501
replace group="Zomis (Chins)" if numcode==77502
replace group="Shans" if numcode==77507
replace group="Moros" if numcode==84003
replace group="Aborigines" if numcode==90001

*note: assigning AMAR numcode 26001 to Asal Turks in Germany with numcode of 25501
because both are Turks in Germany, and only difference is old numcode for Germany used
in Asal data
replace numcode=26001 if numcode==25501

*note: assigning AMAR numcode 36505 to Asal Tatars in USSR with numcode 36405 because
both are Tatars in Russia, and only difference is USSR numcode used for Russia rather
than Russia numcode in Asal data
*see - http://www-personal.umich.edu/~janavs/cowcodes.do
```

```
replace numcode=36505 if numcode==36405

*note: assigning AMAR numcode 81602 to Asal Chinese in S. Vietnam with numcode 81802
because both are Chinese groups, and only difference is Asal used old numcode for
Vietnam
*see - http://www-personal.umich.edu/~janavs/cowcodes.do
replace numcode=81602 if numcode==81802

*note: assigning AMAR numcode 81604 to Asal Montagnards in S. Vietnam with numcode
81804 because both are Montagnard groups, and only difference is Asal used old numcode
*see - http://www-personal.umich.edu/~janavs/cowcodes.do
replace numcode=81604 if numcode==81804

sort numcode year

*saving as "MAR_AsalPoldisEcdis19801989.dta"
save "C:\Users\Aga\Desktop\MAR_AsalPoldisEcdis19801989.1.dta", replace
clear

*2 - merging Asal 1980-1989 data with MAR 1940-2003 database
use "C:\Users\Aga\Desktop\mardata02.2005.4.dta", clear
sort numcode year group

*creating variable based on 1940-2003 poldis and ecdis to compare values that are
replaced with Asal data
gen poldis_check =.
replace poldis_check = poldis
gen ecdis_check=.
replace ecdis_check= ecdis

merge 1:1 numcode group year using
"C:\Users\Aga\Desktop\MAR_AsalPoldisEcdis19801989.1.dta", update replace

/*Note: the following is the result of the merge
. merge 1:1 numcode group year using
"C:\Users\Aga\Desktop\MAR_AsalPoldisEcdis19801989.1.dta", update replace
    Result                           # of obs.
    -----------------------------------------
    not matched                          6,799
        from master                      5,299  (_merge==1)
        from using                       1,500  (_merge==2)

    matched                              1,351
        not updated                        173  (_merge==3)
        missing updated                      0  (_merge==4)
        nonmissing conflict              1,178  (_merge==5)
    -----------------------------------------
*/

*3 - checking merge

*a- *NOTE: the following groups from AMAR were not matched because they do not exist
in the Asal 1980-1989 data
list numcode group year poldis ecdis _merge if _merge==1 & year==1980 | _merge==1 &
year==1985 | _merge==1 & year==1986 | _merge==1 & year==1987 | _merge==1 & year==1988
| _merge==1 & year==1989
/*      +-----------------------------------------------------------------------+
        | numcode                  group   year   poldis   ecdis        _merge |
        |-----------------------------------------------------------------------|
3031. |   47506                    Ijaw   1980        4       4   master only (1) |
3032. |   47506                    Ijaw   1985      -99     -99   master only (1) |
3033. |   47506                    Ijaw   1986      -99     -99   master only (1) |
3034. |   47506                    Ijaw   1987      -99     -99   master only (1) |
```

```
3035. |   47506                    Ijaw  1988     -99     -99   master only (1) |
      |------------------------------------------------------------------------|
3036. |   47506                    Ijaw  1989     -99     -99   master only (1) |
3103. |   48402                 M'boshi  1980       3       0   master only (1) |
3104. |   48402                 M'boshi  1985     -99     -99   master only (1) |
3105. |   48402                 M'boshi  1986     -99     -99   master only (1) |
3106. |   48402                 M'boshi  1987     -99     -99   master only (1) |
      |------------------------------------------------------------------------|
3107. |   48402                 M'boshi  1988     -99     -99   master only (1) |
3108. |   48402                 M'boshi  1989     -99     -99   master only (1) |
3454. |   51001               Zanzibaris  1985    -99     -99   master only (1) |
3455. |   51001               Zanzibaris  1986    -99     -99   master only (1) |
3456. |   51001               Zanzibaris  1987    -99     -99   master only (1) |
      |------------------------------------------------------------------------|
3457. |   51001               Zanzibaris  1988    -99     -99   master only (1) |
3458. |   51001               Zanzibaris  1989    -99     -99   master only (1) |
4251. |   62503                    Nuba  1980       4       4   master only (1) |
4252. |   62503                    Nuba  1985     -99     -99   master only (1) |
4253. |   62503                    Nuba  1986     -99     -99   master only (1) |
      |------------------------------------------------------------------------|
4254. |   62503                    Nuba  1987     -99     -99   master only (1) |
4255. |   62503                    Nuba  1988     -99     -99   master only (1) |
4256. |   62503                    Nuba  1989     -99     -99   master only (1) |
4271. |   62504    Darfur Black Muslims  1980       2       2   master only (1) |
4276. |   62504    Darfur Black Muslims  1985       2       2   master only (1) |
      |------------------------------------------------------------------------|
4277. |   62504    Darfur Black Muslims  1986       1       2   master only (1) |
4278. |   62504    Darfur Black Muslims  1987       1       2   master only (1) |
4279. |   62504    Darfur Black Muslims  1988       1       2   master only (1) |
4280. |   62504    Darfur Black Muslims  1989       3       2   master only (1) |
4685. |   65202                   Kurds  1980       4       3   master only (1) |
      |------------------------------------------------------------------------|
4690. |   65202                   Kurds  1985       4       3   master only (1) |
4691. |   65202                   Kurds  1986       4       3   master only (1) |
4692. |   65202                   Kurds  1987       4       3   master only (1) |
4693. |   65202                   Kurds  1988       4       3   master only (1) |
4694. |   65202                   Kurds  1989       4       3   master only (1) |
      +------------------------------------------------------------------------+
*/

*b - NOTE: the following groups from Asal were not matched because they are not
current in the AMAR data
sort numcode group year
drop n1
by numcode group: gen n1 = (_n == 1)
list numcode group year poldis ecdis _merge if _merge==2 & n1==1
/*    +------------------------------------------------------------------------+
      | numcode                   group  year  poldis  ecdis           _merge |
      |------------------------------------------------------------------------|
1429. |   22003                 Bretons  1980       1       1   using only (2) |
1759. |   31501              Hungarians  1980       2       0   using only (2) |
1769. |   31502                    Roma  1980       4       3   using only (2) |
1779. |   31503                  Slovaks  1980      1       1   using only (2) |
2037. |   34503                Slovenes  1980       0       0   using only (2) |
      |------------------------------------------------------------------------|
2175. |   34508                   Serbs  1980       0       0   using only (2) |
2403. |   36001                 Germans  1980       3       3   using only (2) |
2477. |   36401                Armenians  1980      0       0   using only (2) |
2487. |   36402            Azerbaijanis  1980       0       0   using only (2) |
2497. |   36404                Chechens  1980       1       1   using only (2) |
      |------------------------------------------------------------------------|
2507. |   36406               Estonians  1980       2       0   using only (2) |
2517. |   36407               Georgians  1980       0       0   using only (2) |
```

```
2527. |   36408                  Germans   1980       2          0    using only (2) |
2537. |   36410                     Jews   1980       3          0    using only (2) |
2547. |   36411                 Karachai   1980       1          0    using only (2) |
      |---------------------------------------------------------------------------|
2557. |   36412                  Kazakhs   1980       1          1    using only (2) |
2567. |   36413                  Kirghiz   1980       1          1    using only (2) |
2577. |   36414                    Kurds   1980       2          2    using only (2) |
2587. |   36415                 Latvians   1980       2          0    using only (2) |
2597. |   36416              Lithuanians   1980       2          0    using only (2) |
      |---------------------------------------------------------------------------|
2607. |   36417                     Roma   1980       2          2    using only (2) |
2617. |   36418                   Tajiks   1980       1          1    using only (2) |
2627. |   36419                  Turkmen   1980       1          1    using only (2) |
2637. |   36420               Ukrainians   1980       0          0    using only (2) |
2647. |   36421                   Uzbeks   1980       1          1    using only (2) |
      |---------------------------------------------------------------------------|
2657. |   36422                Moldovans   1980       0          0    using only (2) |
2667. |   36424                    Avars   1980       0          0    using only (2) |
2677. |   36428                   Ingush   1980       1          0    using only (2) |
2687. |   36431                  Lezgins   1980       0          0    using only (2) |
2697. |   36433                   Buryat   1980       3          1    using only (2) |
      |---------------------------------------------------------------------------|
2707. |   36434                   Kumyks   1980       2          0    using only (2) |
2717. |   36437                Tuvinians   1980       3          2    using only (2) |
2727. |   36438                    Yakut   1980       0          2    using only (2) |
2954. |   36904           Crimean Tartars  1980       4          4    using only (2) |
3219. |   43601           Djerema-songhai  1980       0          0    using only (2) |
      |---------------------------------------------------------------------------|
3229. |   43602                    Hausa   1980       0          0    using only (2) |
3267. |   43702                 Lebanese   1980       4          0    using only (2) |
3361. |   45001         Americo-Liberians  1980       0          0    using only (2) |
3703. |   47502             Hausa-Fulani   1980       0          0    using only (2) |
3821. |   48301               Northerners   1980       3          0    using only (2) |
      |---------------------------------------------------------------------------|
3911. |   49001                  Bakongo   1980       1          1    using only (2) |
3949. |   49004                  Lingala   1980       0          0    using only (2) |
3987. |   49006              Kivu Region   1980       2          2    using only (2) |
4108. |   50002                   Ankole   1980       4          4    using only (2) |
4145. |   50004                    Kakwa   1980       4          4    using only (2) |
      |---------------------------------------------------------------------------|
4155. |   50005               Karamojong   1980       2          2    using only (2) |
4165. |   50006              Kongjo/Amba   1980       4          2    using only (2) |
4175. |   50007                    Langi   1980       0          0    using only (2) |
4185. |   50008                  Lugbara   1980       4          4    using only (2) |
4195. |   50009                Nyarwanda   1980       4          4    using only (2) |
      |---------------------------------------------------------------------------|
4205. |   50011                     Madi   1980       4          4    using only (2) |
4323. |   50107                  Turkana   1980       1          1    using only (2) |
4333. |   50108                 Rendille   1980       1          1    using only (2) |
4424. |   50113                   Borana   1980       2          2    using only (2) |
4649. |   53002                 Eritreans   1980       4          0    using only (2) |
      |---------------------------------------------------------------------------|
4922. |   55103                    Tonga   1980       0          0    using only (2) |
5018. |   56002    Black Africans (Other)  1980       4          4    using only (2) |
5713. |   64006                     Roma   1980       4          3    using only (2) |
7579. |   81101                    Chams   1980       0          0    using only (2) |
7940. |   85004            East Timorese   1980       4          3    using only (2) |
      +---------------------------------------------------------------------------+
*/

*c - checking groups that had poldis and ecdis values replaced from Asal 1980-1989
data
list numcode group year poldis_check poldis ecdis_check ecdis if _merge==5
```

```
*4 - dropping scode, gprop variables from Asal 1980-1989 data
drop scode gprop poldis_check ecdis_check

*5 - dropping groups that are coded in Asal 1980-1989 data but are not current in AMAR
1940-2003
drop if _merge==2
tab _merge

*6 - renaming poldis, ecdis, and group to match names in Phase V data
rename group AMAR_Group
rename poldis POLDIS
rename ecdis ECDIS

*saving data
save "C:\Users\Aga\Desktop\mardata02.2005.5.dta", replace
clear

*************************************************************************************

*IV. STATA code for cleaning the MAR 2004-2006 data before it is appended to the MAR
1940-2003 data
use "C:\Users\Aga\Desktop\AMAR Reconciliation_2014\AMAR - MAR 2004 2006 and SB
datasets\MAR 2004-2006 Data_original\marupdate_20042006_for reconciliation.dta", clear

*Note: the following variable was added to the 2004-2006 dataset, in order to identify
that a group is a 2004-2006 group.
gen MAR20042006 =.
replace MAR20042006=1 if numcode>0
save "C:\Users\Aga\Desktop\AMAR Reconciliation_2014\AMAR - MAR 2004 2006 and SB
datasets\MAR 2004-2006 Data_original\marupdate_20042006_for reconciliation.dta",
replace
clear

use "C:\Users\Aga\Desktop\AMAR Reconciliation_2014\AMAR - MAR 2004 2006 and SB
datasets\MAR 2004-2006 Data_original\marupdate_20042006_for reconciliation.dta", clear

*cleaning up MAR 2004-2006 group and country names
rename VMAR_Group AMAR_group
rename VMAR_Region AMAR_region
replace country = "Democratic Republic of Congo" if country == "Dem. Rep. of the
Congo"
replace country = "Republic of Congo" if country == "Rep. of the Congo"

sort country AMAR_Group year

save "C:\Users\Aga\Desktop\marupdate_20042006.1.dta"

clear

*************************************************************************************

*V. STATA code for appending 2004-2006 data

use "C:\Users\Aga\Desktop\mardata02.2005.5.dta", clear

append using "C:\Users\Aga\Desktop\marupdate_20042006.1.dta", force

*creating variable that identifies all groups that are in MAR proper (all MAR 1940-
2003 and MAR 2004-2006 groups)
drop n1
sort country AMAR_Group year
gen MAR_proper=1
```

```
sort country AMAR_Group year
by country AMAR_Group: gen n1 = (_n == 1)
list country AMAR_Group if n1==1

*Below we fix the region coding so that it corresponds to the selection bias data.

replace AMAR_Region="Europe" if AMAR_Region=="Western Democracies and Japan"
replace AMAR_Region="Europe" if AMAR_Region=="Post-communist States"
replace AMAR_Region="India" if country=="India"
replace AMAR_Region="Asia" if country=="Japan"
replace AMAR_Region="Asia" if country=="Australia"
replace AMAR_Region="Asia" if country=="New Zealand"
replace AMAR_Region="North America" if country=="United States of America"
replace AMAR_Region="North America" if country=="Canada"

save "C:\Users\Aga\Desktop\mardata02.2005.6.dta", replace

clear

******************************************************************************************

*IV. STATA code for cleaning selection bias data before it is appended to the MAR
1940-2003 data

use "C:\Users\Aga\Desktop\AMAR Reconciliation_2014\AMAR - MAR 2004 2006 and SB
datasets\AMAR Selection Bias Data_original\MARselectionbiasdata.dta", clear

rename countrycode ccode
rename groupname AMAR_Group
rename countryname country
rename amarregion AMAR_Region
rename gpop GPOP
rename cpop CPOP
rename lang LANG
rename gpro GPRO
rename custom CUSTOM
rename belief BELIEF
rename religs1 RELIGS1
rename race RACE
rename groupcon GROUPCON
rename gc2 GC2
rename gc6b GC6B
rename gc7 GC7
rename gc10 GC10
rename gc11 GC11
rename gc119 GC119
rename autlost AUTLOST
rename yearwt YEARWT
rename magn MAGN
rename prstat PRSTAT
rename autonend AUTONEND
rename transyr TRANSYR
rename sepx SEPX
rename sepkin SEPKIN
rename emig EMIG
rename displace DISPLACE
rename poldis POLDIS
rename ecdis ECDIS
rename culpo1 CULPO1
rename culpo2 CULPO2
rename guarrep GUARREP
rename gojpa GOJPA
```

91

```
rename auton2 AUTON2
rename autgain AUTGAIN
rename autpro AUTPRO
rename legisrep LEGISREP
rename execrep EXECREP
rename polgr POLGR
rename ecgr ECGR
rename culgr CULGR
rename kinsup KINSUP
rename kinmatsup KINMATSUP
rename kinpolsup KINPOLSUP
rename kinmilsup KINMILSUP
rename stasup STASUP
rename stamatsup STAMATSUP
rename stapolsup STAPOLSUP
rename stamilsup STAMILSUP
rename nsasup NSASUP
rename nsamatsup NSAMATSUP
rename nsapolsup NSAPOLSUP
rename nsamilsup NSAMILSUP
rename intracon INTRACON
rename factcc1 FACTCC1
rename factcc2 FACTCC2
rename factcc3 FACTCC3
rename factsev1 FACTSEV1
rename factsev2 FACTSEV2
rename factsev3 FACTSEV3
rename intercon INTERCON
rename ccgroup1 CCGROUP1
rename ccgroup2 CCGROUP2
rename ccgroup3 CCGROUP3
rename ccgroupsev1 CCGROUPSEV1
rename ccgroupsev2 CCGROUPSEV2
rename ccgroupsev3 CCGROUPSEV3
rename prot PROT
rename reb REB
rename repgenciv REPGENCIV
rename repnviol REPNVIOL
rename repviol REPVIOL


/*dropping group that was coded in the selection bias data but then disaggregated for
the AMAR sample frame.
This group no longer appears in the AMAR regional lists and is therefore being dropped
for the reconciled data.*/

*dropping Asians in Bahrain
list AMAR_Group country year if numcode==69202
drop if numcode==69202

/*dropping groups that were coded in the selection bias data but that were then found
to not meet the population threshold criteria in a review of the gpro data for the
AMAR sample frame and regional lists.
Therefore, these groups no longer appear as coded in the AMAR data and are being
dropped for the reconciled data.*/

*dropping Bette-Bende in Nigeria
list numcode AMAR_Group country year if AMAR_Group=="Bette-Bende" &
country=="Nigeria"
drop if numcode==47507

*dropping Romansch in Switzerland
list numcode AMAR_Group country year if AMAR_Group=="Romansch" &
country=="Switzerland"
```

```
drop if numcode==22504

*dropping Sorbs in Germany
drop if AMAR_Group=="Sorbs" & country=="Germany"

*dropping Vietnamese in Germany
drop if AMAR_Group=="Vietnamese" & country=="Germany"

*dropping the Polynesians in the Solomon Islands, based on fact that Solomon Islands
doesn't meet the population threshold criteria.
drop if AMAR_Group == "Polynesians" & country =="Solomon Islands"

*generating variable to identify selection bias groups
gen MARselectionbias =.
replace MARselectionbias=1 if numcode>0

*checking count
bysort numcode: gen n1 = (_n == 1)
sum MARselectionbias if n1==1

/*Note: Two groups were coded in the selection bias data but then disaggregated for
the AMAR group lists:
1) the Makonde-Yao in Mozambique - coded as the aggregate Makonde-Yao in the selection
bias data but then disaggregated into the Makonde, Yao, and Makua/Macua in the AMAR
group lists based on updated information; and

2) Indigenous in Bolivia - coded as the aggregate Indigenous in Bolivia in the
selection bias data but then disaggregated into the Aymara and Quechua based in AMAR
group lists based on udpated information.

Although these groups are disaggreated for the AMAR regional lists, we are including
the aggregated groups in the AMAR database in order to not lose this data.*/

save "C:\Users\Aga\Desktop\MARselectionbiasdata.1.dta", replace

clear

************************************************************************************

*VI. STATA code for appending selection bias data

use "C:\Users\Aga\Desktop\mardata02.2005.6.dta", clear

append using "C:\Users\Aga\Desktop\MARselectionbiasdata.1.dta", force

sort country AMAR_Group year

save "C:\Users\Aga\Desktop\mardata02.2005.7.dta", replace

clear

************************************************************************************

*VIII. Recoding POLGR variable based on modifications noted in codebook.

use "C:\Users\Aga\Desktop\mardata02.2005.7.dta", clear

*STATA code for modifying 2004-2006 POLGR variable.
/*NOTE: collapsing 2004-2006 POLGR coding level 2 into coding level 1.
Since the MAR 1940-2003 political grievance variables were only reconciled as POLGR =
-99, 0, 1, 3, and 4, this does not affect these groups.*/
tab POLGR
recode POLGR (2=1)
```

```
tab POLGR

*recoding remaining POLGR levels after collapsing 2 into 1
recode POLGR (3=2)
recode POLGR (4=3)
tab POLGR

save "C:\Users\Aga\Desktop\mardata02.2005.8.dta", replace

clear

*********************************************************************************

*IX. *STATA code for approved recodes

use "C:\Users\Aga\Desktop\mardata02.2005.8.dta", clear

/*1. The following coding corrects the intercommunal and intracommunal conflict coding
for the Acholi in Uganda; specifically, it corrects the INTERCON, CCGROUP1,
CCGROUPSEV1, INTRACON, FACTCC1, and FACTSEV1 variables for 1965-2003.  These coding
changes were noted during a review of the intercommunal conflict variable in 2013-
2014.  During this review, it was determined that the Acholi engaged in INTERCON in
1970 and 1986 only; therefore, INTERCON was coded 1 and CCGROUP1 and CCGROUPSEV1 were
coded accordingly for 1970 and 1986, and these variables were coded 0 for 1965, 1975,
1980, 1985, and 1987-2003.  Further, reports indicate Acholi engaged in intracommunal
conflict (LRA versus Acholi civilians) in 1990-2003.  Therefore, INTRACON was coded 1,
FACTCC1 was coded LRA versus Acholi civilians, and FACTSEV1 was coded accordingly for
1990-2003, and these variables were coded 0 for 1965, 1970, 1975, 1980, and 1985-1989.
The following code reflects these changes. */

*INTERCON
list year AMAR_Group INTERCON CCGROUP1 CCGROUPSEV1 if numcode==50001

*recoding INTERCON = 0 for all years but 1970 and 1986; 1970 and 1986 coded 1
replace INTERCON=0 if numcode== 50001 & year==1965
replace INTERCON=1 if numcode==50001 & year==1970
replace INTERCON =0 if numcode== 50001 & year==1975
replace INTERCON =0 if numcode== 50001 & year==1980
replace INTERCON =0 if numcode== 50001 & year==1985
replace INTERCON=1 if numcode==50001 & year==1986
replace INTERCON =0 if numcode== 50001 & year==1987
replace INTERCON =0 if numcode== 50001 & year==1988
replace INTERCON =0 if numcode== 50001 & year==1989
replace INTERCON =0 if numcode== 50001 & year==1990
replace INTERCON =0 if numcode== 50001 & year==1991
replace INTERCON =0 if numcode== 50001 & year==1992
replace INTERCON =0 if numcode== 50001 & year==1993
replace INTERCON =0 if numcode== 50001 & year==1994
replace INTERCON =0 if numcode== 50001 & year==1995
replace INTERCON =0 if numcode== 50001 & year==1996
replace INTERCON =0 if numcode== 50001 & year==1997
replace INTERCON =0 if numcode== 50001 & year==1998
replace INTERCON =0 if numcode== 50001 & year==1999
replace INTERCON =0 if numcode== 50001 & year==2000
replace INTERCON =0 if numcode== 50001 & year==2001
replace INTERCON =0 if numcode== 50001 & year==2002
replace INTERCON =0 if numcode== 50001 & year==2003

list year INTERCON CCGROUP1 CCGROUPSEV1 if numcode==50001

*recoding CCGROUP1 = blank for all years but 1970 and 1986; for 1970, replacing with
Langi/Lango and for 1986, replacing with Bugandan traders
replace CCGROUP1="" if numcode==50001 & year!= 1970 & year!= 1986
```

```
replace CCGROUP1="Langi/Lango" if numcode== 50001 & year==1970
replace CCGROUP1="Bugandan traders" if numcode ==50001 & year==1986

list year INTERCON CCGROUP1 CCGROUPSEV1 if numcode==50001

*recoding CCGROUPSEV1 = 0 for all years but 1970 and 1986; recoding CCGROUPSEV = 1 for
1970 and CCGROUPSEV1 = 5 for 1986
replace CCGROUPSEV1=0 if numcode==50001 & year==1965
replace CCGROUPSEV1=1 if numcode==50001 & year==1970
replace CCGROUPSEV1=0 if numcode==50001 & year==1975
replace CCGROUPSEV1=0 if numcode==50001 & year==1980
replace CCGROUPSEV1=0 if numcode==50001 & year==1985
replace CCGROUPSEV1=5 if numcode==50001 & year==1986
replace CCGROUPSEV1=0 if numcode==50001 & year==1987
replace CCGROUPSEV1=0 if numcode==50001 & year==1988
replace CCGROUPSEV1=0 if numcode==50001 & year==1989
replace CCGROUPSEV1=0 if numcode==50001 & year==1990
replace CCGROUPSEV1=0 if numcode==50001 & year==1991
replace CCGROUPSEV1=0 if numcode==50001 & year==1992
replace CCGROUPSEV1=0 if numcode==50001 & year==1993
replace CCGROUPSEV1=0 if numcode==50001 & year==1994
replace CCGROUPSEV1=0 if numcode==50001 & year==1995
replace CCGROUPSEV1=0 if numcode==50001 & year==1996
replace CCGROUPSEV1=0 if numcode==50001 & year==1997
replace CCGROUPSEV1=0 if numcode==50001 & year==1998
replace CCGROUPSEV1=0 if numcode==50001 & year==1999
replace CCGROUPSEV1=0 if numcode==50001 & year==2000
replace CCGROUPSEV1=0 if numcode==50001 & year==2001
replace CCGROUPSEV1=0 if numcode==50001 & year==2002
replace CCGROUPSEV1=0 if numcode==50001 & year==2003

list year INTERCON CCGROUP1 CCGROUPSEV1 if numcode==50001

*INTRACON
list year AMAR_Group INTRACON FACTCC1 FACTSEV1 if numcode==50001

*recoding INTRACON = 0 for 1965, 1970, 1975, 1980-1989; recoding INTRACON =1 for 2001-
2003
replace INTRACON=0 if numcode==50001 & year==1965
replace INTRACON=0 if numcode==50001 & year==1970
replace INTRACON=0 if numcode==50001 & year==1975
replace INTRACON=0 if numcode==50001 & year==1980
replace INTRACON=0 if numcode==50001 & year==1985
replace INTRACON=0 if numcode==50001 & year==1986
replace INTRACON=0 if numcode==50001 & year==1987
replace INTRACON=0 if numcode==50001 & year==1988
replace INTRACON=0 if numcode==50001 & year==1989
replace INTRACON=1 if numcode==50001 & year==2001
replace INTRACON=1 if numcode==50001 & year==2002
replace INTRACON=1 if numcode==50001 & year==2003

list year AMAR_Group INTRACON FACTCC1 FACTSEV1 if numcode==50001

*recoding FACTCC1=blank for 1965, 1970, 1975, and 1980-1989; recoding FACTCC1= LRA and
Acholi civilians for 1990-2003
replace FACTCC1 = "" if numcode ==50001 & year==1965
replace FACTCC1 = "" if numcode ==50001 & year==1970
replace FACTCC1 = "" if numcode ==50001 & year==1975
replace FACTCC1 = "" if numcode ==50001 & year==1980
replace FACTCC1 = "" if numcode ==50001 & year==1985
replace FACTCC1 = "" if numcode ==50001 & year==1986
replace FACTCC1 = "" if numcode ==50001 & year==1987
replace FACTCC1 = "" if numcode ==50001 & year==1988
```

```
replace FACTCC1 = "" if numcode ==50001 & year==1989
replace FACTCC1 = "LRA rebels & Acholi civilians" if numcode==50001 & year>1989

list year AMAR_Group INTRACON FACTCC1 FACTSEV1 if numcode==50001

*recoding FACTSEV1 = 0 for 1965, 1970, 1975, and 1980-1989; recoding FACTSEV1=1 for
1990-1994; recoding FACTSEV1=4 for 1995; and recoding FACTSEV1=5 for 1996-2003
replace FACTSEV1=0 if numcode==50001 & year==1965
replace FACTSEV1=0 if numcode==50001 & year==1970
replace FACTSEV1=0 if numcode==50001 & year==1975
replace FACTSEV1=0 if numcode==50001 & year==1980
replace FACTSEV1=0 if numcode==50001 & year==1985
replace FACTSEV1=0 if numcode==50001 & year==1986
replace FACTSEV1=0 if numcode==50001 & year==1987
replace FACTSEV1=0 if numcode==50001 & year==1988
replace FACTSEV1=0 if numcode==50001 & year==1989
replace FACTSEV1=1 if numcode==50001 & year==1990
replace FACTSEV1=1 if numcode==50001 & year==1991
replace FACTSEV1=1 if numcode==50001 & year==1992
replace FACTSEV1=1 if numcode==50001 & year==1993
replace FACTSEV1=1 if numcode==50001 & year==1994
replace FACTSEV1=4 if numcode==50001 & year==1995
replace FACTSEV1=5 if numcode==50001 & year==1996
replace FACTSEV1=5 if numcode==50001 & year==1997
replace FACTSEV1=5 if numcode==50001 & year==1998
replace FACTSEV1=5 if numcode==50001 & year==1999
replace FACTSEV1=5 if numcode==50001 & year==2000
replace FACTSEV1=5 if numcode==50001 & year==2001
replace FACTSEV1=5 if numcode==50001 & year==2002
replace FACTSEV1=5 if numcode==50001 & year==2003

list year AMAR_Group INTRACON FACTCC1 FACTSEV1 if numcode ==50001

**Correct coding error to the Lowland Indigenous in Ecuador for the REB score for 2003
replace REB=1 if numcode==13003 & year==2003

**Correct coding error to the Croats in Serbia for the GROUPCON scores for 2004-2006
replace GROUPCON=2 if numcode== 34507 & year==2004
replace GROUPCON=2 if numcode== 34507 & year==2005
replace GROUPCON=2 if numcode== 34507 & year==2006

**Correct name from the Muslims in Bosnia to the Bosniaks in Bosnia for all years
replace AMAR_Group = "Bosniaks" if numcode==34603

**Correct coding error to Arabs in Israel for the POLDIS scores for 2004-2006
replace POLDIS=3 if numcode== 66601 & year==2004
replace POLDIS=3 if numcode== 66601 & year==2005
replace POLDIS=3 if numcode== 66601 & year==2006

**Correct coding error to the Assamese in India for the ECDIS scores for 2004-2006
replace ECDIS=1 if numcode== 75015 & year==2004
replace ECDIS=1 if numcode== 75015 & year==2005
replace ECDIS=1 if numcode== 75015 & year==2006

**Correct coding error to the Nagas in India for the FACTSEV2 scores for 2004-2006
replace FACTSEV2 =4 if numcode== 75009 & year==2004
replace FACTSEV2 =4 if numcode== 75009 & year==2005
replace FACTSEV2 =4 if numcode== 75009 & year==2006

**Correct coding error to the Maori in New Zealand for the CUSTOM scores for 2004-2006
replace CUSTOM=1 if numcode== 92001 & year==2004
replace CUSTOM=1 if numcode== 92001 & year==2005
replace CUSTOM=1 if numcode== 92001 & year==2006
```

```
**Correct name from the Ogani to the Ogoni in Nigeria for all years
replace AMAR_Group = "Ogoni" if numcode==47504

*Syntax to make changes to data based on recode requests submitted by coders during
the 2004-2006 update and approved by the research director**

**Correct coding error to the Bakongo in Angola for the POLDIS scores for 2002-2003
replace POLDIS=-99 if numcode==54001 & year==2002
replace POLDIS=2 if numcode==54001 & year==2003

**Correct coding error to the Shi'a in Bahrain for ECDIS scores, 1999-2002
replace ECDIS=3 if numcode==69201 & year==1999
replace ECDIS=3 if numcode==69201 & year==2000
replace ECDIS=3 if numcode==69201 & year==2001
replace ECDIS=3 if numcode==69201 & year==2002

**Correct coding error for Afro-Brazilians in Brazil for GROUPCON scores, all years
replace GROUPCON=1 if numcode==14001

**Correct coding error for Mons in Burma in ECDIS scores, 1995-2003
replace ECDIS=4 if numcode==77505 & year==1995
replace ECDIS=4 if numcode==77505 & year==1996
replace ECDIS=4 if numcode==77505 & year==1997
replace ECDIS=4 if numcode==77505 & year==1998
replace ECDIS=4 if numcode==77505 & year==1999
replace ECDIS=4 if numcode==77505 & year==2000
replace ECDIS=4 if numcode==77505 & year==2001
replace ECDIS=4 if numcode==77505 & year==2002
replace ECDIS=4 if numcode==77505 & year==2003

**Correct coding for Mons in Burma for POLDIS scores, 1980-2003
replace POLDIS=4 if numcode==77505 & year>1989
replace POLDIS=3 if numcode==77505 & year<1990

**Correct coding for CULPO2 for Corsicans in France.
replace CULPO2=0 if numcode==22004 & year>1998

**Correct coding for Basques in France for POLDIS scores, 1982-2003
replace POLDIS=0 if numcode==22002 & year>1981

** Correct coding for Muslims in France for POLDIS scores, 2001-2003
replace POLDIS=3 if numcode==22006 & year==2001
replace POLDIS=3 if numcode==22006 & year==2002
replace POLDIS=3 if numcode==22006 & year==2003

**Correct coding for Mende in Sierra Leone for POLDIS score, 2003
replace POLDIS=0 if numcode==45103 & year==2003

** Correct coding for Greeks in Albania for CULPO2 score, 2004-2006
replace CULPO2=0 if numcode==33901 & year>2003

** Correct coding for Cabindas in Angola for BELIEF, all years
replace BELIEF=0 if numcode==54003 & year>1979

**Correct coding for Cabindans in Angola for CULPO1 for 2002, 2003
replace CULPO1=2 if numcode==54003 & year==2002
replace CULPO1=2 if numcode==54003 & year==2003

**Correct coding for indigenous in Argentina for PROT for 2002
replace PROT=2 if numcode==16002 & year==2002

**Correct coding for Shi'is in Bahrain for GOJPA for 2002-2003
```

```
replace GOJPA=2 if numcode==69201 & year==2002
replace GOJPA=2 if numcode==69201 & year==2003


**Correct coding for Shan in Burma for CULPO2 for 1980-on
replace CULPO2=2 if numcode==77507 & year>1979


**Correct coding for Zomis (Chins) in Burma for CULPO1 for 1990-2006
replace CULPO1=2 if numcode==77502 & year>1989


**Backcode LEGISREP for Kirdi in Cameroon for 1992-2003
replace LEGISREP=1 if numcode==47101 & year==1992
replace LEGISREP=1 if numcode==47101 & year==1993
replace LEGISREP=1 if numcode==47101 & year==1994
replace LEGISREP=1 if numcode==47101 & year==1995
replace LEGISREP=1 if numcode==47101 & year==1996
replace LEGISREP=1 if numcode==47101 & year==1997
replace LEGISREP=1 if numcode==47101 & year==1998
replace LEGISREP=1 if numcode==47101 & year==1999
replace LEGISREP=1 if numcode==47101 & year==2000
replace LEGISREP=1 if numcode==47101 & year==2001
replace LEGISREP=1 if numcode==47101 & year==2002
replace LEGISREP=0 if numcode==47101 & year==2003


**Correct coding for ECDIS for Westerners in Cameroon for 1990-2006
replace ECDIS=3 if numcode==47102 & year>1989


**Correct coding for CULPO2 for Turkmen in China
replace CULPO2=1 if numcode==71003 & year>1979


 **Correct coding for Antillean Blacks in Costa Rica for GOJPA for 1995-2003
replace GOJPA=1 if numcode==9401 & year>1994


**Correct coding for Serbs in Croatia for CULPO2 for 2001-2003
replace CULPO2=1 if numcode==34401 & year==2001
replace CULPO2=1 if numcode==34401 & year==2002
replace CULPO2=1 if numcode==34401 & year==2003


**Correct coding for Slovaks in Czech Republic for GROUPCON, all years
replace GROUPCON=1 if numcode==31603 & year>1993


** Correct coding for Hutus in DRC for INTERCON and related variables, for 1996.
replace INTERCON=1 if numcode==49009 & year==1996
replace CCGROUP1="Rwandan Army" if numcode==49009 & year==1996
replace CCGROUPSEV1=5 if numcode==49009 & year==1996


** Correct coding for Luba in DRC for GOJPA for 2002-2004
replace GOJPA=2 if numcode == 49003 & year==2002
replace GOJPA=2 if numcode == 49003 & year==2003
replace GOJPA=2 if numcode == 49003 & year==2004


**Correct coding for Muslims in France for INTERCON and associated variables for 2002.
replace INTERCON=1 if numcode==22006 & year==2002
replace CCGROUP1="French" if numcode==22006 & year==2002
replace CCGROUPSEV1=3 if numcode==22006 & year ==2002


**Correct POLDIS coding for Russians in Georgia for 2001-2003
replace POLDIS=0 if numcode==37204 & year==2001
replace POLDIS=0 if numcode==37204 & year==2002
replace POLDIS=0 if numcode==37204 & year==2003


**Correct coding for Fulani in Guinea for CULPO1 for 1980-2006
replace CULPO1=0 if numcode==43801
```

```
**Correct coding for Fulani in Guinea for CUSTOM for 1980-2006
replace CUSTOM=0 if numcode==43801

**Correct coding for Fulani in Guinea for GOJPA for 1998-2003
replace GOJPA=2 if numcode==43801 & year==1998
replace GOJPA=2 if numcode==43801 & year==1999
replace GOJPA=2 if numcode==43801 & year==2000
replace GOJPA=2 if numcode==43801 & year==2001
replace GOJPA=2 if numcode==43801 & year==2002
replace GOJPA=2 if numcode==43801 & year==2003

**Correct coding for Malinke in Guinea for GROUPCON, all years
replace GROUPCON=3 if numcode==43802

**Correct coding for Kashmiris in India for CULPO1 for all years
replace CULPO1=1 if numcode==75007 & year<1996
replace CULPO1=2 if numcode==75007 & year>1995

**Correct coding for Kashmiris in India for ECDIS for all years
replace ECDIS=3 if numcode==75007 & year<1990
replace ECDIS=1 if numcode==75007 & year==1990
replace ECDIS=1 if numcode==75007 & year==1991
replace ECDIS=1 if numcode==75007 & year==1992
replace ECDIS=1 if numcode==75007 & year==1993
replace ECDIS=1 if numcode==75007 & year==1994
replace ECDIS=1 if numcode==75007 & year==1995
replace ECDIS=1 if numcode==75007 & year==1996
replace ECDIS=1 if numcode==75007 & year==1997
replace ECDIS=1 if numcode==75007 & year==1998
replace ECDIS=3 if numcode==75007 & year>1998

**Correct coding for Kashmiris in India for POLDIS for all years
replace POLDIS=3 if numcode==75007 & year<1990
replace POLDIS=4 if numcode==75007 & year>1989

**Correct coding for Nagas in India for ECDIS for 2004-2006
replace ECDIS=1 if numcode == 75009 & year==2004
replace ECDIS=1 if numcode == 75009 & year==2005
replace ECDIS=1 if numcode == 75009 & year==2006

**Correct coding for Nagas in India for POLDIS for 2004-2006
replace POLDIS=1 if numcode == 75009 & year==2004
replace POLDIS =1 if numcode == 75009 & year==2005
replace POLDIS =1 if numcode == 75009 & year==2006

**Backcode CULPO2 for 2000s for Arabs in Iran
replace CULPO2=2 if numcode==63009 & year>1999

**Correct coding for Arabs in Iran for ECDIS, all years
replace ECDIS=3 if numcode==63009

**Correct coding for CULPO2 for Baluchis in Iran for all years
replace CULPO2=1 if  numcode==63005

**Correct coding for Baluchis in Iran for all year for ECDIS
replace ECDIS=3 if numcode==63005

**Correct coding for Baluchis in Iran for GOJPA for 2003
replace GOJPA=4 if numcode==63005 & year==2003

**Correct coding for Baluchis in Iran for  SEPX for 2003
replace SEPX=3 if numcode==63005 & year==2003
```

```
**Correct coding for Koreans in Japan for CULPO2, all years
replace CULPO2=1 if numcode==74003

**Correct coding for Luo in Kenya for GROUPCON
replace GROUPCON=3 if numcode==50104 & year<1999
replace GROUPCON=-99 if numcode==50104 & year==1999
replace GROUPCON=-99 if numcode==50104 & year==2000
replace GROUPCON=-99 if numcode==50104 & year==2001
replace GROUPCON=-99 if numcode==50104 & year==2002
replace GROUPCON=-99 if numcode==50104 & year==2003
replace GROUPCON=2 if numcode==50104 & year>2003

**Correct coding for Druze in Lebanon for GOJPA for 1993-2003
replace GOJPA=2 if numcode==66001 & year>1992

**Correct coding for Druze in Lebanon for POLDIS for 2001-2003
replace POLDIS=0 if numcode==66001 & year==2001
replace POLDIS=0 if numcode==66001 & year==2002
replace POLDIS=0 if numcode==66001 & year==2003

**Correct coding for Poles in Lithuania for ECDIS for 1992-on
replace ECDIS=2 if numcode==36801 & year>1991

**Correct coding for Serbs in Macedonia for CULPO1 for 1991-on
replace CULPO1=2 if numcode==34302 & year>1990

**Correct coding for Chinese in Malaysia for CULPO2 for 1980-on
replace CULPO2=1 if numcode==82001 & year>1979

**Correct coding for East Indians in Malaysia for POLDIS for 2000-2003
replace POLDIS=2 if numcode==82003 & year>1999

**Correct coding for Lhotshampas in Bhutan for all years
replace CULPO1=2 if numcode==76001

**Correct coding for Mohajirs in Pakistan for POLDIS, all years
replace POLDIS=3 if numcode==77007 & year<1988
replace POLDIS=0 if numcode==77007 & year==1988
replace POLDIS=0 if numcode==77007 & year==1989
replace POLDIS=0 if numcode==77007 & year==1990
replace POLDIS=0 if numcode==77007 & year==1991
replace POLDIS=3 if numcode==77007 & year==1992
replace POLDIS=3 if numcode==77007 & year==1993
replace POLDIS=3 if numcode==77007 & year==1994
replace POLDIS=3 if numcode==77007 & year==1995
replace POLDIS=3 if numcode==77007 & year==1996
replace POLDIS=3 if numcode==77007 & year==1997
replace POLDIS=3 if numcode==77007 & year==1998
replace POLDIS=3 if numcode==77007 & year==1999
replace POLDIS=3 if numcode==77007 & year==2000
replace POLDIS=3 if numcode==77007 & year==2001
replace POLDIS=0 if numcode==77007 & year>2001

**Correct coding for indigenous in Panama for GROUPCON for all years
replace GROUPCON=2 if numcode==9502

**Correct coding for Lari in Rep of Congo for ECDIS< 2001-2003
replace ECDIS=2 if numcode==48401 & year==2001
replace ECDIS=2 if numcode==48401 & year==2002
replace ECDIS=2 if numcode==48401 & year==2003

**Correct coding for M'Boshi in Rep of Congo for CUSTOM, all years
replace CUSTOM=1 if numcode==48402
```

```
**Correct coding for Roma in Slovakia for ECDIS, 2003
replace ECDIS=1 if numcode==31702 & year==2003

**Correct coding for Magyars in Romania for CULPO1, 1996-2003
replace CULPO1=1 if numcode==36002 & year>1995

**Correct coding for Magyars in Romania for CULPO2, 1999-2003
replace CULPO2=1 if numcode==36002 & year>1998

**Correct coding for Magyars in Romania for GROUPCON, all years
replace GROUPCON=3 if numcode==36002

**Backcode LEGISREP for Mende in Sierra Leone, 2002, 2003
replace LEGISREP=1 if numcode==45103 & year==2002
replace LEGISREP=1 if numcode==45103 & year==2003

**Correct coding for Mende in Sierra Leone for POLDIS, 2002, 2003
replace POLDIS=0 if numcode==45103 & year==2002
replace POLDIS=0 if numcode==45103 & year==2003

**Correct coding for Temne in Sierra Leone for POLDIS, 2002-2003
replace POLDIS=0 if numcode==45104 & year==2002
replace POLDIS=0 if numcode==45104 & year==2003

**Correct coding for Roma in Slovakia for ECDIS, 2003
replace ECDIS=1 if numcode ==31702 & year==2003

**Correct coding for Somalis in Ethiopia for POLDIS, 1992-2006
replace POLDIS=2 if numcode==53005 & year>1991

**Correct coding for Afro-Caribbeans in UK for POLDIS, 1980-2006
replace POLDIS=1 if numcode==20005 & year<1990
replace POLDIS=-99 if numcode==20005 & year==1991
replace POLDIS=-99 if numcode==20005 & year==1992
replace POLDIS=3 if numcode==20005 & year>1992

**Correct coding for Afro-Caribbeans in UK for CULPO1, 1980-2006
replace CULPO1=0 if numcode==20005

**Correct coding for Crimean Tatars in Ukraine for POLDIS, 1997-2006
replace POLDIS=1 if numcode==36904 & year >1996

**Correct coding for Blacks in Venezuela for CULPO1, 1980-2006
replace CULPO1=0 if numcode==10101

**Correct coding for Blacks in Venezuela for CULPO2, 1980-2006
replace CULPO2=0 if numcode==10101

**Correct coding for Indigenous in Venezuela for GROUPCON, all years
replace GROUPCON=2 if numcode==10102

**Correct coding for Montagnards in Vietnam for POLDIS, 2004-2006
replace POLDIS=2 if numcode==81604 & year>2003

**Correct coding for Indigenous in Argentina for POLDIS, 1980-2006
replace POLDIS=4 if numcode==16002 & year<1983
replace POLDIS=3 if numcode==16002 & year>1982

**Backcode Indigenous in Argentina for REPGENCIV, 2002
replace REPGENCIV=4 if numcode==16002 & year==2002

**Backcode Indigenous in Argentina for STAMATSUP, 1993-2003
```

```
replace STAMATSUP=1 if numcode==16002 & year==1993
replace STAMATSUP=1 if numcode==16002 & year==1994
replace STAMATSUP=1 if numcode==16002 & year==1995
replace STAMATSUP=1 if numcode==16002 & year==1996
replace STAMATSUP=1 if numcode==16002 & year==1997
replace STAMATSUP=1 if numcode==16002 & year==1998
replace STAMATSUP=1 if numcode==16002 & year==1999
replace STAMATSUP=1 if numcode==16002 & year==2000
replace STAMATSUP=1 if numcode==16002 & year==2001
replace STAMATSUP=1 if numcode==16002 & year==2002
replace STAMATSUP=1 if numcode==16002 & year==2003
replace STASUP=1 if numcode==16002 & year==1993
replace STASUP=1 if numcode==16002 & year==1994
replace STASUP=1 if numcode==16002 & year==1995
replace STASUP=1 if numcode==16002 & year==1996
replace STASUP=1 if numcode==16002 & year==1997
replace STASUP=1 if numcode==16002 & year==1998
replace STASUP=1 if numcode==16002 & year==1999
replace STASUP=1 if numcode==16002 & year==2000
replace STASUP=1 if numcode==16002 & year==2001
replace STASUP=1 if numcode==16002 & year==2002
replace STASUP=1 if numcode==16002 & year==2003


**Correct coding for Russians in Belarus for GOJPA, 1996-2006
replace GOJPA=3 if numcode==37001 & year>1995


**Correct coding for Croats in BiH for LANG, all years
replace LANG=1 if numcode==34602


**Correct coding for Serbs in BiH for LANG, all years
replace LANG=1 if numcode==34601


**Correct coding for Luba in DRC for GOJPA, 2002-2003
replace GOJPA=2 if numcode==49003 & year==2002
replace GOJPA=2 if numcode==49003 & year==2003


**Backcode AUTON2 for Afro-Ecuadorians in Ecuador, all years
replace AUTON2=0 if numcode==13001


**Backcode AUTPRO for Afro-Ecuadorians in Ecuador, all years
replace AUTPRO=0 if numcode==13001


**Correct coding for CULGR for Afro-Ecuadorians in Ecuador, 1997-2003
replace CULGR=1 if numcode==13001 & year==1997
replace CULGR=1 if numcode==13001 & year==1998
replace CULGR=1 if numcode==13001 & year==1999
replace CULGR=1 if numcode==13001 & year==2000
replace CULGR=1 if numcode==13001 & year==2001
replace CULGR=1 if numcode==13001 & year==2002
replace CULGR=1 if numcode==13001 & year==2003


**Correct coding for Afro-Ecuadorians in Ecuador for CUSTOM, all years
replace CUSTOM=1 if numcode==13001


**Correct coding for Afro-Ecuadorians in Ecuador for ECGR, 1998-2002
replace ECGR=2 if numcode==13001 & year==1998
replace ECGR=2 if numcode==13001 & year==1999
replace ECGR=2 if numcode==13001 & year==2000
replace ECGR=2 if numcode==13001 & year==2001
replace ECGR=2 if numcode==13001 & year==2002


**Backcode EXECREP for Afro-Ecuadorians in Ecuador, 2001-2003
replace EXECREP=0 if numcode==13001 & year==2001
```

```
replace EXECREP=0 if numcode==13001 & year==2002
replace EXECREP=0 if numcode==13001 & year==2003


**Backcode GUARREP for Afro-Ecuadorians in Ecuador, all years
replace GUARREP=0 if numcode==13001


**Backcode for KINSUP and related variables for Afro-Ecuadorians in Ecuador, 2001-2003
replace KINSUP=0 if numcode==13001 & year==2001
replace KINSUP=0 if numcode==13001 & year==2002
replace KINSUP=0 if numcode==13001 & year==2003
replace KINMATSUP=0 if numcode==13001 & year==2001
replace KINMATSUP=0 if numcode==13001 & year==2002
replace KINMATSUP=0 if numcode==13001 & year==2003
replace KINPOLSUP=0 if numcode==13001 & year==2001
replace KINPOLSUP=0 if numcode==13001 & year==2002
replace KINPOLSUP=0 if numcode==13001 & year==2003
replace KINMILSUP=0 if numcode==13001 & year==2001
replace KINMILSUP=0 if numcode==13001 & year==2002
replace KINMILSUP=0 if numcode==13001 & year==2003


**Backcode LANG for Afro-Ecuadorians in Ecuador, all years
replace LANG=0 if numcode==13001


**Backcode LEGISREP for Afro-Ecuadorians in Ecuador, 1999-2003
replace LEGISREP=1 if numcode==13001 & year==1999
replace LEGISREP=1 if numcode==13001 & year==2000
replace LEGISREP=1 if numcode==13001 & year==2001
replace LEGISREP=1 if numcode==13001 & year==2002
replace LEGISREP=1 if numcode==13001 & year==2003


**Backcode NSAMILSUP for Afro-Ecuadorians in Ecuador, 2001-2003
replace NSAMILSUP=0 if numcode==13001 & year==2001
replace NSAMILSUP=0 if numcode==13001 & year==2002
replace NSAMILSUP=0 if numcode==13001 & year==2003


**Backcode NSASUP, NSAPOLSUP and NSAMATSUP for Afro-Ecuadorians in Ecuador, 1995-2003
replace NSASUP=1 if numcode==13001 & year==1995
replace NSASUP=1 if numcode==13001 & year==1996
replace NSASUP=1 if numcode==13001 & year==1997
replace NSASUP=1 if numcode==13001 & year==1998
replace NSASUP=1 if numcode==13001 & year==1999
replace NSASUP=1 if numcode==13001 & year==2000
replace NSASUP=1 if numcode==13001 & year==2001
replace NSASUP=1 if numcode==13001 & year==2002
replace NSASUP=1 if numcode==13001 & year==2003
replace NSAPOLSUP=1 if numcode==13001 & year==1995
replace NSAPOLSUP=1 if numcode==13001 & year==1996
replace NSAPOLSUP=1 if numcode==13001 & year==1997
replace NSAPOLSUP=1 if numcode==13001 & year==1998
replace NSAPOLSUP=1 if numcode==13001 & year==1999
replace NSAPOLSUP=1 if numcode==13001 & year==2000
replace NSAPOLSUP=1 if numcode==13001 & year==2001
replace NSAPOLSUP=1 if numcode==13001 & year==2002
replace NSAPOLSUP=1 if numcode==13001 & year==2003
replace NSAMATSUP=1 if numcode==13001 & year==1995
replace NSAMATSUP=1 if numcode==13001 & year==1996
replace NSAMATSUP=1 if numcode==13001 & year==1997
replace NSAMATSUP=1 if numcode==13001 & year==1998
replace NSAMATSUP=1 if numcode==13001 & year==1999
replace NSAMATSUP=1 if numcode==13001 & year==2000
replace NSAMATSUP=1 if numcode==13001 & year==2001
replace NSAMATSUP=1 if numcode==13001 & year==2002
replace NSAMATSUP=1 if numcode==13001 & year==2003
```

```
**Backcode POLGR for Afro-Ecuadorians in Ecuador, 2000-2003
**Note: This was originally recoded as POLGR=3 (group calling for autonomy) during the
initial backcoding of this variable.  However, since the POLGR levels were collapsed
and recoded for the AMAR Phase I data, and POLGR = 3 is now POLGR =2, this is now
being coded as POLGR=2.
replace POLGR=2 if numcode==13001 & year==2000
replace POLGR=2 if numcode==13001 & year==2001
replace POLGR=2 if numcode==13001 & year==2002
replace POLGR=2 if numcode==13001 & year==2003

**Backcode RELIGS1 for Afro-Ecuadorians in Ecuador, all years
replace RELIGS1=1 if numcode==13001

**Backcode REPGENCIV for Afro-Ecuadorians in Ecuador, 2001-2003
replace REPGENCIV=1 if numcode==13001 & year==2001
replace REPGENCIV=4 if numcode==13001 & year==2002
replace REPGENCIV=1 if numcode==13001 & year==2003

**Backcode REPNVIOL for Afro-Ecuadorians in Ecuador, 2001-2003
replace REPNVIOL=0 if numcode==13001 & year==2001
replace REPNVIOL=0 if numcode==13001 & year==2002
replace REPNVIOL=0 if numcode==13001 & year==2003

**Backcode REPVIOL for Afro-Ecuadorians in Ecuador, 2001-2003
replace REPVIOL=0 if numcode==13001 & year==2001
replace REPVIOL=0 if numcode==13001 & year==2002
replace REPVIOL=0 if numcode==13001 & year==2003

**Backcode SEPX for Afro-Ecuadorian in Ecuador, 2000-2003
replace SEPX=3 if numcode==13001 & year==2000
replace SEPX =3 if numcode==13001 & year==2001
replace SEPX =3 if numcode==13001 & year==2002
replace SEPX =3 if numcode==13001 & year==2003

**Backcode STAPOLSUP and STAMILSUP for Afro-Ecuadorians in Ecuador, 2001-2003
replace STAPOLSUP=0 if numcode==13001 & year==2001
replace STAPOLSUP=0 if numcode==13001 & year==2002
replace STAPOLSUP=0 if numcode==13001 & year==2003
replace STAMILSUP=0 if numcode==13001 & year==2001
replace STAMILSUP =0 if numcode==13001 & year==2002
replace STAMILSUP =0 if numcode==13001 & year==2003

**Backcode STAMATSUP and STASUP for Afro-Ecuadorians in Ecuador, 1998-2003
replace STASUP=1 if numcode==13001 & year==1998
replace STASUP =1 if numcode==13001 & year==1999
replace STASUP =1 if numcode==13001 & year==2000
replace STASUP =1 if numcode==13001 & year==2001
replace STASUP =1 if numcode==13001 & year==2002
replace STASUP =1 if numcode==13001 & year==2003
replace STAMATSUP=1 if numcode==13001 & year==1998
replace STAMATSUP =1 if numcode==13001 & year==1999
replace STAMATSUP =1 if numcode==13001 & year==2000
replace STAMATSUP =1 if numcode==13001 & year==2001
replace STAMATSUP =1 if numcode==13001 & year==2002
replace STAMATSUP =1 if numcode==13001 & year==2003

**Correct coding for GOJPA for Abkhazians in Georgia, 1992-2003
replace GOJPA=4 if numcode==37201& year>1991

**Correct coding for GOJPA for Ashanti in Ghana, 1992-on
replace GOJPA=2 if numcode==45201 & year>1991
```

```
**Correct coding for GOJPA for Malinke in Guinea, all years
replace GOJPA=1 if numcode==43802 & year<1990
replace GOJPA=2 if numcode==43802 & year>1989

**Correct coding for GOJPA for Chinese in Indonesia, 1998-2006
replace GOJPA=2 if numcode==85003 & year>1997

**Correct coding for GOJPA for Arabs in Iran, 1980-2004
replace GOJPA=3 if numcode==63009 & year==1980
replace GOJPA=-99 if numcode==63009 & year==1981
replace GOJPA=-99 if numcode==63009 & year==1982
replace GOJPA=-99 if numcode==63009 & year==1983
replace GOJPA=-99 if numcode==63009 & year==1984
replace GOJPA=-99 if numcode==63009 & year==1985
replace GOJPA=-99 if numcode==63009 & year==1986
replace GOJPA=-99 if numcode==63009 & year==1987
replace GOJPA=-99 if numcode==63009 & year==1988
replace GOJPA=-99 if numcode==63009 & year==1989
replace GOJPA=-99 if numcode==63009 & year==1990
replace GOJPA=-99 if numcode==63009 & year==1991
replace GOJPA=-99 if numcode==63009 & year==1992
replace GOJPA=-99 if numcode==63009 & year==1993
replace GOJPA=-99 if numcode==63009 & year==1994
replace GOJPA=-99 if numcode==63009 & year==1995
replace GOJPA=-99 if numcode==63009 & year==1996
replace GOJPA=-99 if numcode==63009 & year==1997
replace GOJPA=-99 if numcode==63009 & year==1998
replace GOJPA=-99 if numcode==63009 & year==1999
replace GOJPA=1 if numcode==63009 & year==2000
replace GOJPA=1 if numcode==63009 & year==2001
replace GOJPA=1 if numcode==63009 & year==2002
replace GOJPA=1 if numcode==63009 & year==2003
replace GOJPA=1 if numcode==63009 & year==2004

**Correct coding for EMIG for Shi'is in Iraq, 2003
replace EMIG=1 if numcode==64506 & year==2003

**Backcode DISPLACE for Shi'is in Iraq, 2003
replace DISPLACE=1 if numcode==64505 & year==2003

**Correct coding for CUSTOM for Sunnis in Iraq, all years
replace CUSTOM=1 if numcode==64507

**Correct coding for GOJPA for Roma in Italy, all years
replace GOJPA=1 if numcode==32504

**Correct coding for LANG for Sards in Italy, 1980-1984
replace LANG=1 if numcode==32503 & year==1980
replace LANG=1 if numcode==32503 & year==1981
replace LANG=1 if numcode==32503 & year==1982
replace LANG=1 if numcode==32503 & year==1983
replace LANG=1 if numcode==32503 & year==1984

**Correct coding for BELIEF for Luo in Kenya, all years
replace BELIEF=0 if numcode==50104

**Correct coding for GOJPA for Maronite Christians in Lebanon, 2004-2006
replace GOJPA=3 if numcode==66002 & year>2003

**Correct coding for EMIG for Albanians in Macedonia, 2001-2003
replace EMIG=3 if numcode==34301 & year==2001
replace EMIG=1 if numcode==34301 & year==2002
replace EMIG=0 if numcode==34301 & year==2003
```

```
**Correct coding for GOJPA for Slavs in Moldova, 1992-2003
replace GOJPA=3 if numcode==35902 & year>1992

**Correct coding for CUSTOM for Ahmadis in Pakistan, all years
replace CUSTOM=0 if numcode==77001

**Correct coding for GOJPA for Mohajirs in Pakistan, all years
replace GOJPA=-99 if numcode==77007 & year<1984
replace GOJPA=3 if numcode==77007 & year>1983

**Correct coding for GOJPA for Sindhis in Pakistan, all years
replace GOJPA=-99 if numcode==77006 & year<1988
replace GOJPA=3 if numcode==77006 & year==1988
replace GOJPA=3 if numcode==77006 & year==1989
replace GOJPA=3 if numcode==77006 & year==1990
replace GOJPA=3 if numcode==77006 & year==1991
replace GOJPA=3 if numcode==77006 & year==1992
replace GOJPA=3 if numcode==77006 & year==1993
replace GOJPA=3 if numcode==77006 & year==1994
replace GOJPA=3 if numcode==77006 & year==1995
replace GOJPA=3 if numcode==77006 & year==1996
replace GOJPA=3 if numcode==77006 & year==1997
replace GOJPA=3 if numcode==77006 & year==1998
replace GOJPA=2 if numcode==77006 & year==1999
replace GOJPA=2 if numcode==77006 & year==2000
replace GOJPA=2 if numcode==77006 & year==2001
replace GOJPA=2 if numcode==77006 & year==2002
replace GOJPA=2 if numcode==77006 & year==2003
replace GOJPA=3 if numcode==77006 & year==2004

**Backcode GOJPA for Indigenous in Panama, 1980-on
replace GOJPA=2 if numcode==9502

**Correct coding for GOJPA for Highland Indigenous in Peru for 1980-2003
replace GOJPA=3 if numcode==13502

**Correct coding for CUSTOM for M'boshi in Rep. of Congo, all years
replace CUSTOM=1 if numcode==48402

**Backcode EMIG for Magyars in Romania for 2001-2003
replace EMIG=1 if numcode==36002 & year==2001
replace EMIG=1 if numcode==36002 & year==2002
replace EMIG=1 if numcode==36002 & year==2003

**Correct coding for LANG for Magyars in Romania
replace LANG=1 if numcode==36002

**Backcode STASUP and STAMATSUP for Magyars in Romania, 2001-2003
replace STASUP=1 if numcode==36002 & year==2001
replace STASUP=1 if numcode==36002 & year==2002
replace STASUP=1 if numcode==36002 & year==2003
replace STAMATSUP=1 if numcode==36002 & year==2001
replace STAMATSUP=1 if numcode==36002 & year==2002
replace STAMATSUP=1 if numcode==36002 & year==2003

**Correct coding for GOJPA for Temne in Sierra Leone, 2002-2003
replace GOJPA=2 if numcode==45104 & year==2002
replace GOJPA=2 if numcode==45104 & year==2003

**Backcode EMIG for Europeans in S. Africa, 2001-2003
replace EMIG=1 if numcode==56004& year==2001
replace EMIG=1 if numcode==56004& year==2002
```

```
replace EMIG=1 if numcode==56004& year==2003


**Correct coding for GOJPA for Sri Lankan Tamils in Sri Lanka, 1995-2003
replace GOJPA=3 if numcode==78002 & year<1984
replace GOJPA=4 if numcode==78002 & year>1983


**Correct coding for GOJPA for Chinese in Thailand, 1998-2003
replace GOJPA=1 if numcode==80001 & year==1998
replace GOJPA=1 if numcode==80001 & year==1999
replace GOJPA=1 if numcode==80001 & year==2000
replace GOJPA=1 if numcode==80001 & year==2001
replace GOJPA=1 if numcode==80001 & year==2002
replace GOJPA=1 if numcode==80001 & year==2003


**Correct coding for GOJPA for Acholi in Uganda, 1986-2006
replace GOJPA=3 if numcode==50001 & year>1985


**Correct coding for CUSTOM for Afro-Caribbeans in the UK, 2004-2006
replace CUSTOM=1 if numcode==20005


**Backcode LANG for African Americans in the U.S., all years
replace LANG=0 if numcode==201


**Correct GOJPA coding for Blacks in Venezuela, 2000-2003
replace GOJPA=2 if numcode==10101 & year==2000
replace GOJPA=2 if numcode==10101 & year==2001
replace GOJPA=2 if numcode==10101 & year==2002
replace GOJPA=2 if numcode==10101 & year==2003


**Correct LANG coding for Indigenous in Venezuela, all years
replace LANG=1 if numcode==10102


**Correct EMIG coding for Montagnards in Vietnam, 2001-2003
replace EMIG=1 if numcode==81604 & year==2001
replace EMIG=1 if numcode==81604 & year==2002
replace EMIG=1 if numcode==81604 & year==2003


*X.  Changing country name of Yugoslavia to reflect de jure name changes for all
groups except Serbs (country name was already changed in the selection bias data)
sort numcode year
list AMAR_Group country year numcode if ccode==345
replace country="Socialist Federal Republic of Yugoslavia" if country=="Yugoslavia" &
year<1993
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==1993
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==1994
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==1995
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==1996
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==1997
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==1998
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==1999
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==2000
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==2001
replace country="Federal Republic of Yugoslavia" if country=="Yugoslavia" & year==2002
replace country="Serbia and Montenegro" if country=="Yugoslavia" & year>2002
list AMAR_Group country year numcode if ccode==345


*XI.  Changing name of Kosovo Albanians to Albanians to reflect name in sample frame
and AMAR Phase I data
list AMAR_Group country year numcode if AMAR_Group=="Kosovo Albanians"
replace AMAR_Group="Albanians" if numcode==34501
list AMAR_Group country year numcode if numcode==34501


save "C:\Users\Aga\Desktop\AMAR.PhaseVI.19402006.FINAL.dta", replace
```

```
clear

****************************************************************************
*XII. STATA code for adding 2001/2007 gpro data from the AMAR sample frame

insheet using "C:\Users\Aga\Dropbox UMD\Dropbox\A-MAR revisions\AMAR
Reconciliation\Sample_Frame_8-18_FINAL updated June 10 2016 FINAL PUBLIC.csv"

*prepping database for merger
drop  v21- v47
rename amargroupname AMAR_Group
rename region AMAR_Region
rename _7_gpro gpro_AMAR
rename selection_bias_groupinamar selectionbias_in_amar
sort numcode

replace country = "Republic of Congo" if country== "Rep. Of Congo"
replace country = "Democratic Republic of Congo" if country == "Democratic Republic Of
The Congo"
replace country = "Mozambique" if country=="MOZAMBIQUE"
replace country = "United Kingdom" if country == "Uk"
replace country = "United States of America" if country == "Usa"

replace AMAR_Region ="Middle East and North Africa" if AMAR_Region=="Middle_East"
replace AMAR_Region ="Sub-Saharan Africa" if AMAR_Region=="SSAfrica"
replace AMAR_Region ="Latin America and the Caribbean" if AMAR_Region=="Latin_America"
replace AMAR_Region="North America" if AMAR_Region=="North_America"

*saving for merger
sort numcode
save "C:\Users\Aga\Desktop\AMARsampleframeJune102016.dta", replace
clear

****************************************************************************

*now appending the sample frame gpros to the AMAR data

use "C:\Users\Aga\Dropbox UMD\Dropbox\A-MAR revisions\AMAR Reconciliation\AMAR_all
files and data for replicating data FINAL\AMAR.PhaseVI.19402006.FINAL.dta", clear

sort numcode year

drop _merge

merge m:1 numcode using "C:\Users\Aga\Dropbox UMD\Dropbox\A-MAR revisions\AMAR
Reconciliation\AMAR_sample frame FINAL\AMARsampleframeJune102016.dta"

*checking merge
drop n1
bysort ccode numcode: gen n1 = (_n == 1)
sum  MAR_proper MARselectionbias selectionbias_in_amar amar_group mar_in_amar
amar_group if n1==1

/*
    Variable |      Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
  MAR_proper |      288           1           0          1          1
MARselecti~s |       74           1           0          1          1
selectionb~r |       74           1           0          1          1
  amar_group |     1202           1           0          1          1
 mar_in_amar |      291           1           0          1          1
-------------+--------------------------------------------------------
```

108

```
  amar_group |      1202            1           0          1           1
.
```

MAR_proper counts groups in MAR, mar_in_amar counts groups as they appear in the AMAR
data as informed by the Sample Frame.
These numbers are different because groups were split, merged, and/or dropped between
MAR and AMAR in the following manner:

```
MAR group    Indigenous Highland Peoples Bolivia
             - disaggregated to Aymara and Quechua in AMAR (+1)
MAR group    Jurassians   Switzerland
             - aggregated to French Speakers in AMAR (-1)
MAR group    Foreign Workers Switzerland
             - disaggregated into French-Speakers, Portuguese, Spanish, Yugoslavs in
             AMAR (+2)
MAR group    Turks  West Germany and East Germany and Germany
             - aggregated into Turks Germany in AMAR (-2)
MAR group    Croats Socialist Federal Republic of Yugoslavia
             - dropped because they don't meet population criteria in AMAR (-1)
MAR group    Zanzibaris   Tanzania
             - disaggregated to Zanzibari Arabs and Zanzibari Africans in AMAR (+1)
MAR group    Makonde/Yao  Mozambique
             - disaggregated to Makonde, Yao, and Makua in AMAR (+2)
MAR group    Honamese South Korea
             - dropped because they don't meet population criteria in AMAR, but left
             in data (0)
MAR group    Scheduled Tribes India
             - disaggregated into Scheduled Tribes of East India,  Scheduled Tribes Of
             Northeast India, Scheduled Tribes Of North India, Scheduled Tribes Of
             West India, Scheduled Tribes Of South India in AMAR (+4)
MAR group    Chinese    South Vietnam
             - dropped country only coded from 1955-1975 (-1)
MAR group   Montagnards     South Vietnam
             - dropped country only coded from 1955-1975 (-1)
MAR group    Turkmen  China
             - name of group changes from Phase IV to Phase V; now called the Uygur.
             Counted in MAR_proper but not in mar_in_amar as a reminder that name
             differs between the 2 sets of data (-1)
*/

save "C:\Users\Aga\Dropbox UMD\Dropbox\A-MAR revisions\AMAR
Reconciliation\AMAR_database FINAL\AMAR.Phase1data.19402006.FINAL.1.June102016.dta"

clear

*making additional changes to variable names, other changes in final clean of data

*1 - renaming variables from lower-case to upper-case
rename country COUNTRY
rename year YEAR
rename auton AUTON
rename ccode CCODE
rename numcode NUMCODE

*2 - dropping variables from MAR Phase IV that are no longer coded in MAR Phase V or
AMAR Phase I
drop current
drop id
drop marstat
drop name
drop type
drop version
```

109

```
save "C:\Users\Aga\Dropbox UMD\Dropbox\A-MAR revisions\AMAR
Reconciliation\AMAR_database FINAL\AMAR.Phase1data.19402006.FINAL.1.June102016.dta",
replace
```

```
*****************************************************************************
```

VI.   Appendix III – Other information

Numcode assignment for countries that split and/or reunified

Several countries split or reunified during the coding duration of the MAR Phase IV data. Below is an explanation of how numcodes were assigned for these cases.

There are several groups that were originally coded in countries before they split and then were coded in the original and break-away countries after the split. For these cases, the numcode of the group in the original country (that split) remains the same before and after the split (i.e., a new numcode was not assigned for the group after the original country split), regardless of whether the group was also coded in the break-away country after the split.

This applies to the following groups: Hindus in Pakistan (same numcode used before and after split of Pakistan into Bangladesh, even though Hindus are also coded in Bangladesh); Papuans in Indonesia (same numcode used for Papuans before and after split of Indonesia into East Timor, even though Papuans also coded in East Timor); Afars in Ethiopia (same numcode used before and after split of Ethiopia into Eritrea, even though Afars also coded in Eritrea); Serbs in Yugoslavia (same numcode used before and after split of Yugoslavia into Bosnia, Croatia, Macedonia, Slovenia, and Serbia and Montenegro, even though Serbs also coded in Bosnia, Croatia, and Macedonia); Roma in Yugoslavia (same numcode used before and after split of Yugoslavia into Bosnia, Croatia, Macedonia, Slovenia, and Serbia and Montenegro, even though Roma also coded in Macedonia); Croats in Yugoslavia (same numcode used before and after split of Yugoslavia into Bosnia, Croatia, Macedonia, Slovenia, and Serbia and Montenegro, even though Croats also coded in Bosnia); and Bosniaks in Yugoslavia (same numcode used before and after split of Yugoslavia into Bosnia, Croatia, Macedonia, Slovenia, and Serbia and Montenegro, even though Bosniaks also coded in Bosnia). Regarding other countries that split, Czechoslovakia was not coded in the MAR Phase IV data before its split into the Czech Republic and Slovakia; therefore, there was no overlap of groups pre- and post-split.

For countries that reunified, a new numcode was assigned to the group after the reunification of the country to correspond to the new Correlates of War (COW) code used for the country. This applies to the following groups: Turks in West Germany (new numcode used for Turks in Germany after reunification of East and West Germany); Turks in East Germany (new numcode used for Turks in Germany after reunification of East and West Germany). Regarding other countries that unified, the Yemen Arab Republic and Yemen People's Republic were not coded in the MAR Phase IV data before they unified into Yemen; therefore, there is no overlap of groups pre- and post-unification.