© Journal of Peace Research, vol. 34, no. 2, 1997, ISSN: 0022-3433, pp. 163–175

Is Collective Violence Correlated with Social Pluralism?

RUDOLPH J. RUMMEL
Department of Political Science, University of Hawaii at Manoa

In order to determine the contribution of social pluralism (ethnic, religious, and racial differences) to violence, diverse multivariate cross-national analyses were done. These involved 109 variables on conflict and violence; pluralism; and social political, economic, demographic, and cultural differences for all states, 1932 to 1982. The results show that pluralism is a multidimensional empirical concept; that regardless of which dimension is focused upon, it has a lesser relationship to violence than do other national characteristics, such as political freedom. And what relationship is found between pluralism and violence, holding other variables constant, is largely accounted for by the number of ethnic and religious groups in a state. Drawing on this and other studies, the conclusion is that where political power is centralized around a trans-plural group, such as a military junta or monarch, or trans-plural ideology, such as communism or fascism, then violence is highly likely, regardless of what plural units may or may not exist. And where power is centralized, nondemocratic, and highly dependent upon one's social group membership, such as ethnicity or religion, then collective violence is also highly likely.

1. *M. G. Smith on Pluralism*

One of the few social scientists to make social pluralism a central and organizing concept of a theory of political and social behavior was the anthropologist M. G. Smith.[1] Indeed, he focused his theoretical and field research on the nature and consequences of social pluralism. By pluralism he meant the division of society into socially and politically meaningful racial, ethnic, language, religious, and cultural units. These units are most socially and politically meaningful when they form corporate units, with explicitly recognized membership, offices, culture, and unit-internal and external relations. They thus organize and impact upon the behavior of their members. According to Smith, as society becomes divided into such corporate units, and the distribution of power, prestige, and wealth depends on the unit to which one belongs, the likelihood of collective violence increases.

Towards the end of his life, Smith tried to determine more precisely the relationship between pluralism and collective violence. In Smith's own words (Smith, 1991c):

> Having spent my life trying to clarify the conception of pluralism, I now wish to test and demonstrate its relevance for the solution of many urgent problems in the modern world. To that end I have compiled information on the demographic, economic, social and political characteristics of all sovereign nation-states, together with such detailed records as I can gather of internal collective violence in them, (i.e., coups, revolts, attempted secessions, riots, pogroms, terrorism, purges, massacres, genocides), or changes of constitutional regime, from 1932 to 1982. This global compilation should allow me to determine the exact contributions of the plural conditions of these states to their histories of internal violence and disorder in that period.

To this end Smith collected data on 166 sovereign states, between indicated years, on 187 classifications and variables. Among these, of course, were a large number measuring differences in social diversity and pluralism among states.

It was a great loss to anthropology and to our knowledge of violence that Smith died in 1993 before he was able to do his analyses. Fortunately, however, he did complete his data collection and the preparation of his data for analysis, and he was able to roughly sketch out his ideas as to the appropriate analyses to be done (Smith, 1991c).

> The problem now is, on the assumption that the data are clean, and as nearly complete as I can make them, to devise a statistical model to determine whether there is any causal link between the data on social composition and the levels or types of internal collective disorder that the societies experienced from 1932 to 1982. A negative answer to this question indicates that we have found no such bond, despite exhaustive and systematic attempts to do so, while positive results may be explicitly causal or may have the form of correlations. However, as correlations do not indicate causes, by stepwise regression analysis of covariation and other methods, we shall seek to demonstrate and measure the precise contributions of such specific factors as political organization, economy and demography to historical events, including the inci-

dence of internal collective violence or disorder in these states.

Smith went on to make clear that he envisioned holding the political, demographic and economic variables constant while determining the correlational relationships of social composition to collective violence. In this he also envisioned the use of path analysis and/or recursive causal analysis.[2]

After Smith's death, Mary Smith, his wife, made available to several anthropologists his data and code books on a diskette, with the hope that others would carry through the analysis. The data were all neatly compiled on an Excel spreadsheet and the code book and relevant notes were in Word.[3] I became aware of these data in early 1995 and felt it would be a huge waste if someone did not carry through Smith's analysis. Because of my background in analyzing similar data and interest in collective violence (see Rummel, 1972, 1979, 1976–81), I decided to do so.

2. *Data Preparation*

First, I completely reconstructed Smith's nominal categories, making dichotomous variables out of them where possible.[4] Second, since many of Smith's variables count the number of coups, revolts, purges, deaths, and the like from 1932 to 1982, and many states only became independent in the 1960s and 70s, I normed these and other such variables by the number of years of independence since 1932 (also one of Smith's variables). Third, where data were clearly skewed toward a few states with values, I did a $\log_{10} \chi+1)$ transformation.[5]

Throughout this revision of Smith's data, whether creating new variables out of his set or transforming his variables, I kept in mind that his dominant interest was in the causal or predictive relationship of social pluralism to collective violence. I thus made every effort to include Smith's relevant classifications or variables.

Once this initial revision of Smith's data was completed, which created 109 variables, I then tried to minimize the amount of missing data. I first eliminated all states with less than 80% of the data across the variables; and then similarly removed variables with less than 80% data across the remaining states. I ended up with 162 states and ninety-eight variables, with no more than 4% missing data overall.

3. *Research Design*

A regression analysis involving ninety-eight variables is out of the question. Multicollinearity among the independent variables alone would defeat this effort, not to mention that the multiple correlation coefficient would be inflated by the gross capitalization on random error among so many variables. This analysis will be done in four stages, therefore. The first will be separate component (factor analyses) of the violence, pluralism, political, and economic/demographic variables. This will enable me to select statistically independent indicators within each of these domains for the regression analyses. Second, I will do a common factor analysis of all the indicators to determine whether there is a common factor (causal nexus) underlying collective violence and social pluralism. Third, I will follow this with the regression analyses of collective violence, the major aim of this study. And finally, I will do a canonical analysis to determine how well all the independent indicators predict overall to collective violence.

I should say that the results of these analysis should be considered descriptive, where the emphasis is on percent of covariation accounted for or predicted or in common, rather than inferential. I will, however, use significance tests of the regression analysis as simple benchmarks.

4. *The Component Analyses*

For the 28 variables measuring different kinds and aspects of collective violence, I did three kinds of factor analysis. One was a component analysis of data, where any case with at least one missing datum was eliminated from the analysis (called *listwise deletion*). The second was a component analysis with the substitution of a variable's mean for its missing data (called *mean substitution*). And the third was an image (common) factor analysis with mean substitution. To each of these three factor analyses, both orthogonal and oblique (biquartimin) rotation were applied.

Space does not allow me to display the results of these stage-one analyses here, I should note, however, that all the correlations for the internal

violence variables are positive, meaning that there is a tendency for states that have one kind of violence to have others. This was also reflected in the unrotated dimensions, the first of which for the component analysis (mean substitution) accounted for near 35% of the variance among the twenty-eight violence variables; the first three together raised this to 55%. To account for this much variance among this many variables by only three dimensions is impressive.

Table I presents the consolidated dimensions and indicators from the three analyses and their orthogonal and oblique dimensions.[6] I selected them by the size of their correlation with a dimension, their substantive importance, and available data. These six indicators will now define collective violence for the rest of this.

Next, I similarly reduced the twenty-three variables measuring various aspects of social pluralism to their indicators. Like violence, social diversity is highly structured along a few unrotated dimensions, four of which are sufficient to account for over half of the variance among the twenty-three measures in a component analysis (mean substitution). The first of these dimensions alone account for 23.5% of the variance, with the percentage of population of one ethnic group (ETHNIC_) being most highly correlated with it. This means that ETHNIC_ is overall the best indicator of social diversity. Overall, Table II shows the best indicators of the independent clusters of intercorrelation among the variables.

There are eight dimensions that can be consolidated from the separate component and image analysis. Space does not allow their full discussion, but I should note that these dimensions and their indicators define plural dominance, type of pluralism, and the actual diversity along racial, ethnic, and religious lines.[7] And these are fairly statistically independent of each other. The associated indicators will comprise the central independent variables in the forthcoming regression analysis.

Smith also collected data on variables to be held constant while investigating the relationship between violence and pluralism. One set of these defined aspects of the state, government, and politics in or round 1982. From these data it is able to determine twenty political variables. As above, it was possible to use several compo-

Table I. Dimensions and Indicators of Collective Violence*

	Indicator	
Dimension	Code	Name/Measurement
Internal War	INTWAR_R	Years of local civil & guerrilla war/years independent.
Turmoil	RIOTS_YR	Years that involved riots.
Purges	PUR_RAT1	Number of purges/years independent.
Revolution	REV_RATE	Number of general & local revolts & revolutions & attempted successions/years independent.
Coups	COUPS_RA	Number of successful & unsuccessful coups/ years independent.
Violence Intensity	VIOLENCE	1 = no internal violence; 2 = little; 3 = violence (without internal war); 4 = internal war.

* Based on component analyses with mean and list wise substitutions for missing data; and an image analysis with mean substitution.

Table II. Dimensions and Indicators of Social Pluralism*

	Indicator	
Dimension	Code	Name/Measurement
Hierarchic	HIERARCH	Hierarchic society = 1; not = 0.
Native Pluralism	NAT_RACE	Number of indigenous racial stocks.
Ethnics	ETHNICS	Number of ethnic groups.
Plural Emigration	EMIG_RAT	Emigration of plural units sought or achieved/years independent.
Complex Pluralism	PLUR_TYP	1 = not plural; 2 = predominant segmental pluralism; 3 = predominant hierarchic pluralism; 4 = complex pluralism.
Plural Deportation	DEP_RATE	Log 10 (number of plural units deported/years independent).
Religions	RELIGION	Number of local religions.
Pluralism	PLURALIS	Plural = 2; cultural only = 1; not = 0.

* See note Table I.

nent analysis and rotations to reduce this number to the minimum indicators of their clusters of intercorrelation.

These political data are also highly structured (intercorrelated), with the first unrotated component dimension (mean substitution) accounting for 25% of the total variance and the first four dimensions for over half. Noteworthy is that the variable most highly correlated with the first dimension, and thus the best indicator of politics overall, is FREEDOM, which is the Freedom House rating of the civil liberties and political rights of all states. The degree to which the people of a state are free in their rights and liberties is the best measure of the nature, policies, and type of their political system. As to these dimensions of politics and the indicators, see Table III.

Component analyses of political variables have usually found three dimensions: democracy; totalitarianism, and authoritarianism. Democracy and totalitarianism are defined in these analyses (the Freedom and Centralization dimensions), but authoritarianism could not be clearly delineated, since Smith has no measures of the monarchical nature of a regime (as of Jordan, Saudi Arabia, or Kuwait), the prime indicator of authoritarianism.

Smith also collected data on the socio-economic and cultural characteristics of states, from which I included or constructed twenty-three variables. These were also component analyzed as above. The resulting major dimensions (not shown here) are those usually found for states, being wealth (or development) and size (see Rummel, 1972). The other dimensions define different cultural characteristics of states, more specifically whether they were Moslem, Animist, Asian, or in Latin America. The indicators of these are shown in Table IV.

Finally, there is a set of four variables that fit between those measuring violence and pluralism. These are the number of plural units involved in collective violence since 1932 or independence (PLU_VIO), this number divided by the years independent 1932–82 (PLU_VIO_), the number of ethnic groups involved in collective violence since independence or 1932 (ETH_VIO), and this number divided by the years independent 1932–82 (ETHNIC–V). Since these variables measure both the violence and the pluralism of a society (for there to be significant ethnic or plural violence, a society must be ethnically divided or have plural units to begin with), they span both domains. For this reason they cannot be used in a regression analysis (they create a logical dependence between dependent and independent variables), but can be included in the combined image analyses to be presented below. A component analysis was conducted on the four to determine their indicators and found that they reduced to one dimension and one indicator: ETHNIC_V.

We now have the basic set of indicators for violence, pluralism, politics, and other aspects of society and the state. For these data we now can determine whether and how pluralism is related to violence.

Table III. Dimensions and Indicators of Politics*

	Indicator	
Dimension	Code	Name/Measurement
Freedom	FREEDOM	(political rights scores) + (civil liberties scores).
Legislature	LEGISLAT	1 = has a legislature; 0 = none.
StateAge	INDEP_AG	Years independent since 1/1/32 to 12/82.
Centralization	CENTRALI	1 = decentralized; 2 = centralized; 3 = strongly centralized.
Stability	STABILIT	Regime changes since 1/1/32 or independence/years independent.
Presidential	PRES_GOV	1 = presidential govt.; 0 = no

* See note Table I.

Table IV. Dimensions and Indicators of Socio-economic characteristics and Culture*

	Indicator	
Dimension	Code	Name/Measurement
Wealth	GDP_PC	Log 10 (gross domestic product per capita.
Size	AREA	Log 10 (area of the state in square kilometers).
Moslem	MOSLEM	Moslem society = 1; not = 0.
Animist	ANIMIST	Animist society = 1; not = 0.
Asia	ASIA	Asian state = 1; not = 0.
Latin American	LATIN AMERICAN	Latin American state = 1; not = 0.

* See note Table I.

5. *Violence and Pluralism: A Causal Nexus?*

Pluralism and violence indicators were first analyzed through image factor analysis. Image analysis delineates the common factors underlying the intercorrelations among variables. If pluralism is a common dimension to violence, that is, if they form a causal nexus, then image analysis should not only uncover this for the sixteen indicators, but also define the specific aspects of violence and pluralism most and least intercorrelated.

The results of this analysis give the first indication of a relationship, although small, between violence and pluralism, and some idea as to the precise nature of this relationship. I do not wish to over interpret these results here, since there is more analyses to come, but at this point I can note three different factor patterns of collective violence correlated with pluralism. The ethnic division of a society is related to the first factor, which specifically involves ethnic violence (ETHNIC_V) and primarily overall intense and frequent violence (INTWAR_R). Religious divisions (RELIGION) is related to the second factor, which tends to reflect overall violence (VIOLENCE), popular violence (RIOTS_YR), and genocide and mass murder (DEMOCIDE). Note also that there is a high positive correlation of .55 between these two factors. When one kind of violence occurs so tends the other.

The third factor of violence and pluralism involves purges (PUR_RAT1) and coups (COUPS_RA) among the political and military leaders and a high rate of emigration of those of a particular ethnicity, race, religion, and cultural identity (EMIG_RAT). This factor also has a positive correlation of .54 with the intense violence one.

How do these relationships between violence and pluralism hold up when the political, economic, and cultural indicators are included? This now entailed a full image analysis of all twenty-eight indicators. The analysis supported a relationship between the number of ethnic groups (ETHNICS), ethnic violence (ETHNIC_V), and intense violence (INTWAR_R), independent of all the political, socio-economic, and cultural indicators.

However, while the relationship between religious groups (RELIGION), riots (RIOTS_YR), democide, and overall violence (VIOLENCE) also remained, even in the context of all the other indicators, it also included the length of time a state has been independent (INDEP_AG) and its size (AREA). That is, the longer a country has been independent (counting from 1932), the larger, and the more distinct religious groups it has, the more likely it will have extensive violence, riots, and democide.

There was a third factor pattern unrelated to pluralism, but largely to the political variables. This is that the violence among political leaders involving purges (PUR_RATI) and coups (COUPS_RA) is mainly related to how little civil and political rights there are in a state (FREEDOM) and its degree of centralization (CENTRALI).[8] Moreover, there is a tendency for this relationship to hold more for Asian cultures (ASIA) than those of other regions.

Most important, these image analyses identified two possible causal nexi involving both violence and pluralism. One is some kind of relationship between guerrilla war/revolution and the number of ethnic groups; the other is an entirely independent relationship between violence, riots, democide, the number of religious groups, and a state's age and size. The image analysis, however, does not tell us actually how much of the variation in these kinds of violence can be explained by (dependent upon) pluralism. This can be determined by regression and component analysis.

6. *How Much Violence Does Pluralism Predict?*

For the first regression, the rate of a state's guerrilla and civil war (INTWAR_R) was taken as the dependent variable. This is by far the best indicator identified with the first cluster of intercorrelations uncovering by the image analysis of all the indicators. Only one independent variable – the number of ethnic groups (ETHNICS) – is needed. Table V lists the regression results.

This regression shows that 21% of the variation (R Square) in intense violence (INTWAR_R) is accounted for by the number of ethnic groups (ETHNICS).[9] To be able to explain one-fifth of the variation among all states in such intense violence as guerrilla and civil wars from 1932 to 1982 is an accomplishment, and to do this with one variable – the number of ethnic groups – is even more important. And the

Table V. Regression of Intense Violence on Number of Ethnic Groups

Independent Variable:	INTWAR_R
Multiple R	.46
R Square	.21
Adjusted R Square	.21
Standard Error	.30

Analysis of Variance

	DF	*Sum of Squares*	*Mean Square*
Regression	1	4.0	4.0
Residual	163	14.9	.1
F = 44.2	Signif F = .0000		

Variables in the Equation

Variable	*B*	*SE B*	*Beta*	*T*	*Sig. T*
ETHNICS	.01	.002	.46	6.65	.0000
(Constant)	−.001	.03		−.04	.96

factor analyses show clearly that this is a direct relationship, after the effects of the correlation of other plural indicators, and political, social-economic, and cultural indicators have been removed.

What does this say then about predicting violence? In order to determine where in the world revolutionary and guerrilla violence is likely to occur in the future, an important indicator is simply the number of different and distinct ethnic groups a state has. This is not the most important indicator, however, which is the level of civil rights and political liberties, a state's freedom in short. This is clear from other studies.[10] The findings here now add this: *the more nondemocratic a state and the more ethnic groups it has, then the more likely it will have frequent revolutions and guerrilla war.*

There is one more factor involving violence and pluralism to clarify through regression. Be it recalled that previous analysis found one factor that comprised VIOLENCE, RIOTS_YR, and DEMOCIDE, RELIGION, INDEP_AG, AND AREA. The three violence indicators were not well differentiated in their loadings, so I carried out regressions on each of them. The best of these accounted for 28% of the variation in a state's overall violence by its number of religions, area, and to a lesser extent the years of independence since 1932. This is an even better result in variance terms, although mainly one helper variable – area – is required to have this strong a relationship. Moreover, here also we should take democracy as our primary predictor of general violence, then use the number of religions and size as a way of more reliably predicting differences in violence among non-democracies.

At this point one might ask why I did not just regress the separate indicators of violence on all the indicators of pluralism, politics, and socio-economic attributes? The answer is that the regression does not untangle the intercorrelations between the independent variables. Thus, were the regression carried out on all the indicators, it would be unclear how much of what relationships between violence and a pluralism indicator was due to the influence of other pluralism indicators and especially, that between the other indicators and both pluralism and violence. The *common* factor analysis separated out these interrelationships such that when we did a regression of those indicators loaded on separate factors, we know that the regression will deal with the direct effects.

Finally we can do a canonical analysis of the seven violence indicators on the twenty indicators of pluralism, politics, socio-economic characteristics, and culture. Table VI shows the results. The first column presents the best linear combination of dependent indicators (the upper half) fitting that of the independent variables (lower half). Each linear combination produces a variate, and the *canonical correlation* of the dependent variate with the independent one is shown between the two halves of the table. This correlation is similar to the multiple correlation coefficient in regression analysis.

Brackets in the table show the correlations of at least an absolute .30 between the indicators and variates; correlations within each substantive domain have been ordered by rank. Moreover, only three variate pairs with significant canonical correlations have been shown.[11] To the right of the table the communalities of each indicator (the sum of squared correlations across the variates) are given.[12] These show how much of the variance in an indicator is picked up by these three variates. And at the bottom of the table the *trace correlation* is given. This is the overall correlation of the *space* of violence with the *space* of the independent indicators.

With this background, I will step through the interpretation of the first column in the table to

make sure these results are understood. The first column shows the dependent and independent variates have a correlation of .98. VIOLENCE is very highly correlated with the first dependent variate and RIOTS_YR to a much lesser extent; AREA and GDP_PC are most correlated with the associated first independent variate. The extraordinary correlation of .98 between the two variates therefore means that there is a linear combination of mainly the overall violence in a state 1921–82 and its years of riots that is almost completely accounted for (explained, predicted) by a state's characteristics, especially a state's gross domestic product per capita and area. Simply, a state's potential development in 1982 and size have much to do with its degree of overall violence 1932–82.

Table VI. Canonical Analysis of Violence on Pluralism, Politics, Socio-economic Indicators, and Culture

	Indicators	Canonical Variates* 1	2	3	H-sq
Dependent	Violence	[.88]	.03	.24	.84
	RIOTS_YR	[.40]	[−.43]	[.40]	.51
	PUR_RAT1	.25	[.65]	−.19	.52
	COUPS_RA	.22	.23	.19	.14
	INTWAR_R	.17	[.31]	[.75]	.69
	REV_RATE	.21	[.36]	[.50]	.42
	DEMOCIDE	.23	−.09	[.32]	.16
	Correlation	.98	.76	.63	
Independent	PLUR_TYP	.18	.04	.18	.07
	EMIG_RAT	.04	[.30]	.04	.09
	HIERARCH	.06	.21	−.09	.06
	DEP_RATE	.03	.17	.16	.05
	RELIGION	.14	.02	[.39]	.17
	ETHNICS	.10	[.34]	[.37]	.26
	PLURALIS	.03	.03	.23	.05
	NAT_RACE	.14	−.10	.20	.07
	FREEDOM	.16	[−.54]	−.06	.32
	INDEP_AG	.22	[−.43]	.23	.28
	CENTRALI	.25	[.30]	−.16	.18
	PRES_GOV	.07	.20	.06	.05
	LEGISLAT	.03	.17	.01	.03
	STABILIT	.08	[−.35]	[−.36]	.26
	AREA	[.55]	.13	.09	.33
	GDP_PC	[.44]	[−.51]	−.28	.53
	MOSLEM	.04	.13	.04	.02
	LATION_AM	.03	−.21	[.40]	.20
	ANIMIST	.03	.28	[−.30]	.17
	ASIA	.03	.16	.17	.05

Trace Correlation Squared (three variates) = .28

* Only significant variates are shown. Coefficients are correlation of indicators with variates. Those >1.291 are shown in brackets. H-sq is the communality (sum of squared row correlations).

Looking now at all the results in the table, how well is violence accounted for by all the indicators. The trace correlation squared for the three variates is .28, which means that the indicators overall explain 28% of the total variation among the seven indicators of violence.

7. *Discussion and Conclusion*

First, social pluralism as defined by the 23 variables taken or created from Smith's data is highly structured (or patterned), with the variation among states in their social pluralism being along 8 separate and statistically independent clusters of intercorrelation among the pluralism variables. This shows that trying to define pluralism by just one or two scales or indices could well miss very important variation in the social diversity of states.

Second, collective violence can be well accounted for by variation among states in their various characteristics, such as potential and actual development, freedom, and their stability, age, size, and cultural region. Pluralism overall, by contrast, has the lesser relationship to collective violence.

And third, there is, however, two specific relationships between pluralism and violence that exist in the data, taking into account the direct and indirect effects of the political, socio-economic, and cultural aspects of states. The more ethnic groups in a state, the more likely it will have a high rate of guerrilla and revolutionary warfare. And the more religious groups in a society, the more intense the general violence. This is largely moderated by the size of a state. Thus, the larger and older (counting from 1932) a state in addition to the more religious groups, the more the general violence.

In general, then, *pluralism is important, but less so than other aspects of society. And the importance largely resides in the number of ethnic and religious groups a state has.* This does confirm Smith's belief that there is a relationship between social pluralism and violence, although in specifics the results depart from his theory. He believed that pluralism had a much stronger causal effect on collective violence and that certain aspects of pluralism, such as the hierarchical distribution of power among plural units, their segmentation, and corporate nature would be the main predictors. What we have actually

found is that the more interesting theoretical measures, those of hierarchy and plural type, segmentation, potential separatism, and incorporation mode (some of these were not indicators, but related to the indicators – see Table III), among others, had no meaningful correlation with violence. *We end up with two rather simple and ordinary measures – numbers of ethnic and religious groups.*

More specifically, drawing other studies and this one together, where power is centralized around a trans-plural group, such as a military junta or monarch, or transplural ideology, such as communism or fascism, then violence is highly likely, regardless of what plural units may or may not exist. However, *when political power is centralized, nondemocratic, and highly dependent upon one's social group membership, be it race, religion, ethnicity, or some cultural division, then collective violence is also highly likely.*

NOTES

1. See in particular Smith (1975, 1984, 1991a), Kuper & Smith (1969), and Kallab (1994). From 1978 to 1986 Smith was the Franklin M. Crosby Professor of the Human Environment, Department of Anthropology, Yale University; and Franklin M. Crosby Professor Emeritus from 1986 until his death in 1993. For a similar but distinct approach to social pluralism, see Kuper's chapters in Kuper & Smith (1969).
2. There may be some doubt from his writings as to whether Smith thought social pluralism caused violence in general, or only in societies meeting particular criteria. The above quotes and the methods of analyses he suggested make clear that he thought of social pluralism as a general cause of violence, possibly modified in its effect by political and economic institutions and demographic attributes.
3. As of this writing Mary Smith is in contact with the Human Relations Area Files about distributing all these data.
4. The reconstructed list of variables from Smith's data set is given in Appendix I.
5. Units of measurement and transformations are listed in Appendix I for each variable.
6. The variable codes (which are alphabetized in Appendix I for convenience) were dictated by the limits on variable names in SPSS for the Macintosh.
7. I know of no comparable cross-national factor analysis of diversity *per se*, except for Rummel (1996b). In that analysis I collected data on eight measures of diversity for 204 political regimes, 1900–87. A component analysis uncovered two orthogonally rotated dimensions in these data, the major one most correlated with ethnic divisions (as here), and the second with the number of minorities at risk of genocide.
8. One might expect that freedom and centralization would have even a higher relationship to the other violence factors. But Smith measured freedom and centralization for only 1982, whereas he counted violence for all the years from 1932 to 1982, and thus would include that part of a regime's history before or after it was free and decentralized. Therefore, what I correlated here is the violence of a nation 1932–82 with its potential to be free or decentralized in 1982.

   Then why did I not include an appropriate variable that would measure the freedom with each state, 1932–82? To do this would have required changing the whole nature of Smith's data – from states to political regimes. And since a state may have had many regimes over the period 1932–82 this would have necessitated my recollecting from scratch all the violence and political data, an effort beyond the purpose of this study.
9. I also tried a polynomial (cubic) fit, and got an R^2 of .27.
10. These are all summarized in Rummel (1976–81, 1985, 1996a). As to why freedom does not come out here also as an indicator, and why I did not include an appropriate measure of freedom, see note 8.
11. This is by a test of the residual roots (eigenvalues) using the chi-square.
12. This is equivalent to the communalities in factor analysis.

REFERENCES

Kallab, Majda, 1994. *Testament: Life and Work of M. G. Smith 1921–1993*. New York: Research Institute for the Study of Man.

Kuper, Leo & M. G. Smith, eds, 1969. *Pluralism in Africa*. Berkeley, CA: University of California Press.

Rummel, Rudolph J., 1963. 'Dimensions of Conflict Within and Between Nations'. *General Systems: Yearbook of the Society for General Systems Research*, vol. 8. pp. 1–50.

Rummel, Rudolph J., 1972. *The Dimensions of Nations*. Beverly Hills, CA: SAGE.

Rummel, Rudolph J., 1976–81. *Understanding Conflict and War*. Volumes 1–5. Beverly Hills, CA: SAGE.

Rummel, Rudolph J., 1979. *National Attributes and Behavior*. Beverly Hills, CA: SAGE.

Rummel, Rudolph J., 1985. 'A Test of Libertarian Propositions on Violence'. *The Journal of Conflict Resolution*, vol. 29, September, pp. 419–55.

Rummel, Rudolph J., 1996a. *Power Kills*. Rutgers, RI: Transaction.

Rummel, Rudolph J., 1996b. *Statistics of Democide: Estimates, Sources, and Calculations on 20th Century Genocide and Mass Murder*. Charlottesville, VA: Center for National Security Law, University of Virginia.

Smith, M. G., 1975. *Corporations and Society: The Social Anthropology of Collective Action*. Chicago, IL: Aldine Publishing.

Smith, M. G., 1984. 'The Nature and Variety of Plural Unity', pp. 146–186 in David Maybury-Lewis, ed. *The Prospects for Plural Societies: 1982 Proceedings of the American Ethnological Society*. Washington, DC: American Ethnological Society.

Smith, M. G., 1991a. *Pluralism, Politics, and Ideology in*

the Creole Caribbean. New York: Research Institute for the Study of Man.
Smith, M. G., 1991b. 'World Survey: Notes on Codes and Analysis II'. Paper. Np: 18 November.
Smith, M. G., 1991c. 'World Survey: Notes on Data Analysis I'. Paper. Np: 18 November:

APPENDIX 1

Following is the list of variables and their codes for the analyses, and the name and number of the identical or related variables coded by M. G. Smith. The format of the variables are: Smith variable number; Rummel variable number, variable name, code name, and description/comments. For greater clarity as to the type of variable, variables are classified by substantive domain.

The list of the variables and their codes is an alphabetical listing of the code names. Smith's variable number, names, and descriptions are fro Smith (1991b).

Smith #	My # Variable Name*	Code	Description/Comments
DATA DESCRIPTION			
1	country	COUNTRY	
2	region	REG	1 = Europe; 2 = N. America; 3 = USSR; 4 = N. Africa; 5 = S.W. Asia and ME; 6 = S. Asia; 7 = S.E. Asia; 8 = China; 9 = N.E. Asia; 10 = Pacific; 11 = none; 12 = W. Africa; 13 = S.W. Africa; 14 = Southern Africa; 15 = S.E. Africa 16 = W. Africa; 17 = N.W. Africa; 18 = W. African Is.; 19 = Central America 20 = S. America; 21 = Caribbean Islands.
GENERAL CONFLICT/VIOLENCE			
69+70	1 coups	COUPS	Number of successful and unsuccessful coups 1/1/32 to 12/82 (Smith's var. 69+70) per year.
	2 revolts	REVOLTS	Number of general and local revolts/insurrections, attempted successions, revolution. Sum of Smith's var. 80–82, 83b–85.
104	3 purges	PURGES	Number of known purges.
73	4 political deaths	DEATHS	Average political deaths per year since independence or 1932 per 10,000 pop at 12/31/82
72	5 rate of violence	VIOL_RATE	(years of internal collective violence since independence or 1/1/32)/(years independent – Smith's var. 66).
DURATION			
83a	6 revolt years	REV_YRS	Years of revolt, insurrection or successionist violence.
91	7 years of indep. war	INDWAR_Y	Years of independence war.
	8 internal war years	INTWAR_Y	Years of local civil/guerrilla war (Smith's var. 92+93).
94	9 external war years	EXTWAR_Y	Years of external war.
RATES			
	10 purges per year	PUR_RAT1	Number of known purges (Smith's var. 104)/(years independent – Smith's var. 66).
	11 purge years rate	PUR_RAT2	%: 100* (no. of years purges lasted – Smith's var. 105)/(years independent – Smith's var. 66).
97	12 riot years	RIOTS_YR	% of known years that involved riots.
98	13 terrorism years	TERR_YR	% of known years that involved terrorism or lynching.
99	14 police/mil. viol. years	POL_YRS_	% of known years that involved police/military violence vs. plural units.
151	15 police/mil. action years	MIL_POL_	% of known years with military/police action.
153	16 armed attack years	ATTACK_Y	% of known years with armed attacks.
155	17 political killing years	POLKILL_	% of known years with political killings.
	18 coups per year	COUPS_RA	(number of successful and unsuccessful coups 1/1/32 to 12/82 (Smith's var. 69+70) per year)/(years independent – Smith's var. 66).
	19 revolts per year	REV_RATE	(Number of general and local revolts, insurrections, attempted successions, revolution. Sum of Smith's car. 80–82, 83b–85)/(years indepen. – Smith's var. 66).
	20 insurr. per year	INS_RATE	(Years of revolt, insurrection or successionist violence – var. 83A)/(years independent – Smith's var. 66).
	21 internal war rate	INTWAR_R	(Years of local civil/guerrilla war – Smith's var. 92, 93)/(years independent – Smith's var. 66).

Smith #	My # Variable Name*	Code	Description/Comments
	22 peaceful change rate	PEACE_CH	Rate of peaceful regime change: Smith's (var. 86+1)/(years independent – Smith's var. 66).
	23 imposed change rate	IMP_CHG	Rate of externally imposed regime change: Smith's (var. 87+1)/(years independent – Smith's var. 66).
	24 intervention years	INTERVEN	Log 10 (1+(% of known years that involved violent external intervention) – Smith's var. 95)).
	25 democide	DEMOCIDE	Log 10 (1+(% of known years that involved local pogroms, massacres, or genocides – Smith's var. 95)).
SCALES			
71	26 violence	VIOLENCE	1 = no internal collective violence since independence or 1/1/32; 2 = little; 3 = violence without war; 4 = internal war.
90	27 violent independence	VIO_IND	1 = independence was won by independent armed struggle (1 in Smith's rating); 0 = no (2 in Smith's).
184	28 violence intensity	VIO_INTE	1 = little or no violence; 2 = violence without war; 3 = internal war.
PLURALISM/VIOLENCE			
74	29 plural violence	PLU_VIO	Number of plural units involved in collective violence since independence or 1/1/32.
	30 plural violence rate	PLU_VIO_	(Number of plural units involved in collective violence since independence or 1/1/32)/(years independent – Smith's var. 66).
75	31 ethnic violence	ETH_VIO	Number of ethnic groups involved in collective violence since independence or 1/1/32.
	32 ethnic violence rate	ETHNIC_V	(Number of ethnic groups involved in collective violence since independence or 1/1/32)/(years independent – Smith's var. 66).
PLURALISM			
UNITS/INDICES			
32A	33 no. of native races	NAT_RACE	Number of indigenous racial stocks.
32B	34 no. of foreign races	FOR_RACE	Number of foreign racial stocks.
33A	35 no. of ethnic groups	ETHNICS	Number of ethnic units.
34A	36 plurality	PLUNITS	Number of plural units.
35	37 ethnolinguistic group	ETHNOLIN	Ethnolinguistic index.
43	38 cultures	CULTURES	Number of plural cultures.
44A	39 languages	LANGUAGE	Number of indigenous first languages.
45A	40 religions	RELIGION	Number of local religions.
46A	41 religions	REGION	Number of regional units. USSR estimated as much as India.
101	42 plunits emigration	PLU_EMIG	Number of plural units known to have sought or achieved emigration (5 = 5 or more).
	43 plunits relocated	PLU_RELO	Log 10 (1 + (no of plural units known to have been internally relocated – Smith's var. 102)).
103	44 plunits deported	PLU_DEP	Number of plural units known to have been deported.
SCALES			
	45 minority dominance	MIN_DOM	For 1982: dominant plural unit (plunit) from Smith's var. 47: 1 = minor plunit/ethnic group; 0 = none, or largest plunit/ethnic group.
	46 dominant plunit	DOM_PLUN	4-(Smith's var. 183: 1 = largest plural or ethnic unit; 2 = minor; 3 = none dominant).
	47 potential separatism	SEPARATE	(Potential separatism – Smith's var. 149)* (potential separatism – Smith's var. 150) = log 10 (1+(Smith's var. 149 *Smith's var. 150)).
185	48 de facto incorporation mode	INCORP	1 = no mode of incorporation; 2 = differential incorporation; 3 = segmental incorporation; 4 = universalistic incorporation.
RATES			
	49 plunits emig. rate	EMIG_RAT	%: 100* (number of plural units that sought or achieved emigration – Smith's var. 101)/(years independent – Smith's var. 66).
	50 plunits relocated rate	RELOC_RA	log 10 (1+(number of plural units to have been internally relocated – Smith's var. 102)).
	51 plunits deport. rate	DEP_RATE	log 10 (number of plural units deported – Smith's var. 103)/(years independent – Smith's var. 66)).
POPULATION PERCENTS			
40	52 plunit %	PLUNIT_	% population in largest plural unit.
41	53 ethnic %	ETHNIC_	% population in largest ethnic unit.
42	54 race %	RACE_	% population in largest racial unit.

Smith #	My # Variable Name*	Code	Description/Comments
TYPES			
	55 hierarchic society	HIERARCH	Hierarchic = 1, not = 0; (from Smith's var. 38, with reference to his var. 39).
	56 plural society	PLURALIS	Plural = 2, cultural only = 1, not = 0; (from Smith's var. 38, with reference to his var. 39).
	57 segmentation	SEGMENTE	Segmented = 1, not = 0; (from Smith's var. 38, with reference to his var. 39).
182	58 types of pluralism	PLUR_TYP	1 = not plural; 2 = predominant segmental pluralism; 3 = predominant hierarchic pluralism; 4 = complex pluralism.
POLITICS			
STATE/NATION			
48	59 new nation	NEW	For 1982: 1 = recently emergent nation (Smith code 2); 0 = not (any other code).
65	60 state age	STATE_AGE	Age of state since 1/1/32 and to 12/82.
66	61 independence age	INDEP_AG	Years independent since 1/1/32 to 12/82.
	62 boundary change rate	BOUND_CHG	(Number of boundary changes – Smith's var. 79)/(years independent – Smith's var. 66).
GOVERNMENT NATURE			
53	63 de facto centralized	CENTRALI	For 1982: 1 = decentralized (Smith's codes 3, 4); 2 = centralized (Smith's 2); 3 = strongly centralized (Smith's code 1). (Smith's 5, 6 treated as missing).
54	64 de facto localism	LOCALISM	For 1982: 1 = Marxist (Smith's 1); 2 = appointed mil. officials (Smith's 4); 3 = appointed civilian (3); 4 = elected local councils without power (2); 5 = same with power (1). (Smith's 6, 7 treated as missing).
	65 secular govt	SEC_GOV	1 = secular govt. (Smith's var. 57A, codes 18, 19, 20); 0 = not (Smith's other codes).
	66 Christian govt	CHRIST_GOV	1 = Christian govt. (Smith's var. 57A, codes 7–9); 0 = not (Smith's other codes).
	67 Muslim govt	MUS_GOV	1 = Muslim govt. (Smith's var. 57A, codes 1–5); 0 = not (Smith's other codes).
59	68 military govt	MIL_GOV	For 1982: 1 = civilian govt. (Smith's 1); 2 = mixed (3); 3 = military (2). (Smith's 5 treated as missing).
	69 presidential govt	PRES_GOV	de facto for 1982: 1 = presidential govt. (Smith's var. 60, code 3); 0 = not (all other codes).
	70 ministerial govt	PM_GOV	de facto for 1982: 1 = prime minister (Smith's var. 60, code 4); 0 = not (all other codes).
GOVERNMENT OUTPUT			
55	71 de facto welfare	WELFARE	For 1982: 1 = no social welfare (Smith's 2); 2 = limited (3); 3 = yes (1).
	72 pol. discrimination	POL_DISC	(Political discrimination (Smith's var. 145))* (political discrimination intensity (Smith's var. 146)) = log 10(1 + (Smith's var. 145.* Smith's var. 146)).
IDEOLOGY			
56	73 socialist	SOCIALIST	1 = govt. ideology is socialist (Smith's 1–6); not = 0 (other codes).
62	74 de facto totalitarian	TOTAL	For 1982: 1 = democracy (Smith's 5, 6); 2 = authoritarian (1, 2, 7–12); 3 = totalitarian = Marxist (3, 4).
	75 pol. and civil rights	FREEDOM	(8 – political rights scores (Smith's var. 143)) + (8 – civil rights score (Smith's var. 144)).
CHARACTERISTICS			
61	76 elected leadership	ELECT	For 1982: 1 = direct or indirect election of top leadership (Smith's 3, 4, 16, 17); 0 = other codes.
63	77 legislature or not	LEGISLAT	For 1982: 0 = no legislature (Smith's 8); 1 = legislature (all other codes).
64	78 party system	PARTIES	For 1982: 0 = no party (Smith's 5, 7); 1 = one party (1, 2, 6); 2 = multiparties (3, 4).
68	79 regime duration	REG_DUR	Years of current regime at 12/82. Regime apparently means type of governing/ideology – not leadership (see Smith's var. 78b).
8B	80 government's duration	GOV_DUR	Duration of current government at 12/31/82 – see Smith's var. 68 (govt. apparently means current leadership.

Smith #	My # Variable Name*	Code	Description/Comments
CHARACTERIZATION			
	81 stability	STABILIT	Rate of regime change: (years independent – Smith's var. 66)/(1 + (no. of regime changes 1932–82 – Smith's var. 67)).
ECONOMIC			
NATURE			
	82 economy	ECON	1 = command economy (Smith's var. 29, code 5, 60; 2 = mixed (code 2, 4); 3 = free mkt (code 1, 3). Identical to Smith's var. 186, except for missing data on the latter.
30	83 development status	DEV	UN economic category; 1 = low; 2 = LDC; 3 = MDC; 4 = high.
31A	84 economic autonomy	ECON_AUTO	de facto.
OUTPUT			
14	85 GDP pc	GDP_PC	Gross domestic product per capita; 19 missing data replaced with GNP pc (Smith's var. 15).
17	86 GDP growth %	GDP_GWTH	% annual GDP growth; 47 cases of missing data replaced with % GNP growth.
31B	87 Life quality	LIFE_QUAL	PQLI 1977 (physical quality of life index).
STRUCTURE			
9	88 agricultural LF %	AGR_LF_	% agricultural labor force.
12	89 agricultural ratio	AGR_RATIO	Ratio of labor force in agriculture to that in industry and services.
EQUALITY			
20	90 top 5% share %	TOP_5_	% income share of top 5% of population.
22	91 bot 20% share %	BOT_20_	% income share of bottom 20% of population.
26	92 mid 40% share %	MID_40_	% income share of middle 40% of population.
	93 econ. discrimination	ECON_DISC	(Economic discrimination (Smith's var. 147))* (economic discrimination intensity (Smith's var. 148)) = log 10 (1 + (Smith's var. 147* Smith's var. 148)).
157	94 inc. inequality ratio	INEQUALITY	(Income share of top 5% – Smith's var. 20)/(share of bottom 20% – Smith's var. 22).
CULTURE			
REGIONAL			
	95 Europe	EUROPE	Smith's var. 2, code 1: Europe = 1, not = 0.
	96 N. Africa & Middle E.	N. AF & ME	Smith's var. 2, codes 4–5: N. AF & ME = 1, not = 0.
	97 Africa	AFRICA	Smith's var. 2, codes 12–17: Africa = 1, not = 0.
	98 Asia	ASIA	Smith's var. 2, codes 7–9: Asia = 1, not = 0.
	99 Central/S. America	LATIN_AM	Smith's var. 2, codes 19–20: Latin American = 1, not = 0. Includes Mexico.
RELIGION			
	100 Christian	CHRIST	Christian = 1; not = 0. From Smith's var. 37A and 37B.
	101 Moslem	MOSLEM	Animist = 1; not = 0. From Smith's var. 37A and 37B.
	102 Animist	ANIMIST	Moslem = 1; not = 0. From Smith's var. 37A and 37B.
SIZE, DEMOGRAPHIC, EDUCATION, HEALTH			
3	103 area	AREA	Log 10 (sq. km).
5B	104 life expectancy	LIVE_EXP	Years.
4	105 population	POP	Log 10 (millions).
13	106 population growth	POP_GWTH	Per year.
5A	107 density	DENSITY	People per sq. km.
6	108 urban %	URBAN	% urban population.
7	109 literate %	LIT	% population literate.

* Plunit means plural unit

Variable Code Alphabetical List

97 Africa	AFRICA	64 de facto localism	LOCALISM
88 agricultural LF %	AGR_LF_	92 mid 40% share %	MID_40_
89 agricultural ratio	AGR_RATIO	68 military govt	MIL_GOV
102 Animist	ANIMIST	15 police/mil. action years	MIL_POL_
103 area	AREA	45 minority dominance	MIN_DOM

Variable Code Alphabetical List

98 Asia	ASIA	101 Moslem	MOSLEM
16 armed attack years	ATTACK_Y	67 Muslim govt	MUS_GOV
91 bot 20% share %	BOT_20_	96 N. Africa & Middle E.	N. AF & ME
62 boundary change rate	BOUND_CHG	33 number of native races	NAT_RACE
63 de facto centralized	CENTRALI	59 new nation	NEW
100 Christian	CHRIST	78 party system	PARTIES
66 Christian govt	CHRIST_GOV	22 peaceful change rate	PEACE_CH
1 coups	COUP	36 plurality	PLUNITS
18 coups per year	COUPS_RA	52 plunit	PLUNIT_
38 cultures	CULTURES	56 plural society	PLURALIS
4 political deaths	DEATHS	58 type of pluralism	PLUR_TYP
25 democide	DEMOCIDE	44 plunits deported	PLU_DEP
107 density	DENSITY	42 plunits emigration	PLU_EMIG
51 plunits deport. rate	DEP_RATE	43 plunits relocated	PLU_RELO
83 development status	DEV	29 plural violence	PLU_VIO
46 dominant plunit	DOM_PLUN	30 plural violence rate	PLU_VIO_
82 economy	ECON	70 ministerial govt	PM_GOV
84 economic autonomy	ECON_AUTO	17 political killing years	POLKILL_
93 econ. discrimination	ECON_DISC	72 pol. discrimination	POL_DISC
76 elected leadership	ELECT	14 police/mil. viol. years	POL_YRS_
49 plunits emig. rate	EMIG_RAT	105 population	POP
35 number of ethnic groups	ETHNICS	106 population growth	POP_GWTH
53 ethnic %	ETHNIC_	69 presidential govt	PRES_GOV
32 ethnic violence rate	ETHNIC_V	3 purges	PURGES
37 ethnolinguistic	ETHNOLIN	10 purges per year	PUR-RAT1
31 ethnic violence	ETH_VIO	11 purge years rate	PUR-RAT2
95 Europe	EUROPE	54 race %	RACE_
9 external war years	EXTWAR_Y	41 regions	REGION
34 number of foreign races	FOR_RACE	79 regime duration	REG_DUR
75 pol. and civil rights	FREEDOM	40 religions	RELIGION
86 GDP growth %	GDP_GWTH	50 plunits relocated rate	RELOC_RA
85 GDP pc	GDP_PC	2 revolts	REVOLTS
80 government's duration	GOV_DUR	19 revolts per year	REV_RATE
55 hierarchic society	HIERARCH	6 revolt years	REV_YRS
23 imposed change rate	IMP_CHG	12 riot years	RIOTS_YR
48 de facto incorporation	INCORP	65 secular govt	SEC_GOV
61 independence age	INDEP_AG	57 segmentation	SEGMENTE
7 years of indep. war	INDWAR_Y	47 potential separatism	SEPARATE
94 inc. inequality ratio	INEQUALITY	73 socialist	SOCIALIST
20 insurr. per year	INS_RATE	81 stability	STABILIT
24 intervention years	INTERVEN	60 state age	STATE_AGE
21 internal war rate	INTWAR_R	13 terrorism years	TERR_YR
8 internal war years	INTWAR_Y	90 top 5% share %	TOP_5_
39 languages	LANGUAGE	74 de facto totalitarian	TOTAL
99 Central/S. America	LATIN_AM	108 urban %	URBAN
77 legislature or not	LEGISLAT	26 violence	VIOLENCE
87 Life quality	LIFE_QUAL	5 rate of violence	VIOL_RATE
109 literate %	LIT	27 violent independence	VIO_IND
104 life expectancy	LIVE_EXP	28 violence intensity	VIO_INTE
		71 de facto welfare	WELFARE

RUDOLPH J. RUMMEL, b. 1932, PhD in Political Science (Northwestern University, 1963); Professor Emeritus of Political Science, University of Hawaii (1995–). Most recent books: *Death By Government* (Transaction, 1994), *Statistics of Democide: Genocide and Mass Murder Since 1900* (Center for National Security Law, University of Virginia, 1996), *Power Kills* (Transaction, 1996).