

Socially relevant ethnic groups, ethnic structure, and AMAR

Journal of Peace Research
2015, Vol. 52(1) 110–115
© The Author(s) 2014
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0022343314536915
jpr.sagepub.com



Jóhanna K Birnir

Department of Government and Politics, University of Maryland

Jonathan Wilkenfeld

Department of Government and Politics, University of Maryland

James D Fearon

Department of Political Science, Stanford University

David D Laitin

Department of Political Science, Stanford University

Ted Robert Gurr

Department of Government and Politics, University of Maryland

Dawn Brancati

Department of Political Science, Washington University in St Louis

Stephen M Saideman

Norman Paterson School of International Affairs, Carleton University

Amy Pate

National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland

Agatha S Hultquist

Department of Government & Politics, University of Maryland

Abstract

Protracted conflicts over the status and demands of ethnic and religious groups have caused more instability and loss of human life than any other type of local, regional, and international conflict since the end of World War II. Yet we still have accumulated little in the way of accepted knowledge about the ethnic landscape of the world. In part this is due to empirical reliance on the limited data in the Minorities at Risk (MAR) project, whose selection biases are well known. In this article we tackle the construction of a list of ‘socially relevant’ ethnic groups meeting newly justified criteria in a dataset we call AMAR (A for All). We find that one of the principal difficulties in constructing the list is determining the appropriate level of aggregation for groups. To address this issue, we enumerate subgroups of the commonly recognized groups meeting our criteria so that scholars can use the subgroup list as one reference in the construction of the list of ethnic groups most appropriate for their study. Our conclusion outlines future work on the data using this expanded dataset on ethnic groups.

Keywords

AMAR, ethnic groups, MAR

Corresponding author:

jkbirnir@umd.edu

Protracted conflicts over the status and demands of ethnic and religious groups have caused more instability and loss of human life than any other type of local, regional, and international conflict since the end of World War II (Harff & Gurr, 1989). Yet prior to the mid-1980s with the publication of Horowitz's seminal volume (Horowitz, 1985), ethnic conflict was an important topic of empirical research mainly for sociologists concerned with interethnic relations in immigrant societies and political scientists tracing the rise of nationalism 'from peoples to states'. In the early 1990s, with the breakup of the Soviet Union and Yugoslavia along ethnonational lines, research on ethnic conflict, perhaps overpredicting its prevalence (Fearon & Laitin, 1996), flourished, and focused not only on post-communist states but on self-determination movements in other multi-ethnic states as well. As the focus broadened, research came to include all ethnic and religious identity groups that provide a basis for political mobilization and action.

In spite of a growing social scientific interest in the topic in the 1990s, the Minorities at Risk (MAR) project was the only sustained effort to collect systematic and replicable data on politically active communal groups and their political actions. Originally designed in the late 1980s by Ted Robert Gurr with encouragement from James Scarritt and assistance by Monty G Marshall to enumerate minorities 'at risk', that is, any group that 'collectively suffers, or benefits from, systematic discriminatory treatment vis-à-vis other groups in a society; and/or collectively mobilizes in defense or promotion of its self-defined interests' (Minorities at Risk Project, 2009: 1), the dataset and associated activities became a public good, used by academics, journalists, governments, and nongovernmental organizations to answer a variety of descriptive and analytic questions unanticipated at the project's inception.

This intensified scholarly attention also revealed limitations with the MAR dataset, highlighting the selection of groups 'at risk' that is not necessarily representative of the larger population of ethnic groups (Fearon & Laitin, 1996, 2002, 2003; Fearon, 2003; see also Öberg, 2002a; Hug, 2003, 2013; Birnir, 2007; Brancati, 2006, 2009).

Subsequent data collections have put forth expanded lists of included groups. One of these well-known efforts is the Ethnic Power Relations (EPR) dataset (Wimmer, Cederman & Min, 2009) that includes information on politically mobilized groups only (see also Scarritt & Mozaffar, 1999; Posner, 2004). However, similar to MAR in their basic approach, the EPR data do not enumerate comparison groups that do not engage in the specified activity. In contrast, Öberg (2002b) does

enumerate some comparison groups for MAR but the field of ethnic politics still lacks a sampling frame including a more complete set of politicized and unpoliticized minority groups – and majorities – for drawing less biased samples for the study of ethnic politics.

This lack of a more complete sampling frame is a substantial obstacle to the accumulation of knowledge about relationships between ethnicity and a multitude of outcomes including ethnic conflict, the design of political institutions, the conflict-management strategies of governments in multi-ethnic states, and the international consequences of and responses to ethnic warfare. The objective of this article is to provide one such sampling frame of an expanded group list AMAR (A for all) that includes nearly 1,200 socially relevant groups that are not selected on any politically defined criteria such as being 'at risk' (MAR) or 'politically relevant' (EPR).

The article is organized as follows. We first define the types of groups that meet our new criteria on inclusion. We then discuss an arguably 'best practice' in preparing such a list that centers on transparency of subgroup listing, which facilitates re-aggregation for purposes of examining divergent ethnic configurations. We conclude with discussion of future research.

Socially relevant ethnic groups

Our new criterion of inclusion in this sample frame is of groups that are socially relevant without any necessary political activation. By 'socially relevant', as defined by Fearon (2006: 852), we mean 'when people notice and condition their actions on ethnic distinctions in everyday life'. Fearon contrasts this to the politicization of ethnicity, that is, 'when political coalitions are organized along ethnic lines, or when access to political or economic benefits depends on ethnicity' (Fearon, 2006: 852). Social (and political) identities, in turn, are subsets of all existing ethnic structures (Chandra & Wilkinson, 2008: 523). Importantly, social relevance of an identity does not refer to political mobilization (though socially relevant groups may become mobilized), and does not have inherent political connotations, but only refers to the salience of the identity in guiding an individual's actions in life. If the criteria for selecting groups is political mobilization or being 'at risk', then studies that attempt to estimate the impact of some variable on the likelihood that an ethnic group experiences some outcome can suffer from selection bias. For example, if protest is associated with higher risks of ethnic violence, and discrimination causes both politicization of ethnicity and protest, then failing to consider socially

relevant but not political mobilized ethnic groups can lead a study to underestimate or entirely miss the effect of discrimination on conflict.

Consequently, the AMAR criteria that aim to outline socially relevant groups at a given point in time are that:

- (1) Membership in the group is determined primarily by descent by both members and non-members.¹
- (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 1% of a country's population.

It cannot be overemphasized that social relevance of ethnic identity is fluid and context dependent, albeit sticky. We are well aware of the theoretical complexities in the creation of any list of ethnic groups or practices. Because of the fluidity of identity, no one list of ethnic groups is correct or comprehensive in any absolute sense. Furthermore, types of socially relevant identity vary between countries. Therefore, we endeavored to let the list of socially relevant identities emerge organically for each country from the sources consulted. To this end we consulted a wide variety of general and country-specific sources including but not limited to Ethnologue (a valuable source that does not select groups based on their activation), Minority Rights Group International, various encyclopedias, census data, academic articles and books, news articles, and prior accumulations of data enumerating ethnic groups.

Following construction, the list was then reviewed by a number of regional experts and revised repeatedly.² Applying the above selection criteria to the world's ethnic groups resulted in the enumeration of roughly 1,200 groups, over 900 of which were not in the original MAR dataset.

AMAR's coding transparency

The most difficult aspect of the list construction was to decide upon the most appropriate aggregation of overlapping groups and groups that contain many subgroups and how to handle cross-cutting identities. Within a country many of the socially relevant ethnic groups defined on the same axis of identity are mutually exclusive. These include, for example, groups defined by their primary language, caste, race or religion. Importantly, however, in other cases socially relevant ethnic groups are not mutually exclusive and individuals may identify with different socially relevant aggregate groups at different times. For example, in Italy Neapolitano-Calabrese and Lombards likely at times consider themselves simply Italian, which AMAR also lists as an ethnic category for Italy. For such overlapping identities AMAR lists both when the sources suggest both are socially relevant.

In other cases our sources suggested an aggregate classification with a complex subgroup structure. The Indian and the Nepalese caste systems are excellent examples of aggregate groups with complex subgroup structures. In India and in Nepal, the caste system is a widely recognized form of social organization. Within each caste, however, many different groups coalesce culturally and/or organize politically around smaller (often regional) subgroupings such as tribe, clan or other types of subcommunities. These 'set/subset' structures are very common, and are the basis for what Okamura (1981) termed 'situational ethnicity' in his analysis of the phenomenon in Africa. For enumeration and analysis there is no 'one size fits all' solution with respect to the most appropriate aggregation of subgroups. Instead, alternative possible aggregations need to be considered in any empirical analysis and different projects will likely choose different levels of aggregation.

Taking seriously the idea that a different sample frame of ethnic groups may be appropriate for different questions – and taking the first step toward the creation of a map of ethnic structure that allows for variation in aggregation – we endeavored to construct the AMAR list in as transparent a way as possible. This is – in our view – a 'best practice' that aims to make the sampling frame of available ethnic identity dimensions amenable to project-specific researcher manipulation. To this end, in addition to the aggregate groups we list as an appropriate sample frame for a given set of questions, the list itself also enumerates by name the principal nested-groups (subgroups). This is the first list of ethnic groups that encourages users to critically assess and even re-aggregate the ethnic dimensions most appropriate

¹ The group may be a caste determined by descent.

² In the future the list will have to be updated periodically because while sticky, ethnic practice is also fluid over time.

Table I. Snapshot of aggregate socially relevant ethnic groups meeting the AMAR criteria as nested in the overall list of ethnic structure: India as example

<i>18 commonly recognized socially relevant AMAR groups</i>	<i>Illustrative number of subgroups* (structure) listed by name in the AMAR data</i>
Bodos	3
Forward classes	576
Kashmiri Muslims	53
Mizos	0
Muslims	241
Nagas	0
Officially backward classes	413
Scheduled classes (Dalits)	363
Sikhs	0
Syrian/Malabar Christians	0
Tripuras	0
Jains	0
Northeast scheduled tribes	50
North scheduled tribes	35
West scheduled tribes	22
South scheduled tribes	26
East scheduled tribes	31
Assamese	0

*Subgroup enumeration is not necessarily complete and further disaggregation of subgroups is certainly possible if required by research. *Principal source:* Singh (1992–98).

for their analysis in this way. The enumeration of subgroups is not comprehensive (it may not ever be as identity changes over time), but it gives a good idea of the underlying structure of each aggregate group. In India, for example, the classifications outlined in Table I meet our aggregate AMAR criteria – but we also make available a listing by name of the subgroups included. Further, by providing data on different levels of nested ethnic categories, AMAR makes possible richer and more empirically based constructivist research on the factors that determine which levels of ethnic identification are politicized under different conditions (see Posner’s [2005] study of political coalitions in Zambia, for an example).

Yet another configuration is cases where ethnic identity dimensions cross-cut rather than overlap. For example, AMAR counts Muslims in France as an aggregate identity group because members of the group share a religion that is practiced by a majority of the group, our sources suggest the group is recognized and viewed as important by members and non-members, and membership is primarily ascriptive. At the same time our aggregate classification of Muslims in France lists the names of 22 distinct Muslim subgroups as disparate

as Algerians and Wolof. The subgroups cross-cut the aggregate religious cleavage on a number of dimensions including race. Researchers interested in race in domestic French politics might prefer to reconfigure our classification of Muslim subgroups to create, for example, an aggregate group of black immigrants. Under this aggregate classification of black immigrants, the researcher might then list as subgroups black Muslim immigrant groups such as the Wolof along with non-Muslim black immigrant groups such as the mostly Christian Fon that we currently classify under the aggregate heading of Afro-French along with five other groups. The aggregate groups we list emerged as salient from the sources we consulted, but the transparent listing of the principal subgroups subsumed under every aggregate AMAR category is intended to allow researchers to reclassify and examine the effect of such cross-cutting cleavages. Importantly, not all possible identities (cross-cutting and other) will appear in AMAR. In India, for example, the data are organized around religion, caste, and tribe, with national identities such as Bengali and Marathi omitted. Meanwhile for Nigeria, the data are organized around tribe, with religious identities (Muslim, Tradition, and Christian) omitted. Follow-on work maps additional cross-cutting group identities on a specific dimension such as religion (Birbir & Satana, 2013), but much work remains before the complex mosaic of world ethnic identity is more complete.

Conclusion

In this article we outlined the construction of the new AMAR list of socially relevant ethnic groups. This is the first attempt at constructing a list of ethnic groups that is not defined by any political criteria – such as ‘at risk’ in MAR or ‘politically relevant’ in EPR. The construction of the AMAR list is also a first attempt at outlining socially relevant groups in a more transparent fashion, beginning to account for not only the aggregate groups but underlying ethnic structure as well. This type of list construction facilitates examination of the aggregate constructed groups and re-aggregation as appropriate for different types of research.

Our follow-on work compares the AMAR list with other collections of lists of ethnic groups including MAR, EPR, Fearon (2003), and Alesina et al. (2003). To deal with the effects of selection bias on the dependent variable in MAR, this forthcoming work also uses the AMAR list to draw a random sample from new groups not in MAR. We then code this random sample of new AMAR groups for the 40 most commonly used MAR variables and

match it with the original MAR data. The resultant data constitute an unbiased sample more representative of the universe of socially relevant groups. These data can then be used by researchers to verify extant MAR analyses and carry out new unbiased analyses using the suite of the most commonly used MAR variables. Some of the new questions we can answer with this unbiased sample include the causes of mobilization and the relationship between ethnic heterogeneity and conflict, neither of which could previously be fully understood due to known selection issues and the truncation of extant data. Researchers may also use these data to help sort through the differing effects of divergent cleavages and cross-cutting cleavages on political outcomes ranging from ethnic war to peace.

The scholarly communities working in this area are engaged with each other, taking the criticisms of current lists of ethnic groups seriously, so that we can advance the scholarship on the causes and dynamics of ethnic conflict. In line with the ideas of best practices in research on the complex subject that is ethnicity, AMAR represents one transparent advance in this conversation. The next steps focus on attempting to verify that which we think we know about ethnic conflict with the expectation that we will have greater confidence that our new work has fewer biases than previous efforts.

Replication data

The AMAR list can be found at <http://www.cidcm.umd.edu/mar/> and <http://www.prio.no/jpr/datasets>.

Acknowledgements

The work in this article was funded by NSF grant no. SES0718957 Minorities at Risk: Addressing Selection Bias Issues and Group Inclusion Criteria for Ethno-Political Research. Parts of this work have been presented at the UCLA Workshop on Ethnicity Datasets, the Stanford IR Workshop, the Maryland Comparative Workshop, the Uppsala Workshop on Ethnic Conflict Data, and the Folke Bernadotte and Penn State University Conflict Data Workshop. We thank Andreas Wimmer, Kathleen Cunningham, Borjan Zic, Tanja Ellingsen, and workshop participants for helpful comments and suggestions.

References

Alesina, Alberto; Arnaud Devleeschauwer, William Easterly, Sergio Kurlat & Romain Wacziarg (2003) Fractionalization. *Journal of Economic Growth* 8(2): 155–194.
 Birnir, Jóhanna K (2007) *Ethnicity and Electoral Politics*. Cambridge & New York: Cambridge University Press.

Birnir, Jóhanna K & Nil Satana (2013) Religious groups and terrorism: A violent match? Paper presented at the Annual Convention of the International Studies Association, San Francisco, April.
 Brancati, Dawn (2006) Decentralization: Fueling the fire or dampening the flames of ethnic conflict and secessionism. *International Organization* 60(3): 651–685.
 Brancati, Dawn (2009) *Peace by Design: Managing Intrastate Conflict through Decentralization*. Oxford: Oxford University Press.
 Chandra, Kanchan & Steven Wilkinson (2008) Measuring the effect of ethnicity. *Comparative Political Studies* 41(4–5): 515–563.
 Fearon, James D (2003) Ethnic and cultural diversity by country. *Journal of Economic Growth* 8(2): 195–222.
 Fearon, James D (2006) Ethnic mobilization and ethnic violence. In: Barry R Weingast & Donald Wittman (eds) *Oxford Handbook of Political Economy*. Oxford: Oxford University Press, 852–868.
 Fearon, James D & David D Laitin (1996) Explaining inter-ethnic cooperation. *American Political Science Review* 90(4): 715–735.
 Fearon, James D & David D Laitin (2002) Ethnicity, insurgency and civil war (research project) (<http://www.stanford.edu/group/ethnic/>).
 Fearon, James D & David D Laitin (2003) Ethnicity, insurgency, and civil war. *American Political Science Review* 97(1): 75–90.
 Harff, Barbara & Ted Robert Gurr (1989) Victims of the state: Genocide, politicide and group repression since 1945. *International Review of Victimology* 1(1): 23–41.
 Horowitz, Donald (1985) *Ethnic Groups in Conflict*. Berkeley, CA: University of California Press.
 Hug, Simon (2003) Selection bias in comparative research: The case of incomplete data sets. *Political Analysis* 11(3): 255–274.
 Hug, Simon (2013) The use and misuse of the ‘Minorities at Risk’ Project. *Annual Review of Political Science* 16: 191–208.
 Minorities at Risk Project (2009) *Minorities at Risk Codebook*. College Park, MD: Center for International Development and Conflict Management (http://www.cidcm.umd.edu/mar/data/mar_codebook_Feb09.pdf).
 Öberg, Magnus (2002a) The onset of ethnic war as a bargaining process: Testing a costly signaling model. Doctoral dissertation. Department of Peace and Conflict Research, Uppsala University.
 Öberg, Magnus (2002b) Minorities Not ‘at Risk’: A control group for use with Minorities at Risk data. Dataset. Department of Peace and Conflict Research, Uppsala University.
 Okamura, Jonathan Y (1981) Situational ethnicity. *Ethnic and Racial Studies* 4(4): 452–465.
 Posner, Daniel (2004) Measuring ethnic fractionalization in Africa. *American Journal of Political Science* 48(4): 849–863.

- Posner, Daniel (2005) *Institutions and Ethnic Politics in Africa*. New York: Cambridge University Press.
- Scarrit, James R & Shaheen Mozaffar (1999) The specification of ethnic cleavages and ethnopolitical groups for the analysis of democratic competition in contemporary Africa. *Nationalism and Ethnic Politics* 5(1): 82–117.
- Singh, Kumar Suresh (1992–98) *People of India: National Series*, Volumes 1–6. Calcutta: Anthropological Survey of India; Delhi: Oxford University Press.
- Wimmer, Andreas; Lars-Erik Cederman & Brian Min (2009) Ethnic politics and armed conflict: A configurational analysis of a new global dataset. *American Sociological Review* 74(2): 316–337.
- JÓHANNA K BIRNIR, b. 1969, PhD (University of California, Los Angeles, 2001); Associate Professor, University of Maryland, College Park.
- JONATHAN WILKENFELD, b. 1942, PhD (Indiana University, 1969); Professor, University of Maryland, College Park.
- JAMES D FEARON, b. 1962, PhD (University of California, Berkeley, 1992); Geballe Professor in the School of Humanities and Sciences and Professor, Department of Political Science, Stanford University.
- DAVID D LAITIN, b. 1945, PhD (University of California, Berkeley, 1974); James T Watkins IV and Elise V Watkins Professor of Political Science, Stanford University.
- TED ROBERT GURR, b. 1936, PhD (New York University, 1965); Distinguished University Professor, Emeritus, University of Maryland.
- DAWN BRANCATI, b.1975, PhD (Columbia University, 2003); Assistant Professor, Washington University in St Louis.
- STEPHEN M SAIDEMAN, b. 1966, PhD (University of California, San Diego, 1993); Paterson Chair in International Affairs, Carleton University.
- AMY PATE, b. 1975, PhD (University of Maryland, 2007); Senior Researcher, National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland, College Park.
- AGATHA S HULTQUIST, b. 1978, MA (University of Denver, 2003), PhD expected (University of Maryland, College Park, 2016); PhD Candidate and Graduate Research Assistant, University of Maryland, College Park.