Encyclopedia of Associations
Report for the APSA Short Course, August 30, 2006

Contact:
 Frank R. Baumgartner, Frankb@psu.edu
 John McCarthy, jxm516@psu.edu
 Shaun Bevan, graduate assistant, sxb964@psu.edu

Overview:
The *Encyclopedia of Associations* (EA) has been published nearly annually since 1956.
Currently it lists approximately 23,000 voluntary associations active at the US national level.
We have started by coding the entire 1995 edition. We are then moving both forwards and
backwards and expect to complete a full database covering the entire history, in 5-year
intervals. We are not doing each edition of the EA because we have found that the
information is not typically updated each year in any case. We can extrapolate based on
overall counts in each of 16 sections to get estimates of annual numbers of groups per topic
area. We are coding by the major topic level only. Note that the EA also includes keywords
which are included in the database as well. For some subtopic categories, these keywords
may correspond quite well. Before releasing the data we will identify those matches as well.

Like the New York Times dataset, the EA dataset has a number of topic categories that have
no corresponding topics in the government-focused datasets, as is shown below. This is
because there are great numbers of hobby-, recreation- and other non-policy related
associations listed.

As of August 2006, we have completed two full editions of the EA. By January 2007 we
expect to have data at ten-year intervals and by Summer 2007 we expect to release data at the
full 5-year interval level.

Note: We have previously created a complete database for the smallest of the EA subject
categories, labor unions. In addition, we created a partial database consisting of
approximately 25 editions (out of 40) for the Public Affairs section. Results from these more
comprehensive databases led to the current structure of the database.

More information:
Complete topic codebook, with examples and coder instructions:
http://polisci.la.psu.edu/faculty/Baumgartner/Ency/Ency_Codebook.pdf

Published paper based on the labor union database:
http://polisci.la.psu.edu/faculty/Baumgartner/SSR_06.pdf

Working paper summarizing the growth of Associations over time:
http://polisci.la.psu.edu/faculty/Baumgartner/EA_Data_Source.pdf

Encyclopedia of Associations, List of Cases per topic area, 1995 edition.

Code	Description	Count
1	Macroeconomics, taxes, and the economy	85
2	Civil Rights, Minority Issues and Civil Liberties	794
3	Health	2321
4	Agriculture	898
5	Labor, Employment, and Immigration	435
6	Education	1148
7	Environment	511
8	Energy	213
10	Transportation	541
11[1]	Family	409
12	Law and Crime	462
13	Social Welfare	474
14	Community Development and Housing Issues	183
15	Banking, Finance, and Domestic Commerce	2930
16	Defense	672
17	Telecommunications, Media, Space, Science and Technology	1147
18	Foreign Trade	199
19	International Affairs and Foreign Aid	1563
20	Federal Government, Public Policy Generally	456
21	Public Lands, Indian Affairs, Forests, Water, and Irrigation	160
24	State and Local Government and Policy	107
26	Weather and Natural Disasters	12
27	Fires	13
29	Sports and Recreation	884
30	Death Notices	23
31	Churches and Religion	1023
32[2]	Ideological, Social Cause, and Political Groups	571
40[3]	Performing, Fine and Creative Arts	1400
41	Culture, Heritage and History	1436
42	Hobbies, Collectors, Amusements and Clubs	1496
99[4]	Miscellaneous	68
	Total for the 1999 edition	22634

[1] Section 11 here is a subset of section 12, law, crime, and family in the PAP
[2] Section 32 here is a subset of section 20, public policy, in the PAP.
[3] Sections 40, 41, and 42 have no corresponding values in other parts of the PAP.
[4] Section 99, miscellaneous, will hopefully be resolved and eliminated before the data are released.

Memo

To: Those involved in PAP
From: Frank Baumgartner and John McCarthy
Re: Protests dataset
Date: August 29, 2006

Sociologists John McCarthy, Doug McAdam, and Sarah Soule have gathered data on all protests reported in the *New York Times* from 1960 to 1995, several tens of thousands of protests in all. Their dataset includes a comprehensive topic coding system as well as considerable information about each protest reported: size, date, place, target of the demand, and other information.

The topic coding system used in the Protest database is not compatible with the PAP system of subtopics since the data were collected independently and the list of topics was constructed inductively from the subject matter of the protests. There are many areas in the PAP topic system with no protests, but some particular areas (anti-war activities, nuclear power, civil rights) with many protests.

Baumgartner and McCarthy propose to recode the entire set of protest data into the 225 subtopic codes used in the PAP during the 2006–07 academic year. We can do this with available resources. It involves only recoding the topic systems with global revisions, not individual-level recoding of the data.

Future grant proposals may allow us to update the data as well.

Therefore we expect to be able to produce a time series of US national protests from 1960 to present within a year or two.