Ethnic Polarization, Potential Conflict, and Civil Wars

American Economic Review (2005)

Jose G. Montalvo   Marta Reynal-Querol

October 6, 2014

Introduction

- Many studies on ethnic diversity and its effects on conflicts/civil wars.
- Political scientists emphasize *institutions* for the dynamics of conflict.
- Economists have connected ethnic diversity with economic phenomena like investment, growth, etc.
- Empirical work: Include as a regressor in empirical growth estimations some index of *ethnic fractionalization*.
- Why? At least three reasons...
  1. Ethnically diverse socieies have more civil wars which lower investment/growth.
  2. Ethnic diversity may increase corruption and lower investment/growth.
  3. Diffusion of technology more difficult in ethnically diverse socities.

Relation between ethnic fractionalization and conflict

- NO relation! Why?
  1. Misclassification of ethnic groups in the Atlas Nadorov Mira? Others have used alternative sources but corr very high (over 0.8).
  2. Measure "ethnic distance" across groups. Again, corr with original ELF is around 0.82.
  3. Perhaps fractionalization is NOT the right concept/measure?
- The third alternative is the one examined in this paper.
- Horowitz (1985): There is less violence in highly homogeneous and highly heterogeneous societies, and more conflict in societies where a large ethnic minority faces an ethnic majority.
- So, *polarization* should capture the likelihood of conflicts better than *fractionalization*.

- The Ethnolinguistic Fractionalization index (ELF):
  $FRAC = 1 - \sum_{i=1}^{N} \pi_i^2 = \sum_{i=1}^{N} \pi_i(1 - \pi_i)$
- *Interpretation*: prob that two randomly selected individuals will belong to different groups.
- A measure of ethnic polarization:
  $RQ = 4\sum_{i=1}^{N} \pi_i^2(1 - \pi_i)$
- Based on the Esteban and Ray measure of polarization.
- How far is the distribution from the $(1/2, 0, ..., 0, 1/2)$ distribution (bipolar) which represents maximum polarization?

From Income to Ethnicity

- From Income Inequality to Ethnic Fractionalization: Gini and ELF
- From Income Polarization to Discrete Polarization and the RQ Index.
- The dichotomous nature (belong/do not belong) of distance across groups has important implications.
- The RQ index is precisely the index $DP(1, 4)$, i.e., $\alpha = 1, k = 4$.

The Empirical Relationship between Ethnic Fractionalization and Polarization

- The difference between ethnic polarization and fractionalization is BOTH theoretical and actual.

- There are basically three sources of ethnolinguistic diversity across countries:
  1. The World Christian Encyclopedia (WCE)
  2. The Encyclopedia Britannica (EB)
  3. The Atlas Nadorov Mira

- There are also several sources of data on religious diversity.
  1. The World Christian Encyclopedia (WCE)
  2. LEtat des religions dans le monde (ET), which is based on a combination of national data sources and the WCE, provides information on the proportions of followers of Animist and Syncretic cults. Used in this paper.

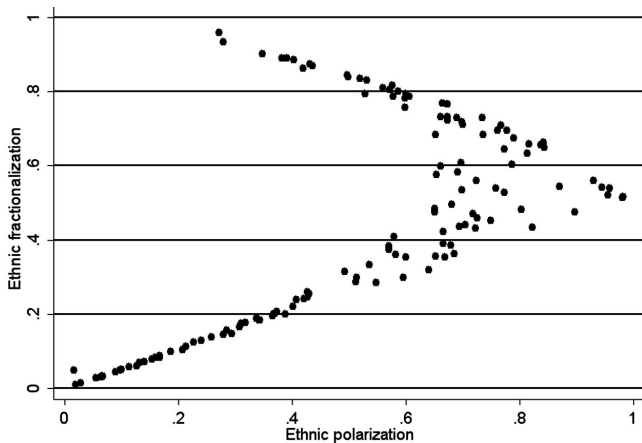# Are Empirical Measures of Ethnic Polarization and Fractionalization Very Different?

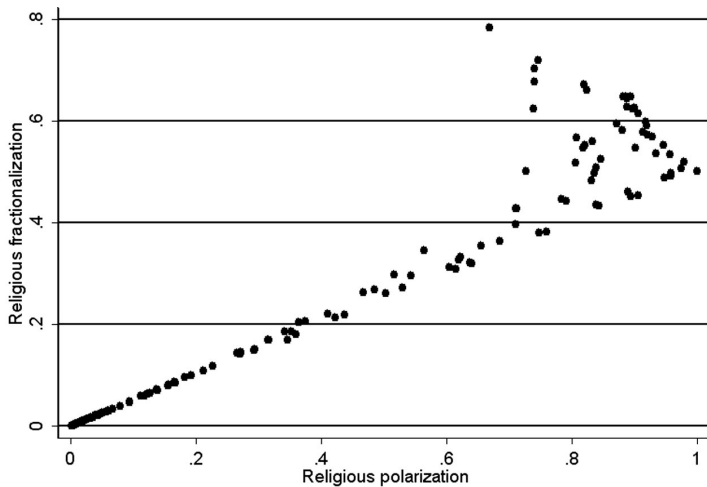FIGURE 1. ETHNIC FRACTIONALIZATION VERSUS POLARIZATION

Source: WCE.

Religion

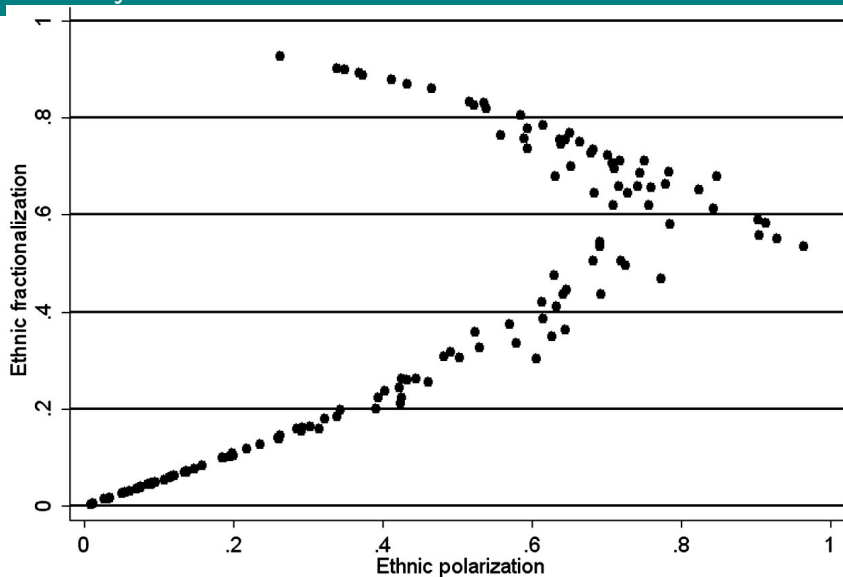FIGURE 2. RELIGIOUS FRACTIONALIZATION VERSUS POLARIZATION

Source: ET.

FIGURE 3. ETHNIC FRACTIONALIZATION VERSUS POLARIZATION

Source: ANM.

Regression Analysis

- Estimate a logit model for the incidence of civil wars as a function of polarization and fractionalization measures of ethnic and religious heterogeneity.
- The sample includes 138 countries during the 19601999 period. Sample grouped into 5-year periods.
- The endogenous variable is the incidence of a civil war from the Peace Research Institute of Oslo (PRIO) dataset.
- Basic variable corresponds to the definition of intermediate and high-intensity civil wars of PRIO.
- PRIO defines an intermediate and high-intensity armed conflict as a contested incompatibility that concerns government and/or territory, where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 yearly battle-related deaths and a minimum of 1,000 during the course of the civil war.

TABLE 1—LOGIT REGRESSIONS FOR THE INCIDENCE OF CIVIL WARS: BASIC INDICATORS OF ETHNOLINGUISTIC/RELIGIOUS HETEROGENEITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	−5.82	−6.26	−6.29	−5.27	−6.03	−6.89	−6.77	−7.47
	(2.06)	(1.93)	(2.01)	(1.66)	(1.85)	(2.26)	(1.94)	(2.32)
LGDPC	−0.28	−0.44	−0.42	−0.40	−0.32	−0.33	−0.37	−0.37
	(1.27)	(1.99)	(1.79)	(1.44)	(1.11)	(1.13)	(1.32)	(1.33)
LPOP	0.34	0.41	0.40	0.40	0.39	0.43	0.40	0.43
	(2.18)	(2.40)	(2.21)	(2.47)	(2.39)	(3.01)	(2.31)	(2.72)
PRIMEXP	−0.90	−1.01	−1.07	−0.36	−0.56	−0.35	−1.21	−0.89
	(0.52)	(0.54)	(0.57)	(0.21)	(0.32)	(0.21)	(0.64)	(0.48)
MOUNTAINS	0.00	0.00	−0.00	0.00	0.00	0.00	−0.00	−0.00
	(0.49)	(0.25)	(0.19)	(0.36)	(0.41)	(0.29)	(0.15)	(0.16)
NONCONT	0.08	0.29	0.28	0.04	0.08	0.31	0.32	0.47
	(0.13)	(0.49)	(0.48)	(0.07)	(0.13)	(0.49)	(0.52)	(0.79)
DEMOCRACY	0.07	0.03	0.03	0.10	0.10	0.01	0.03	−0.03
	(0.21)	(0.09)	(0.09)	(0.29)	(0.28)	(0.05)	(0.08)	(0.09)
ETHFRAC	1.19		0.17					0.04
	(1.89)		(0.19)					(0.05)
ETHPOL		2.37	2.28				2.27	2.09
		(2.97)	(2.23)				(2.84)	(2.03)
RELFRAC				0.37		−4.97		−4.45
				(0.36)		(1.65)		(1.39)
RELPOL					0.73	3.90	0.44	3.29
					(1.00)	(1.97)	(0.65)	(1.59)
Pseudo R^2	0.10	0.12	0.12	0.09	0.10	0.11	0.12	0.13
N	846	846	846	846	846	846	846	846

Notes: The sample includes 138 countries for the period 1960–1999. The dependent variable is the incidence of civil wars following the definition of PRIO, which includes intermediate and high-intensity armed conflicts (PRIOCW). The method of estimation is logit. The absolute z-statistics in parentheses are calculated using standard errors adjusted for clustering. Explanatory variables: LGDPC, log of real GDP per capita in the initial year; LPOP, the log of the population at the beginning of the period; PRMEXP, primary exports (Collier and Hoeffler); MOUNTAINS, mountains; NONCONT, noncontiguous states; DEMOCRACY, degree of democracy (Polity IV); ETHFRAC, ethnic fractionalization (Source: WCE); ETHPOL, ethnic polarization (Source: WCE); RELFRAC, religious fractionalization (Source: ET); RELPOL, religious polarization (Source: ET).

TABLE 2—LOGIT REGRESSIONS FOR THE INCIDENCE OF CIVIL WARS: ROBUSTNESS TO ALTERNATIVE INDICATORS OF ETHNOLINGUISTIC HETEROGENEITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
C	−6.29	−4.82	−6.37	−5.07	−6.22	−5.10	−6.41
	(2.01)	(1.59)	(2.03)	(1.74)	(1.93)	(1.70)	(1.96)
LGDPC	−0.42	−0.49	−0.42	−0.40	−0.43	−0.49	−0.41
	(1.79)	(2.35)	(1.94)	(1.85)	(1.95)	(2.15)	(1.76)
LPOP	0.40	0.40	0.41	0.40	0.40	0.42	0.39
	(2.21)	(2.46)	(2.43)	(2.40)	(2.29)	(2.64)	(2.32)
PRIMEXP	−1.07	−0.17	−1.11	1.19	−0.52	−0.20	−1.25
	(0.57)	(0.10)	(0.60)	(0.50)	(0.18)	(0.11)	(0.69)
MOUNTAINS	−0.00	0.00	−0.00	0.00	−0.00	0.00	−0.00
	(0.19)	(0.03)	(0.21)	(0.38)	(0.22)	(0.11)	(0.26)
NONCONT	0.28	0.22	0.26	0.03	0.28	0.18	0.28
	(0.48)	(0.37)	(0.46)	(0.06)	(0.49)	(0.30)	(0.46)
DEMOCRACY	0.03	0.06	0.04	0.09	0.03	0.07	0.03
	(0.09)	(0.18)	(0.11)	(0.25)	(0.10)	(0.22)	(0.08)
ETHPOL	2.28		2.54		2.35		2.91
	(2.23)		(2.79)		(2.82)		(2.62)
ETHFRAC	0.17						
	(0.19)						
ETHDOM		0.44	−0.14				
		(1.16)	(0.34)				
ETHLRG*PRIMEXP				−2.92	−0.98		
				(0.78)	(0.21)		
LARMINOR						2.22	−1.36
						(1.32)	(0.61)
N	846	846	846	846	846	846	846
Pseudo R^2	0.12	0.09	0.12	0.09	0.13	0.10	0.12

Notes: The sample includes 138 countries for the period 1960–1999. The dependent variable is the incidence of civil wars following the definition of PRIO, which includes intermediate and high-intensity armed conflicts (PRIOCW). The method of estimation is logit. The absolute *z*-statistics in parentheses are calculated using standard errors adjusted for clustering. Explanatory variables: LGDPC, log of real GDP per capita in the initial year; LPOP, the log of the population at the beginning of the period; PRMEXP, primary exports (Collier and Hoeffler); MOUNTAINS, mountains; NONCONT, noncontiguous states; DEMOCRACY, degree of democracy (Polity IV); ETHFRAC, ethnic fractionalization (Source: WCE); ETHPOL, ethnic polarization (Source: WCE); ETHDOM, ethnic dominance (Source: WCE); ETHLARG*PRIMEXP, largest ethnic group by primary exports; LARMINOR, size of the largest minority (Source: WCE).

TABLE 4—LOGIT REGRESSIONS FOR THE INCIDENCE OF CIVIL WARS: COMPARING ALTERNATIVE DATA ON CIVIL WARS

Dependent variable	DSCW (1)	FLCW (2)	PRIO1000 (3)	PRIO25 (4)	PRIOCW (5)	DSCW (6)	FLCW (7)	PRIO1000 (8)	PRIO25 (9)
Constant	−1.84	−2.47	−4.32	−4.16	−7.29	−3.26	−4.25	−5.22	−5.69
	(0.76)	(0.89)	(1.54)	(1.61)	(2.59)	(1.46)	(1.67)	(2.02)	(2.44)
LGDPC	−0.86	−1.19	−0.62	−0.62	−0.45	−0.82	−1.08	−0.66	−0.60
	(3.89)	(4.65)	(2.78)	(3.28)	(1.93)	(3.63)	(4.52)	(3.22)	(3.14)
LPOP	0.35	0.51	0.30	0.39	0.47	0.44	0.58	0.38	0.48
	(2.36)	(2.68)	(1.94)	(2.35)	(2.77)	(3.04)	(3.22)	(2.56)	(3.05)
PRIMEXP	−0.91	−0.55	−0.01	0.24	−0.87	−0.68	−0.37	0.04	0.27
	(0.54)	(0.37)	(0.01)	(0.20)	(0.49)	(0.45)	(0.27)	(0.03)	(0.26)
MOUNTAINS	−0.00	0.00	0.00	0.00	−0.00	−0.00	0.00	0.00	0.00
	(0.80)	(1.04)	(0.54)	(0.22)	(0.20)	(0.63)	(1.03)	(0.50)	(0.18)
NONCONT	0.25	0.90	0.30	0.69	0.16	0.18	0.82	0.13	0.54
	(0.45)	(1.59)	(0.50)	(1.59)	(0.29)	(0.35)	(1.64)	(0.24)	(1.37)
DEMOCRACY	0.43	0.53	0.03	0.18					
	(1.25)	(1.65)	(0.09)	(0.68)					
ETHFRAC	−0.52	0.01	0.57	−0.06	0.18	−0.73	−0.14	0.57	−0.17
	(0.65)	(0.01)	(0.62)	(0.09)	(0.20)	(0.92)	(0.15)	(0.63)	(0.23)
ETHPOL	2.31	1.95	2.33	2.05	2.31	2.32	2.11	2.35	2.13
	(2.76)	(1.97)	(2.16)	(2.41)	(2.23)	(2.74)	(2.02)	(2.12)	(2.48)
Pseudo R^2	0.13	0.25	0.13	0.14	0.15	0.16	0.26	0.16	0.17
N	846	846	846	846	990	990	990	990	990

Notes: The sample includes 138 countries for the period 1960–1999. The method of estimation is logit. The absolute *z*-statistics in parentheses are calculated using standard errors adjusted for clustering. The endogenous variables are: PRIOCW, intermediate and high-intensity armed conflict (PRIO); DSCW, Doyle and Sambanis (2000) definition of civil war; FLCW, Fearon and Laitin (2003) definition of civil war; PRIO1000, armed conflict generating more than 1,000 deaths yearly (PRIO); PRIO25, armed conflict generating more than 25 deaths yearly (PRIO). Explanatory variables: LGDPC, log of real GDP per capita in the initial year; LPOP, the log of the population at the beginning of the period; PRMEXP, primary exports (Collier and Hoeffler); MOUNTAINS, mountains; NONCONT, noncontiguous states; DEMOCRACY, degree of democracy (Polity IV); ETHFRAC, ethnic fractionalization (Source: WCE); ETHPOL, ethnic polarization (Source: WCE).

TABLE 5—ROBUSTNESS OF THE RESULTS TO THE INCLUSION OF REGIONAL DUMMIES AND THE ELIMINATION OF COUNTRIES IN SPECIFIC REGIONS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
C	−6.17	−6.07	−4.59	−4.23	−7.47	−7.60	−5.50	−5.39
	(1.74)	(1.68)	(1.19)	(1.06)	(1.93)	(2.01)	(1.69)	(1.66)
LGDPC	−0.43	−0.41	−0.45	−0.43	−0.37	−0.34	−0.40	−0.43
	(1.84)	(1.71)	(1.78)	(1.57)	(1.59)	(1.43)	(1.72)	(1.78)
LPOP	0.40	0.38	0.33	0.29	0.46	0.45	0.33	0.35
	(2.34)	(2.08)	(1.75)	(1.37)	(2.35)	(2.24)	(2.01)	(1.93)
PRIMEXP	−1.08	−1.15	−0.94	−1.14	−0.55	−0.60	−0.92	−0.81
	(0.56)	(0.59)	(0.45)	(0.55)	(0.30)	(0.33)	(0.43)	(0.37)
MOUNTAINS	−0.00	−0.00	−0.01	−0.01	−0.00	−0.00	−0.00	−0.00
	(0.15)	(0.11)	(0.64)	(0.65)	(0.57)	(0.50)	(0.02)	(0.09)
NONCONT	0.11	0.09	0.12	0.05	−0.06	−0.07	0.12	0.14
	(0.17)	(0.14)	(0.20)	(0.09)	(0.09)	(0.11)	(0.15)	(0.17)
DEMOCRACY	0.09	0.09	0.01	−0.00	0.06	0.06	0.16	0.16
	(0.25)	(0.26)	(0.02)	(0.02)	(0.16)	(0.16)	(0.43)	(0.44)
ETHPOL	2.48	2.35	2.40	1.98	2.20	2.12	2.46	2.61
	(3.10)	(2.23)	(2.60)	(1.32)	(2.80)	(2.15)	(2.92)	(2.43)
ETHFRAC		0.26		0.63		0.19		−0.30
		(0.29)		(0.42)		(0.21)		(0.32)
Reg. Dummies	Yes	Yes	No	No	No	No	No	No
Eliminated region	None	None	SAfrica	SAfrica	Laam	Laam	Asiae	Asiae
N	846	846	580	580	678	678	781	781
Pseudo R^2	0.13	0.13	0.12	0.12	0.12	0.12	0.11	0.11

Notes: The sample includes 138 countries for the period 1960–1999. The dependent variable is the incidence of civil wars following the definition of PRIO, which includes intermediate and high-intensity armed conflicts (PRIOCW). The method of estimation is logit. The absolute *z*-statistics in parentheses are calculated using standard errors adjusted for clustering. Explanatory variables: LGDPC, log of real GDP per capita in the initial year; LPOP, the log of the population at the beginning of the period; PRMEXP, primary exports (Collier and Hoeffler); MOUNTAINS, mountains; NONCONT, noncontiguous states; DEMOCRACY, degree of democracy (Polity IV); ETHFRAC, ethnic fractionalization (Source: WCE); ETHPOL, ethnic polarization (Source: WCE). Regional dummies: SAFRICA, sub-Saharan Africa; LAAM, Latin America; ASIAE, Asia.

TABLE 6—ROBUSTNESS REGRESSIONS: CROSS-SECTION LOGIT REGRESSIONS FOR THE INCIDENCE OF CIVIL WARS

Endogenous variable	PRIOCW (1)	PRIOCW (2)	DSCW (3)	DSCW (4)	FLCW (5)	FLCW (6)
C	−1.19	−1.04	2.23	5.84	4.91	8.37
	(0.37)	(0.26)	(0.68)	(1.31)	(1.41)	(1.76)
LGDPC	−0.63	−0.63	−1.01	−1.40	−1.23	−1.64
	(2.03)	(1.61)	(2.95)	(2.94)	(3.34)	(3.20)
LPOP	0.35	0.34	0.37	0.34	0.25	0.24
	(1.61)	(1.51)	(1.57)	(1.33)	(1.06)	(0.94)
PRIMEXP	1.19	1.29	−0.34	0.23	−0.20	0.90
	(0.55)	(0.57)	(0.15)	(0.10)	(0.08)	(0.36)
MOUNTAINS	−0.00	−0.01	0.00	−0.01	0.01	0.01
	(0.58)	(0.45)	(0.42)	(0.83)	(0.70)	(0.46)
NONCONT	0.02	0.05	−0.53	−0.56	−0.17	0.14
	(0.03)	(0.06)	(0.59)	(0.57)	(0.19)	(0.14)
DEMOCRACY	0.32	0.35	−0.02	−0.09	−0.19	−0.03
	(0.55)	(0.59)	(0.04)	(0.14)	(0.30)	(0.06)
ETHPOL	3.35	3.42	3.26	3.53	2.95	3.27
	(2.46)	(2.48)	(2.37)	(2.44)	(2.15)	(2.26)
Reg. dummies	No	Yes	No	Yes	No	Yes
N	90	90	90	90	90	90
Pseudo R^2	0.17	0.17	0.26	0.27	0.29	0.31

Notes: The sample includes 138 countries for the period 1960–1999. The method of estimation is logit. The endogenous variables are: PRIOCW, intermediate and high-intensity definition of armed conflict of PRIO; DSCW, Doyle and Sambanis (2000) definition of civil war; FLCW, Fearon and Laitin (2003) definition of civil war. Explanatory variables: LGDPC, log of real GDP per capita in the initial year; LPOP, the log of the population at the beginning of the period; PRIMEXP, primary exports (Collier and Hoeffler); MOUNTAINS, mountains; NONCONT, noncontiguous states; DEMOCRACY, degree of democracy (Polity IV); ETHPOL, ethnic polarization (Source: WCE). Regional dummies: SAFRICA, sub-Saharan Africa; LAAM, Latin America; ASIAE, Asia.

Conclusion

- This paper argues that the concept of ethnic heterogeneity appropriate to capture potential conflict should be polarization.
- The index developed here is related to the original index of income polarization of Esteban and Ray (1994).
- In this paper, it is shown that the difference between ethnic polarization and fractionalization is *both* theoretical and actual.
- In the empirical section, they show that ethnic fractionalization does not have a significant effect on the likelihood of conflicts.
- In contrast, ethnic polarization is a significant explanatory variable for the incidence of civil wars.
- This result is robust to the use of other proxies for ethnic heterogeneity, alternative sources of data, regional dummies, and the use of a single cross section of data.