

Education and Conflict

Evidence from a Policy Experiment in Indonesia

DOMINIC ROHNER¹² ALESSANDRO SAIA²

¹ University of Lausanne and CEPR ² University of Bologna

ABCDE, June 2021

Motivation

Education and conflict

- What policies can curb conflict ?
- Widespread believe that education could foster peace and stability held by journalists and NGOs alike ...
- ... but education may also boost revolt

A still open question

- Lack of causal evidence
- Channels and mechanisms are unexplored

What we do

- We exploit the INPRES school construction program in Indonesia taking place from 1974 to 1978
 - Has been studied by several papers in the past (e.g. by Duflo, 2001), but never related to conflict
 - We carry out a Difference-in-Difference analysis
- We had to collect our own conflict data, drawing on information from over 820,000 newspaper pages
 - Our panel dataset contains 289 districts (*Kabupate*) over 40 years, resulting in 11,560 observations.
- We study potential channels of transmission
 - Using census and survey data from various sources ...
 - ... we are able to distinguish economic mechanisms and channels linked to religious cleavages

What we find

- Finding #1. **Education reduces the likelihood of conflict**
 - Findings robust to a wide array of robustness checks
 - Quantitatively important, as one school more per 1000 children can several years after the program's end reduce the baseline conflict risk by more than half.
- Finding #2. **The beneficial effect of education increases over time**
- Finding #3. **Both economic and societal channels of transmission matter**
 - Effect of schooling magnified in places with higher returns to schooling and bigger religious polarization
 - Education increases religious tolerance and local community action, but not by lowering religiosity

Related Literature [Short]

- Pooled cross-country regressions find that education correlates negatively with conflict (Collier and Hoeffler, 2004 ; Thyne, 2006).
- Education and individual-level outcomes :
 - Educated people less likely to enlist in armed rebellion (Humphreys and Weinstein (2008) ; Tezcür (2016)).
 - Education also associated to ...
 - ... bigger taste for cultural diversity (Hainmueller and Hiscox (2007) ; Roth and Sumarto (2015)),
 - ... and more civic involvement (see Dee (2004) ; Milligan et al. (2004) ; Wantchekon et al. (2014) ; Campante and Chor (2012, 2014) ; Friedman et al. (2016))
- Other papers studying school construction in Indonesia, but unrelated to conflict : Duflo (2001) ; Duflo (2004) ; Breierova and Duflo (2004) ; Somanathan (2008) ; Roth and Sumarto (2015) ; Ashraf et al. (2016) ; Martinez-Bravo (2017).

Our contributions :

- ① First paper studying causal impact of education on conflict
- ② Provide novel evidence of specific channels of transmission

Historical Context - INPRES Program

- Between 1973-1974 and 1978-1979, more than 61,000 primary schools were built, more than doubling the number of schools in Indonesia.
- Enrollment rates among children aged 7 to 12 increasing from 69 percent in 1973 to 83 percent by 1978
- Designed by the central government
- Stipulating that the number of schooling places to be built had to be roughly proportional to school-aged children not enrolled before the program

Historical Context -Conflict

- A former Dutch colony, Indonesia has won independence in 1949.
- Great ethnic and religious diversity
- During sample period 1955-1994 substantial amount of conflict, with heterogenous causes
 - Separatist rebellions (Aceh, Western Sumatra, North Sulawesi, Darul Islam in West Java, Free Papua Movement, Maluku, Timor Leste)
 - Anti-communist purges
 - Religious / ethnic violence (e.g. riots in Tanjung Priok)

Construction of Conflict Variables from Newspaper Data I

- No existing disaggregate conflict data reaches far enough back
- Main data source Sydney Morning Herald (SMH)
 - Digital archive with unlimited access and reaching far enough back in time
 - Founded in 1831, the SMH is the oldest continuously published newspaper in Australia
 - High quality and readership of roughly half a million people
 - No particular political bias
 - Australian news quite detailed coverage of Indonesia
- Robustness analysis with Canberra Time and with existing conflict datasets (GDEL, ICEWS, NVMS).

Construction of Conflict Variables from Newspaper Data II

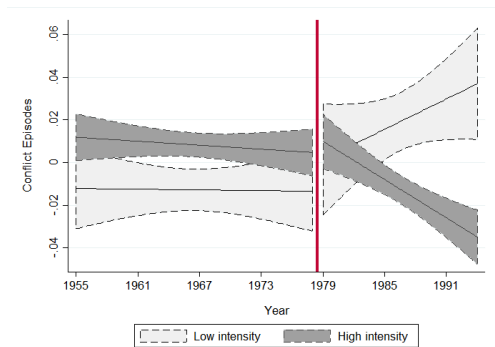
- We searched over 820,000 articles available in SMH archive and downloaded all those containing at least once the word "Indonesia" (the resulting set of articles was of around 34,000).
- Use of natural language processing algorithms to analyse the content of all articles, storing all sentences where at least one conflict related term was present.
- Apply Named Entity Recognition algorithm to identify all locations referred to, and then matched locations to geo-coordinates.
- In 86 percent of cases GDELT and SMH have the same value, while GDELT and ICEWS only have the same value in 71 percent of cases.

▶ Graphically

Identification Strategy and Common pre-trend

We will exploit the differential boost in education across districts and study whether more intense school construction under INPRES has led to a relative decline of conflict with respect to other districts less affected by INPRES

Figure – Intensity of INPRES school construction and conflict events



Regression equation I

The dataset covers 289 districts (*Kabupate* in Indonesian) across 26 provinces over the period 1955-1994.

$$Conflict_{it} = \alpha + \beta \frac{\#SchoolsBuilt}{\#Children}_i * Post1978_t + FE_i + FE_t + \epsilon_{it}.$$

- $Conflict_{it}$ is a dummy for violent event in district i in year t
- $(\#SchoolsBuilt/\#Children)_i$ is number of primary schools constructed under the INPRES program
- Dummy $Post1978_t$ takes a value of 1 for the first year when we expect the program to deplete major effects
- Always district fixed effects (FE_i) and year fixed effects (FE_t) included. In some specifications district-specific linear time trends and Province times year fixed effects (FE_{pt})
- Standard errors are clustered at the level of the 289 districts

Baseline Results I

Table – Main Results

<i>Dep. Variable : Conflict Episode_{it}</i>	(1)	(2)	(3)	(4)	(5)	(6)
# Inpres Schools / N. Children (1.000) * Post-1978	-0.0127*** (0.00448)	-0.0101* (0.00573)	-0.0173*** (0.00610)			
# Schools / N. Children * Years since 1978				-0.00146*** (0.000421)	-0.00175*** (0.000654)	-0.00305*** (0.000705)
Observations	11,560	11,560	11,560	11,560	11,560	11,560
R-squared	0.345	0.450	0.506	0.346	0.450	0.506
Kabupate FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Kabupate Linear Trend	No	Yes	Yes	No	Yes	Yes
Prov * Year FEs	No	No	Yes	No	No	Yes
Mean Dependent Variable	.08	.08	.08	.08	.08	.08
Mean # INPRES Schools per Children	2.35	2.35	2.35	2.35	2.35	2.35

Notes : Dependent variable is a dummy that takes value 1 when a violent event in district i in year t is observed. The variable ($\#SchoolsBuilt/\#Children$) $_i$ represents the number of primary schools constructed under the INPRES program. Dummy $Post1978_i$ takes a value of 1 for all years after 1978. The variable defined as *Years since 1978* until 1978 takes value 0, in 1979 takes value 1, in 1980 takes value 2, and so on. Standard errors are clustered at the level of the 289 districts (kabupate). *** Significant at the 1% level, ** at the 5% level, * at the 10%.

Baseline Results II

Table – Evolution of coefficients over time

	(1)	(2)	(3)
<i>Dep. Var. : Conflict Event_{it}</i>	OLS	OLS	OLS
# Schools / N. Children [1979-1984]	-0.00658 (0.00427)	-0.00952* (0.00536)	-0.0168*** (0.00591)
# Schools / N. Children [1985-1989]	-0.0111** (0.00540)	-0.0151* (0.00813)	-0.0253*** (0.00837)
# Schools / N. Children [1990-1994]	-0.0218*** (0.00594)	-0.0268*** (0.0101)	-0.0489*** (0.0110)
Observation	11,560	11,560	11,560
R-squared	0.346	0.450	0.507
Kabupate FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Kabupate Linear Trend	No	Yes	Yes
Prov * Year FEs	No	No	Yes
Mean Dependent Variable	.08	.08	.08
Mean # INPRES Schools per Children	2.35	2.35	2.35

Notes : Dependent variable is a dummy that takes value 1 when a violent event in district i in year t is observed. The variable ($\#SchoolsBuilt/\#Children$); represents the number of primary schools constructed under the INPRES program. Standard errors are clustered at the level of the 289 districts (kabupate). *** Significant at the 1% level, ** at the 5% level, * at the 10%.

Channels I

- Proxies for social versus economic factors (de-meanned)
- Social : Religious polarization computed at the district level (see Montalvo and Reynal-Querol, 2005)
- Economic : Compute "returns" on education from survey data (i.e. brick houses / entrepreneurship)

$$RoE[Bricks] = - \frac{\frac{Bricks(NoPS)}{Bricks(NoPS)+NoBricks(NoPS)}}{\frac{Bricks(PS)}{Bricks(PS)+NoBricks(PS)}}$$

Channels II

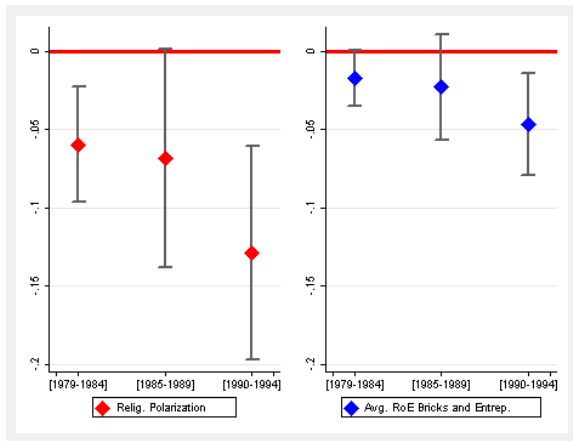
Table – Mechanism : Economic and societal channels

Dep. Variable : Conflict Episode _{it}	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(# Schools / # Children) _{it} * Post-1978 _{it}	-0.0133** (0.00661)	-0.00651 (0.00585)	-0.00996 (0.00621)	-0.00897 (0.00588)				
(# Schools / # Children) _{it} * Years since 1978 _{it}					-0.00312*** (0.000909)	-0.00268*** (0.000720)	-0.00296*** (0.000777)	-0.00286*** (0.000831)
(# Schools / # Children) _{it} * Post-1978 _{it} * Religious Polarization _{it}	-0.0392* (0.0214)	-0.0351* (0.0192)	-0.0473** (0.0211)	-0.0375* (0.0203)				
(# Schools / # Children) _{it} * Post-1978 _{it} * Return to Education [Bricks] _{it}	-0.0236* (0.0121)							
(# Schools / # Children) _{it} * Post-1978 _{it} * Return to Education [Entrep.] _{it}		-0.00111 (0.00474)						
(# Schools / # Children) _{it} * Post-1978 _{it} * Average RoE Bricks and Entrep. _{it}			-0.00633 (0.00868)					
(# Schools / # Children) _{it} * Post-1978 _{it} * Princ. Comp. RoE Bricks and Entrep. _{it}				-0.0116 (0.00821)				
(# Schools / # Children) _{it} * Years since 1978 _{it} * Religious Polarization _{it}					-0.00767*** (0.00277)	-0.00711*** (0.00233)	-0.00799*** (0.00231)	-0.00845*** (0.00257)
(# Schools / # Children) _{it} * Years since 1978 _{it} * Return to Education [Bricks] _{it}					-0.00347** (0.00147)			
(# Schools / # Children) _{it} * Years since 1978 _{it} * Return to Education [Entrep.] _{it}						-0.00103** (0.000494)		
(# Schools / # Children) _{it} * Years since 1978 _{it} * Average RoE Bricks and Entrep. _{it}							-0.00224** (0.000921)	
(# Schools / # Children) _{it} * Years since 1978 _{it} * Princ. Comp. RoE Bricks and Entrep. _{it}								-0.00233*** (0.000850)
Observations	9,480	9,040	10,480	8,040	9,480	9,040	10,480	8,040
R-squared	0.522	0.529	0.517	0.534	0.523	0.530	0.519	0.536
District FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Time Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Mean	.081	.076	.079	.078	.081	.076	.079	.078

NOTE: The unit of observation is a district i and year t . The sample covers 289 districts across 26 provinces over the period 1995-1994. LPM estimates are reported in all columns. The variable Religious Polarization corresponds to the level of religious polarization in district i , whereas the variable Return to Education (RoE) indicates the relative economic advantages at the district level from having completed primary school. Religious polarization and returns to education measures were computed using the 1971 Census. Robust standard error clustered at the district level are reported in parenthesis. Statistical significance is represented by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Channels III

Figure – Channels - Evolution of coefficients over time



How education may ease religious tensions I

- We use wave 5 of IFLS Survey in 2014 on (i) religious tolerance, (ii) community participation and (iii) religiosity.
- Focus on answers provided by individuals born between 1945 and 1972.
- Our identification strategy exploits that date and region of birth jointly determine exposure to INPRES (born before 1962 too old to enroll in new schools)

How education may ease religious tensions III

We estimate the effect of INPRES program using the following econometric specification :

$$Survey_n = \alpha + \beta \frac{\#SchoolsBuilt}{\#Children}_i * Born\ after\ 1962_{nic} + FE_i + FE_c + \epsilon_n,$$

- $(\#SchoolsBuilt/\#Children)_i$ is number of schools constructed under the INPRES program per 1,000 children in district i
- $Born\ after\ 1962$ is a dummy that takes value 1 if the individual n was born after 1962 in district i
- FE_i and FE_c are district and cohort FEs.

How education may ease religious tensions IV

Table – Societal channels : Religious tolerance and local community involvement

<i>Dep. Variable :</i>	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Trust_n</i>	<i>Marriage_n</i>	<i>Roscas_n</i>	<i>Trust_n</i>	<i>Marriage_n</i>	<i>Roscas_n</i>
<i>(# Schools / # Children)_i * Born after 1962_n</i>	0.0344** (0.0140)	0.0321** (0.0127)	0.0189*** (0.00665)	0.0322** (0.0137)	0.0263** (0.0114)	0.0166*** (0.00607)
Observations	10,521	10,522	11,229	10,521	10,522	10,461
R-squared	0.107	0.179	0.154	0.134	0.223	0.237
District FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Gender FEs	No	No	No	Yes	Yes	Yes
Ethnicity FEs	No	No	No	Yes	Yes	Yes
Religion FEs	No	No	No	Yes	Yes	Yes

NOTE : The unit of observation is an individual n born in district i . The sample covers all individuals surveyed in the Wave 5 of the IFLS SURVEY, born between 1945 and 1972. OLS estimates are reported in all columns. $Trust_n$ and $Marriage_n$ variables are used as continuous variables ranging from 0 to 3, treating the scales of the survey questions as cardinal. $Roscas_n$ is a dummy that take a value of 1 if the individual participated to a *arisan* community group over the previous 12 months. Robust standard error clustered at the district level are reported in parenthesis. Statistical significance is represented by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Conclusions

- Novel evidence about the impact of education going beyond cross-country correlations
- We find a statistically significant, quantitatively sizeable impact (one more school per 1,000 children can reduce the baseline conflict risk by more than half at end of period)
- Both social and economic forces seem at work, with returns to schooling taking longer to kick in
- Inter-religious tolerance boosted by education

Thank you for your attention !

Balancing of school construction

[▶ Back](#)

Table – Balancing covariates - Dep. variable : $(\log) \#$ INPRES schools

Dep. Variable : (\log) INPRES Schools;	(1)	(2)	(3)	(4)	(5)	(6)
(\log) Children 5-14;	0.731*** (0.0267)	0.731*** (0.0268)	0.735*** (0.0270)	0.736*** (0.0272)	0.736*** (0.0271)	0.736*** (0.0272)
(\log) School Attendance;	-1.042*** (0.261)	-1.033*** (0.262)	-1.011*** (0.262)	-1.020*** (0.262)	-0.999*** (0.264)	-1.008*** (0.263)
(\log) Enrollment Population;	0.0427 (0.0689)	0.0413 (0.0691)	0.0455 (0.0689)	0.0491 (0.0692)	0.0437 (0.0691)	0.0475 (0.0694)
(\log) Rural Population;	0.0801 (0.171)	0.0688 (0.173)	0.0834 (0.171)	0.0808 (0.171)	0.0700 (0.173)	0.0677 (0.173)
(\log) Primary Industries Employment;	0.0976 (0.197)		0.103 (0.197)	0.105 (0.197)		
(\log) Mining Employment;		-0.267 (0.955)			-0.359 (0.958)	-0.341 (0.958)
(\log) Agricultural Employment;		0.111 (0.199)			0.119 (0.199)	0.120 (0.199)
Dummy Conflict [Pre-1979];			-0.0509 (0.0461)		-0.0528 (0.0463)	
(\log) Years with Conflict [Pre-1979];				-0.0282 (0.0283)		-0.0292 (0.0284)
Observations	289	289	289	289	289	289
R-squared	0.802	0.802	0.803	0.803	0.803	0.803

NOTE : The unit of observation is a district i . The sample covers 289 districts. OLS estimates are reported in all columns. The dependent variable is the (\log) number of primary schools constructed under the INPRES program in a district i . The variable (\log) Children 5-14; represents the number of school-aged children in district i . The variable (\log) Enrollment Population; represents the population-wide pre-INPRES enrollment rates observed in district i . The variable (\log) Rural Population; represents the share of population of district i living in rural areas. The variable (\log) Primary Industries Employment; represents share of population working in primary industries observed in district i . The variable Dummy Conflict [Pre-1979]; is a dummy that takes a value of 1 if a violent event is observed in district i in the period [1955-1979]. The variable (\log) Years with Conflict [Pre-1979]; is the (\log) number of years with conflict episodes observed in district i in the period [1955-1979]. Standard error are reported in parenthesis. Statistical signif. depicted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Control for covariates of school construction

▶ Back

Table – Robustness : Controlling for interactions with socioeconomic covariates

Dep. Variable : Conflict Episode _{<i>i</i>}	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(# Schools / # Children), * Post-1978,	-0.0173*** (0.00610)	-0.0166*** (0.00617)	-0.0160*** (0.00585)	-0.0174*** (0.00611)	-0.0181*** (0.00610)	-0.0192*** (0.00645)	-0.0156** (0.00612)							
(# Schools / # Children), * Years since 1978,								-0.00305*** (0.000705)	-0.00267*** (0.000718)	-0.00349*** (0.000766)	-0.00303*** (0.000707)	-0.00304*** (0.000705)	-0.00327*** (0.000715)	-0.00271*** (0.000810)
School Attendance (5-14), * Post-1978,	No	Yes	No	No	No	No	Yes	No	No	No	No	No	No	No
Enrollment Rate, * Post-1978,	No	No	Yes	No	No	No	Yes	No	No	No	No	No	No	No
Primary Industries, * Post-1978,	No	No	No	Yes	No	No	Yes	No	No	No	No	No	No	No
Rural Population, * Post-1978,	No	No	No	No	Yes	No	Yes	No	No	No	No	No	No	No
Employment Rate, * Post-1978,	No	No	No	No	No	Yes	Yes	No	No	No	No	No	No	No
School Attendance (5-14), * Years since 1978,	No	No	No	No	No	No	No	No	Yes	No	No	No	No	Yes
Enrollment Rate, * Years since 1978,	No	No	No	No	No	No	No	No	No	Yes	No	No	No	Yes
Primary Industries, * Years since 1978,	No	No	No	No	No	No	No	No	No	No	Yes	No	No	Yes
Rural Population, * Years since 1978,	No	No	No	No	No	No	No	No	No	No	No	Yes	No	Yes
Employment Rate, * Years since 1978,	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes
Observations	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560
R-squared	0.506	0.506	0.506	0.506	0.506	0.506	0.506	0.506	0.507	0.507	0.507	0.506	0.507	0.508
District FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Mean	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08

NOTE: The unit of observation is a district i and year t . The sample covers 289 districts across 26 provinces over the period 1995-1994. LPM estimates are reported in all columns. The dependent variable is a dummy that takes a value of 1 if a violent event was observed in district i and year t . The variable (# Schools/# Children) represents the number of primary schools constructed under the INPRES program per 1,000 school-aged children in a district i . The variable School Attendance (5-14), represents the pre-INPRES share of school-aged children enrolled in school in district i . The variable Enrollment Rate, represents the population-wise pre-INPRES enrollment rates observed in district i . The variable Rural Population, represents the share of the population of district i living in rural areas. The variable Primary Industries Employment, represents the share of the population working in primary industries (i.e., Agricultural and Mining Industries) observed in the district i . The variable Employment Rate, represents the share of the population working observed in the district i . All socio-economic variables were computed using the 1971 Census. Robust standard error clustered at the district level are reported in parenthesis. Statistical significance is represented by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Control for water sanitation program

[▶ Back](#)

Table – Robustness Main Results - SMH - with Water Sanitation Program

<i>Dep. Variable : Conflict Episode_{it}</i>	(1)	(2)	(3)	(4)	(5)	(6)
<i>(# Schools / # Children)_i * Post-1978_t</i>	-0.0138** (0.00646)	-0.0120* (0.00683)	-0.0187*** (0.00707)			
<i>(# Schools / # Children)_i * Years Since 1978_t</i>				-0.00162*** (0.000595)	-0.00213*** (0.000801)	-0.00313*** (0.000877)
<i>Intensity Water and Sanitation Program_i * Post-1978_t</i>	0.00435 (0.0156)	0.00783 (0.0137)	0.00593 (0.0193)			
<i>Intensity Water and Sanitation Program_i * Years since 1978_t</i>				0.000655 (0.00155)	0.00152 (0.00162)	0.000349 (0.00215)
Observations	11,560	11,560	11,560	11,560	11,560	11,560
R-squared	0.345	0.450	0.506	0.346	0.450	0.506
District FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trend	No	Yes	Yes	No	Yes	Yes
Province x Year FEs	No	No	Yes	No	No	Yes
Sample Mean	.08	.08	.08	.08	.08	.08

Notes : The variable *Intensity Water and Sanitation Program_i* represents the intensity of a water sanitation program implemented contemporaneously with INPRES in district *i*. Standard errors are clustered at the level of districts (kabupate). *** Significant at the 1% level, ** at the 5% level, * at the 10%.

Control for migration

▶ Back

Table – Robustness : Controlling for time-varying migration and rural population

Dep. Variable : Conflict Episode _{it}	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(# Schools / # Children) _i * Post-1978 _t	-0.0173*** (0.00610)	-0.0174*** (0.00610)	-0.0173*** (0.00611)	-0.0174*** (0.00610)				
(# Schools / # Children) _i * Years Since 1978 _t					-0.00305*** (0.000705)	-0.00306*** (0.000706)	-0.00305*** (0.000708)	-0.00306*** (0.000710)
Observations	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560
R-squared	0.506	0.506	0.506	0.506	0.506	0.506	0.506	0.506
Migration _{it}	No	Yes	No	Yes	No	Yes	No	Yes
Rural Population _{it}	No	No	Yes	Yes	No	No	Yes	Yes
District FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTE : The unit of observation is a district *i* and year *t*. The sample covers 289 districts across 26 provinces over the period 1955-1994. LPM estimates are reported in all columns. The dependent variable is a dummy that takes a value of 1 if a violent event was observed in district *i* and year *t*. *Migrations_{it}* represents the share of population in a district *i* and year *t* having immigrated from another province. *Rural Population_{it}* corresponds to the share of population in a district *i* and year *t* living in rural areas. Time-varying measures of migration and rural population were computed using the Indonesian population census of 1971, 1980 and 1990. Robust standard error clustered at the district level are reported in parenthesis. Statistical significance is represented by * *p* < 0.10, ** *p* < 0.05, *** *p* < 0.01.

Control for climate and oil rents shocks

[▶ Back](#)

Table – Robustness : Controlling for climate and oil rent shocks

Dep. Variable : $Conflict\ Episode_{it}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$(\# Schools / \# Children) * Post-1978$	-0.0173*** (0.00610)	-0.0174*** (0.00609)	-0.0174*** (0.00607)	-0.0187*** (0.00616)				
$(\# Schools / \# Children) * Years\ Since\ 1978$					-0.00305*** (0.000705)	-0.00305*** (0.000704)	-0.00309*** (0.000703)	-0.00294*** (0.000690)
Observations	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560
R-squared	0.506	0.506	0.506	0.506	0.506	0.506	0.507	0.507
$Precipitations_{it}$	No	Yes	Yes	Yes	No	Yes	Yes	Yes
$Temperature_{it}$	No	No	Yes	Yes	No	No	Yes	Yes
$Oil [PRIO-Grid]_i \times Oil\ Prices_t$	No	No	No	Yes	No	No	No	Yes
District FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTE : The unit of observation is a district i and year t . The sample covers 289 districts across 26 provinces over the period 1955-1994. LPM estimates are reported in all columns. The dependent variable is a dummy that takes a value of 1 if a violent event was observed in district i and year t . The variable $(\# Schools / \# Children)_i$ represents the number of primary schools constructed under the INPRES program per 1,000 school-aged children in a district i . Time-varying measures of precipitations and temperature were obtained from the Prio-Grid data. The variable $Oil [Prio-Grid]$ takes a value of 1 if oil has been depleted in district i over the period. World oil prices were retrieved from the BP Statistical Review of World Energy Prices (additional details are provided in the text). Robust standard error clustered at the district level are reported in parenthesis. Statistical significance is represented by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Intensive margin

▶ Back

Table – Robustness : Intensive margin

Dep. Variable :	(1)	(2)	(3)	(4)	(5)	(6)
	$\log(\text{Days}+1)_{it}$	$\log(\text{Weeks}+1)_{it}$	$\log(\text{Months}+1)_{it}$	$\log(\text{Days}+1)_{it}$	$\log(\text{Weeks}+1)_{it}$	$\log(\text{Months}+1)_{it}$
$(\# \text{ Schools} / \# \text{ Children})_i * \text{Post-1978}_t$	-0.0122** (0.0493)	-0.0135** (0.0111)	-0.0139*** (0.00378)			
$(\# \text{ Schools} / \# \text{ Children})_i * \text{Years Since 1978}_t$				-0.00191 (0.211)	-0.00227* (0.0515)	-0.00250*** (0.00148)
Observations	11,560	11,560	11,560	11,560	11,560	11,560
R-squared	0.688	0.681	0.654	0.688	0.681	0.654
District FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes
Province x Year FEs	Yes	Yes	Yes	Yes	Yes	Yes

NOTE : The unit of observation is a district i and year t . The sample covers 289 districts across 26 provinces over the period 1955-1994. OLS estimates are reported in all columns. The dependent variable is defined as the (log) number of days, weeks or months featuring newspaper articles in the *Sydney Morning Herald* referring to conflict events in district i in year t . Robust standard error clustered at the district level are reported in parenthesis. Statistical significance is represented by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Alternative conflict data : GDELT

▶ Back

Table – Results with GDELT

Dep. Variable : Conflict Episode _{it}	All Conflict Events _{it}				Conflicts [18-20] _{it}			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(# Schools / # Children) _i * Years Since 1978 _t	-0.000311 (0.000883)	-0.00170** (0.000716)	-0.000965 (0.000971)	-0.00252*** (0.000676)	-0.000454 (0.000605)	-0.00153** (0.000613)	-0.000919 (0.000677)	-0.00217*** (0.000620)
Observations	4,624	9,826	4,624	9,826	4,624	9,826	4,624	9,826
R-squared	0.353	0.564	0.419	0.616	0.325	0.510	0.388	0.568
District FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Year FEs	No	No	Yes	Yes	No	No	Yes	Yes
Time-Window	1979-1994	1979-2012	1979-1994	1979-2012	1979-1994	1979-2012	1979-1994	1979-2012
Sample Mean	.1	.39	.1	.39	.06	.25	.06	.25

Notes : Standard errors are clustered at the level of districts (kabupate). *** Significant at the 1% level, ** at the 5% level, * at the 10%.

Alternative conflict data : ICEWS

▶ Back

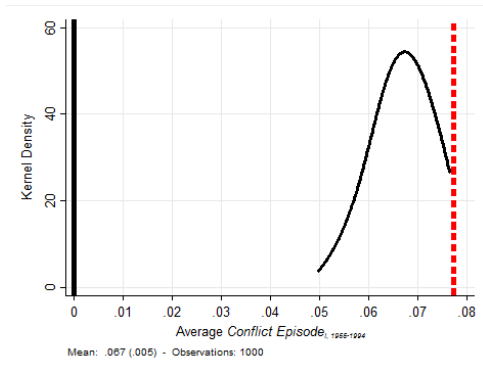
Table – Results with ICEWS

<i>Dep. Variable : Conflict Episode_{it}</i>	All Conflict Episodes _{it}		Conflicts [18-20] _{it}	
	(1)	(2)	(3)	(4)
<i>(# Schools / # Children)_i * Years Since 1978_t</i>	-0.00130 (0.000819)	-0.00225** (0.00112)	-0.00117* (0.000627)	-0.00134* (0.000782)
Observations	5,202	5,202	5,202	5,202
R-squared	0.447	0.516	0.372	0.447
District FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Province x Year FEs	No	Yes	No	Yes
Time-Window	1995-2012	1995-2012	1995-2012	1995-2012
Sample Mean	.4	.4	.26	.26

Notes : Standard errors are clustered at the level of districts (kabupate). *** Significant at the 1% level, ** at the 5% level, * at the 10%.

Monte Carlo with subset keywords I

Figure – Average conflict episodes estimated using 1,000 groups of conflict-related keywords

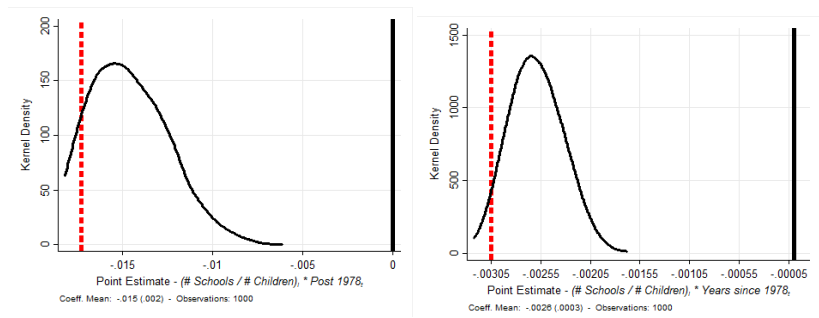


NOTE : The figure shows the distribution of the average conflict likelihood obtained using 1,000 different samples. Each sample was created by randomly drawing only two-thirds of the baseline keywords. The dashed line represents the average number of conflict episodes obtained using in the full baseline set of keywords.

Monte Carlo with subset keywords II

[▶ Back](#)

Figure – Distribution of coefficients estimated using 1,000 groups of conflict-related keywords



A) Column 3 in baseline Table

B) Column 6 in baseline Table

NOTE : The figure shows the distribution of coefficients estimated using 1,000 different samples. Each sample was created by randomly drawing two thirds of the baseline keywords. The dashed line represents the point estimate of the corresponding coefficient obtained using all baseline keywords.

Extended keywords and alternative source II

▶ Back

Table – Robustness : Alternative sources 2/2

Dep. Variable : Conflict Episode _{it}	(1) SMH	(2) SMH	(3) SMH-Broad	(4) SMH-Broad	(5) CT	(6) CT	(7) SMH+CT	(8) SMH+CT
(# Schools / # Children) _i * Years Since 1978 _t	-0.00146*** (0.000421)	-0.00305*** (0.000705)	-0.00133*** (0.000449)	-0.00256*** (0.000856)	-0.000882** (0.000348)	-0.00118** (0.000512)	-0.00176*** (0.000432)	-0.00321*** (0.000786)
Observations	11,560	11,560	11,560	11,560	11,560	11,560	11,560	11,560
R-squared	0.346	0.506	0.375	0.545	0.325	0.460	0.392	0.546
District FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trend	No	Yes	No	Yes	No	Yes	No	Yes
Province x Year FEs	No	Yes	No	Yes	No	Yes	No	Yes
Sample Mean	.08	.08	.13	.13	.06	.06	.06	.06

The unit of observation is a district i and year t . The dataset covers 289 districts across 26 provinces over the period 1955-1994. LPM estimates are reported in all columns. The dependent variable is a dummy that takes a value of 1 if a violent event was observed in district i in year t . Robust standard error clustered at the district level are reported in parenthesis. Statistical significance is represented by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Robustness province level

▶ Back

Table – Robustness Province Level

<i>Dep. Variable :</i>	Dummy Conflict _{pt}		(log) Districts with Conflict _{pt}	
	(1)	(2)	(3)	(4)
$(\# \text{ Schools} / \# \text{ Children})_p * \text{Post-1978}_t$	-0.0434** (0.0220)		-0.0695*** (0.0219)	
$(\# \text{ Schools} / \# \text{ Children})_p * \text{Years Since 1978}_t$		-0.00432** (0.00212)		-0.00777*** (0.00215)
Observations	1,040	1,040	1,040	1,040
R-squared	0.620	0.620	0.679	0.680
Province FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Sample Mean	.45	.45	.4511	.4511

Notes : *** Significant at the 1% level, ** at the 5% level, * at the 10%.

Logit

▶ Back

Table – Impact of INPRES school construction on conflict : Fixed effects logit estimator

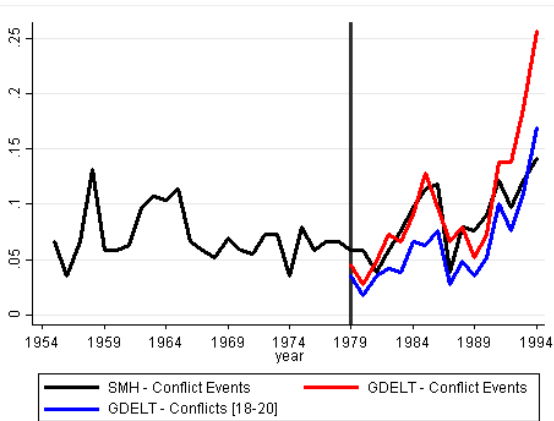
<i>Dep. Variable : Conflict Episode_{it}</i>	(1)	(2)	(3)	(4)
$(\# \text{ Schools} / \# \text{ Children})_i * \text{Post-1978}_t$	-0.435** (0.209)	-0.970*** (0.307)		
$(\# \text{ Schools} / \# \text{ Children})_i * \text{Years since 1978}_t$			-0.0323* (0.0169)	-0.0783** (0.0374)
Pseudo R-squared	0.281	0.386	0.280	0.386
Observations	5,920	5,920	5,920	5,920
District FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
District-Specific Linear Trend	No	Yes	No	Yes
Sample Mean	.14	.14	.14	.14

NOTE : The unit of observation is a district i and year t . The full sample covers 289 districts across 26 provinces over the period 1955-1994. Fixed effects logit estimates are reported in all columns. The dependent variable is a dummy that takes a value of 1 if a violent event was observed in district i and year t . Robust standard error clustered at the district level are reported in parenthesis. Statistical significance is represented by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Evolution of conflict episodes

[▶ Back](#)

Figure – Evolution of conflict episodes across alternative sources



SOURCE : Authors' computations from GDELT and own conflict data.