# Appendix for
# Identity conflict, ethnocentrism and social cohesion

Matteo Sestito*

June 30, 2021

*PhD candidate at Aix-Marseille School of Economics

# 1  Introduction

This appendix accompanies "Identity conflict, ethnocentrism and social cohesion" by Matteo Sestito. Section 2 describes the data used in the paper, while sections 3 and 4 report additional tables and robustness tests for the baseline and mechanism estimating equations, respectively.

# 2  Data

This section reports the exact text of the Afrobarometer variables as well as their values. Where variables have been recoded, a note explain the nature of the change. The body of the question is taken from the sixth Afrobarometer round.

## Outcome variables

- Social identity

  ◇ *Afrobarometer question*: "Let us suppose that you had to choose between being a [enter nationality] and being a [Respondent's ethnic group]. Which of the following best expresses your feelings?"

  ◇ *Values (recoded)*: 1=I feel only [enter nationality]; 2=I feel more [enter nationality] than (Respondent's ethnic group); 3=I feel equally [enter nationality] and (Respondent's ethnic group); 4=I feel more (Respondent's ethnic group) than [enter nationality]; 5=I feel only (Respondent's ethnic group)

  ◇ *Values in binary recoding*: 0=I feel only (Respondent's ethnic group) & I feel more (Respondent's ethnic group) than [enter nationality], 1=Otherwise

  ◇ *Afrobarometer question number (rounds 3 to 6)*: Q82, Q83, Q85b, Q88b

  ◇ *Further comments*: The original coding of the variable is inverted around its median (*e.g.* value 1 in the original coding corresponds to value 5 in the recoded variable used in the analysis)

- Attendance community meetings

  ◇ *Afrobarometer question*: "Here is a list of actions that people sometimes take as citizens. For each of these, please tell me whether you,

personally, have done any of these things during the past year. If not, would you do this if you had the chance: Attended a community meeting?"

◇ *Values*: 0=No, would never do this; 1=No, but would do if had the chance; 2=Yes, once or twice; 3=Yes, several times; 4=Yes, often

◇ *Values in binary recoding*: 0=No, would never do this & No, but would do if had the chance; 1=Otherwise

◇ *Afrobarometer question number (rounds 3 to 6)*: Q31a, Q23a, Q26a, Q20a

- Collective action

  ◇ *Afrobarometer question*: "Here is a list of actions that people sometimes take as citizens. For each of these, please tell me whether you, personally, have done any of these things during the past year. If not, would you do this if you had the chance: Got together with others to raise an issue?"

  ◇ *Values*: 0=No, would never do this; 1=No, but would do if had the chance; 2=Yes, once or twice; 3=Yes, several times; 4=Yes, often

  ◇ *Values in binary recoding*: 0=No, would never do this & No, but would do if had the chance; 1=Otherwise

  ◇ *Afrobarometer question number (rounds 3 to 6)*: Q31b, Q23b, Q26b, Q20b

- Membership community associations

  ◇ *Afrobarometer question*: "Let's turn to your role in the community. Now I am going to read out a list of groups that people join or attend. For each one, could you tell me whether you are an official leader, an active member, an inactive member, or not a member: Some other voluntary association or community group?"

  ◇ *Values*: 0=Not a member; 1=Inactive member; 2=Active member; 3=Official leader

  ◇ *Values in binary recoding*: 0=Not a member & Inactive member; 1=Otherwise

  ◇ *Afrobarometer question number (rounds 3 to 6)*: Q28d, Q22b, Q25b, Q19b

◇ *Further comments*: The question in the third round is phrased slightly differently. Moreover, in this survey, contrarily to the others, respondents were previously asked also for their membership in trade unions and business associations

## Control variables

- Age

  ◇ *Afrobarometer question*: "How old are you?"
  ◇ *Values*: 18-130
  ◇ *Afrobarometer question number (rounds 3 to 6)*: Q1, Q1, Q1, Q1

- Gender

  ◇ *Afrobarometer question*: "Respondent's gender"
  ◇ *Values (recoded)*: 0= Female; 1= Male
  ◇ *Afrobarometer question number (rounds 3 to 6)*: Q101, Q101, Q101, Q101

- Urban

  ◇ *Afrobarometer question*: "Urban or Rural Primary Sampling Unit"
  ◇ *Values (recoded)*: 0=Rural; 1=Urban
  ◇ *Afrobarometer question number (rounds 3 to 6)*: Urbrur

- Education

  ◇ *Afrobarometer question*: "What is your highest level of education?"
  ◇ *Values*: 0=No formal schooling; 1=Informal schooling only (including Koranic schooling); 2=Some primary schooling, 3=Primary school completed; 4=Intermediate school or Some secondary school / high school; 5=Secondary school / high school completed, 6=Post-secondary qualifications, other than university *e.g.* a diploma or degree from a polytechnic or college; 7=Some university; 8=University completed; 9=Post-graduate
  ◇ *Afrobarometer question number (rounds 3 to 6)*: Q90, Q89, Q97, Q97

- Employment status

  - *Afrobarometer question*: "Do you have a job that pays a cash income? If yes, is it full-time or part-time? If no, are you presently looking for a job?"

  - *Values (recoded)*: 0=No (not looking); 1=No (looking); 2=Yes part time; 3= Yes, full time

  - *Afrobarometer question number (rounds 3 to 6)*: Q94, Q94, Q96, Q95

  - *Further comments*: In Afrobarometer rounds 3 and 4 respondents can distinguish between "not looking" and "looking" also in reference to the part- and full-time categories. This distinction is eliminated to express the variables as in the value-scale reported above

- Household wealth

  - The covariate is equal to the sum of five variables measuring household wealth

  - *Afrobarometer question*: "Over the past year, how often, if ever, have you or anyone in your family gone without: $X$?"
  Where $X \in$ {enough food to eat, enough clean water for home use, medicines or medical treatment, enough fuel to cook your food, a cash income}

  - *Values*: 0=Never; 1=Just once or twice; 2=Several times; 3=Many times; 4=Always

  - *Afrobarometer question number (rounds 3 to 6)*: Q8a-e, Q8a-e, Q8a-e, Q8a-e

- Trust president

  - *Afrobarometer question*: "How much do you trust each of the following, or haven't you heard enough about them to say: The President?"

  - *Values*: 0=Not at all; 1=Just a little; 2=Somewhat; 3=A lot

  - *Afrobarometer question number (rounds 3 to 6)*: Q55a, Q49a, Q59a, Q52a

- Trust local government

  - *Afrobarometer question*: "How much do you trust each of the following, or haven't you heard enough about them to say: Your Metropolitan, Municipal or District Assembly?"

       ⋄ *Values*: 0=Not at all; 1=Just a little; 2=Somewhat; 3=A lot

       ⋄ *Afrobarometer question number (rounds 3 to 6)*: Q55d, Q49d, Q59e, Q52e

- Voted

       ⋄ *Afrobarometer question*: "Understanding that some people were unable to vote in the most recent national election in [20xx], which of the following statements is true for you?"

       ⋄ *Values (recoded)*: 0=Did not vote; 1=Voted

       ⋄ *Afrobarometer question number (rounds 3 to 6)*: Q30, Q23d, Q27, Q21

       ⋄ *Further comments*: The original question list various reasons for not having voted, which change from one Afrobarometer round to another. They all have been coded within the same no-vote category

# 3 Robustness tests baseline analysis

This section reports additional robustness tests of the baseline models estimated in section 4 of the paper. The estimating equation of interest is:

$$y_{i,e,k,t} = \mu_{k,t} + \lambda_{e,t} + \beta C_{e,k,t} + \delta' x_{i,e,k,t} + u_{i,e,k,t} \tag{A1}$$

Where all the variables have the same meaning as in the paper. The coefficient of interest is the marginal impact of ethnic conflict ($\beta$), which is expected to be positive.

## 3.1 *Outlier observations*

As clear from Figure 2 in the paper, ethnic conflict is spread very unevenly across Sub-Saharan Africa. For each country, Figure A1 reports the cell-level average of conflict events per-months.[1] Uganda is the coun-

---

[1] That is, the event-based definition of conflict – which is normalised by the number of within-sample months of each cell – is first aggregated at the country level and then divided by the number of cells in each country. Formally:

$$\frac{1}{|\mathcal{K}_p|} \sum_{k \in \mathcal{K}_p} \widetilde{C}_{e,k,t}$$

where $\mathcal{K}_p$ is the set of cells in country $p$ and $\widetilde{C}_{e,k,t}$ is the event-based definition of conflict defined in the data section of the paper.

**Figure A1:** Ethnic conflict in selected Sub-Saharan countries
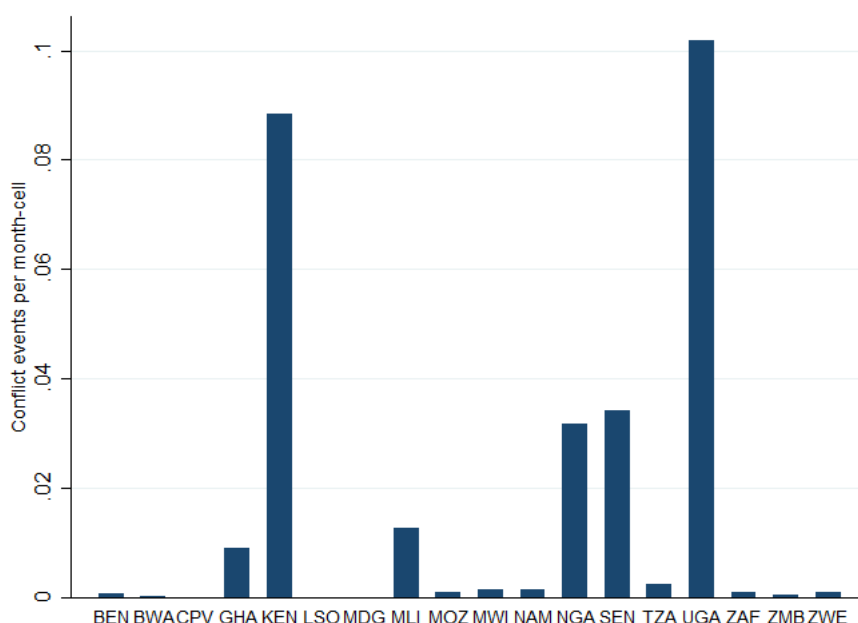


*Note:* Conflict events-per-month are aggregated at the country-level throughout the whole sample period.

try with the highest density of ethnic conflict events, primarily reflecting the toll taken by the bloody rebellion of the Lord's Resistance Army (LRA) in the early 2000s. On the other end of the spectrum are Cape Vert, Lesotho, and Madagascar, that did not experience any conflict event during the whole period. Table A1 checks robustness to the exclusion of these outlier countries. The exclusion of Uganda is particularly relevant in terms of external validity, inasmuch as most of the previous evidence on the relationship between conflict and social cohesion came from this country. The first two columns of Table A1 substantiate that the baseline results are not an Ugandan peculiarity: the estimated coefficient remains always positive, albeit is not significantly different than zero for the collective action variable and for the regression of attendance at community meetings when Afrobarometer controls are included. When dropping the three countries without any conflict activity (columns 3 and 4), results barely change from the baseline both in terms of size and standard errors.

A related concern is that the impact of conflict on social cohesion might be driven by few influential ethnic groups. Table A2 reports the total number of per-month conflict events for the 65 ethnolinguistic groups

6

**Table A1:** OLS regressions of Social capital on Ethnic conflict – Outlier countries

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| SOCIAL IDENTITY |  |  |  |  |
| Conflict indicator | 0.088 | 0.086 | 0.083 | 0.079 |
|  | (0.039)** | (0.039)** | (0.036)** | (0.036)** |
| $R^2$ | 0.21 | 0.21 | 0.22 | 0.22 |
| $N$ | 95,161 | 92,990 | 93,406 | 91,247 |
| ATTENDANCE COMMUNITY MEETINGS |  |  |  |  |
| Conflict indicator | 0.081 | 0.056 | 0.094 | 0.077 |
|  | (0.042)* | (0.040) | (0.039)** | (0.037)** |
| $R^2$ | 0.23 | 0.27 | 0.20 | 0.25 |
| $N$ | 102,247 | 99,960 | 96,843 | 94,578 |
| COLLECTIVE ACTION |  |  |  |  |
| Conflict indicator | 0.045 | 0.029 | 0.068 | 0.058 |
|  | (0.043) | (0.042) | (0.039)* | (0.038) |
| $R^2$ | 0.24 | 0.27 | 0.25 | 0.27 |
| $N$ | 102,009 | 99,737 | 96,652 | 94,399 |
| MEMBERSHIP COMMUNITY ASSOCIATIONS |  |  |  |  |
| Conflict indicator | 0.098 | 0.091 | 0.085 | 0.082 |
|  | (0.034)*** | (0.034)*** | (0.031)*** | (0.031)*** |
| $R^2$ | 0.17 | 0.19 | 0.17 | 0.19 |
| $N$ | 101,860 | 99,597 | 96,481 | 94,239 |
| Fixed effects | Yes | Yes | Yes | Yes |
| Individual cov. | No | Yes | No | Yes |
| Drop Uganda | Yes | Yes | No | No |
| Drop MDG, CPV, LSO | No | No | Yes | Yes |

The sample includes Afrobarometer rounds 3 to 6. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and a urban-rural indicator. Standard errors are clustered at the cell-ethnolinguistic-period level and are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

that have experienced at least one conflict incident. There is, indeed, some substantial heterogeneity in warfare exposure. Acholi, Kikuyu and Ijaw are the three most war-prone groups, the only experiencing more than three conflict incidents per month. The first has been involved in the LRA insurgency and the ensuing disorders.[2] The Kikuyu of Kenya have been exposed to different types of conflict, ranging from tribal clashes to urban disorders in the slums of Nairobi at the hands of the 'Mungiki Militia'. The Ijaw of Nigeria have taken part in conflicts interesting the oil-rich Delta region of the Niger river. The area has known "decades of environmental pollution, economic underdevelopment and political marginalization",[3] which have ultimately resulted in fertile ground for inter-ethnic tensions. Hence, the involvement of the Ijaw through various ethnic militias and loose armed groups such as the 'MEND: Movement for the Emancipation of the Niger Delta' (Hazen and Horner 2007).

The baseline model is thus re-estimated excluding these outlier groups. Results are reported in Table A3. The first two columns substantiate the robustness of results to the exclusion of Acholi respondents from the sample. Columns three and four run a more demanding specification whereby Acholi, Kikuyu and Ijaw ethnolinguistic groups are all excluded from the regressions. Marginal impacts maintain the expected sign, albeit they fail to achieve statistical significance in the case of two social capital outcomes. Finally, in the last two columns I drop all the ethnolinguistic groups not experiencing any conflict activity. Even though the sample size is roughly halved, the estimated coefficients are positive and virtually always statistically different than zero.

---

[2]For an overview of the conflict the reader is referred to Rohner *et al* (2013) and the summary in the UCDP database at: `https://ucdp.uu.se/actor/488`.

[3]Asuni (2009), p. 3.

**Table A2:** Conflict incidence by ethnolinguistic group

| ETHNOLINGUISTIC GROUP | CONFLICT INCIDENCE | ETHNOLINGUISTIC GROUP | CONFLICT INCIDENCE |
|---|---|---|---|
| Acholi | 3.85 | Lugbara | .18 |
| Afrikaans | .13 | Lugwere | .02 |
| Akan | .15 | Luhya | .08 |
| Alur | .08 | Lukhonjo | .17 |
| Ateso | .37 | Lumasaaba | .02 |
| Chewa | .05 | Luo | 1.12 |
| Dagbani | .1 | Lusoga | .02 |
| Diola | 2.24 | Mampruli | .08 |
| Edo | .03 | Ndebele | .02 |
| Eggon | .12 | Ngakarimajong | .71 |
| Ewe | .22 | Ogoni | .08 |
| Fulani | .42 | Oroma | .87 |
| Hausa | .08 | Oshiwambo | .09 |
| Herero | .04 | Pokot | .42 |
| Idoma | .04 | Pulaar/Toucouleur | .02 |
| Igbo | 1.3 | Rufumbira | 1.57 |
| Ijaw | 3.25 | Rukiga | .17 |
| Isoko | .03 | Rukwangali | .02 |
| Japadhola | .05 | Runyankole | .02 |
| Kalabari | .04 | Runyoro | .12 |
| Kalenjin | 1.06 | Shangaan/Tsonga | .2 |
| Kanuri | .03 | Shona | .04 |
| Karanga | .04 | Somali | .24 |
| Kikuyu | 3.72 | Sotho | .02 |
| Kimasai | .2 | Tamasheq | 1.71 |
| Kimeru | .13 | Tarok | .27 |
| Kisii | .08 | Tiv | .07 |
| Konkomba | .08 | Tonga | .02 |
| Kupsabinyi | .02 | Turkana | .22 |
| Kuria | .33 | Urhobo | .12 |
| Kusal | .14 | Xhosa | .05 |
| Langi | .02 | Yoruba | 1.13 |
| Lozi | .02 | | |

Conflict incidence refers to the monthly average number of conflict events experienced by an ethnolinguistic group throughout the whole period of analysis. Only ethnolinguistic groups experiencing at least a conflict event are included in the table.

**Table A3:** OLS regressions of Social capital on Ethnic conflict – Outlier ethnic groups

|  | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **SOCIAL IDENTITY** | | | | | | |
| Conflict indicator | 0.084 | 0.080 | 0.073 | 0.068 | 0.068 | 0.066 |
|  | (0.036)** | (0.036)** | (0.040)* | (0.041)* | (0.040)* | (0.041) |
| $R^2$ | 0.22 | 0.22 | 0.22 | 0.22 | 0.22 | 0.22 |
| $N$ | 103,333 | 101,124 | 101,509 | 99,309 | 50,348 | 49,816 |
| **ATTENDANCE COMMUNITY MEETINGS** | | | | | | |
| Conflict indicator | 0.098 | 0.080 | 0.062 | 0.056 | 0.123 | 0.098 |
|  | (0.039)** | (0.037)** | (0.042) | (0.041) | (0.042)*** | (0.042)** |
| $R^2$ | 0.22 | 0.27 | 0.22 | 0.27 | 0.23 | 0.28 |
| $N$ | 110,449 | 108,121 | 108,622 | 106,303 | 52,426 | 51,870 |
| **COLLECTIVE ACTION** | | | | | | |
| Conflict indicator | 0.067 | 0.055 | 0.039 | 0.036 | 0.124 | 0.104 |
|  | (0.039)* | (0.038) | (0.043) | (0.043) | (0.043)*** | (0.043)** |
| $R^2$ | 0.23 | 0.26 | 0.24 | 0.27 | 0.26 | 0.29 |
| $N$ | 110,195 | 107,883 | 108,375 | 106,072 | 52,330 | 51,776 |
| **MEMBERSHIP COMMUNITY ASSOCIATIONS** | | | | | | |
| Conflict indicator | 0.089 | 0.085 | 0.097 | 0.097 | 0.078 | 0.075 |
|  | (0.031)*** | (0.031)*** | (0.034)*** | (0.033)*** | (0.038)** | (0.037)** |
| $R^2$ | 0.17 | 0.19 | 0.17 | 0.19 | 0.16 | 0.19 |
| $N$ | 110,066 | 107,761 | 108,240 | 105,944 | 52,278 | 51,728 |
| Fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Individual cov. | No | Yes | No | Yes | No | Yes |
| Drop Acholi | Yes | Yes | Yes | Yes | No | No |
| Drop Kikuyu & Ijaw | No | No | Yes | Yes | No | No |
| Drop peaceful groups | No | No | No | No | Yes | Yes |

The sample includes Afrobarometer rounds 3 to 6. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Standard errors are clustered at the cell-ethnolinguistic-period level and are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

## 3.2 *Standard error clustering*

In the baseline specification standard errors are allowed to be correlated at the cell-ethnolinguistic-period level. The level of clustering was chosen so as to match the level of assignment of the conflict indicator. I here relax this assumption and allow standard errors to be correlated within each ethnolinguistic group and within each $0.5° \times 0.5°$ cell. As shown in Figure A2, apart for the models of collective action clustering standard errors spatially, sign and statistical significance of the baseline estimates are always reproduced: the validity of the results does not hinge upon the level of clustering.

**Figure A2:** OLS regressions of Social capital on Ethnic conflict – Alternative standard error clustering
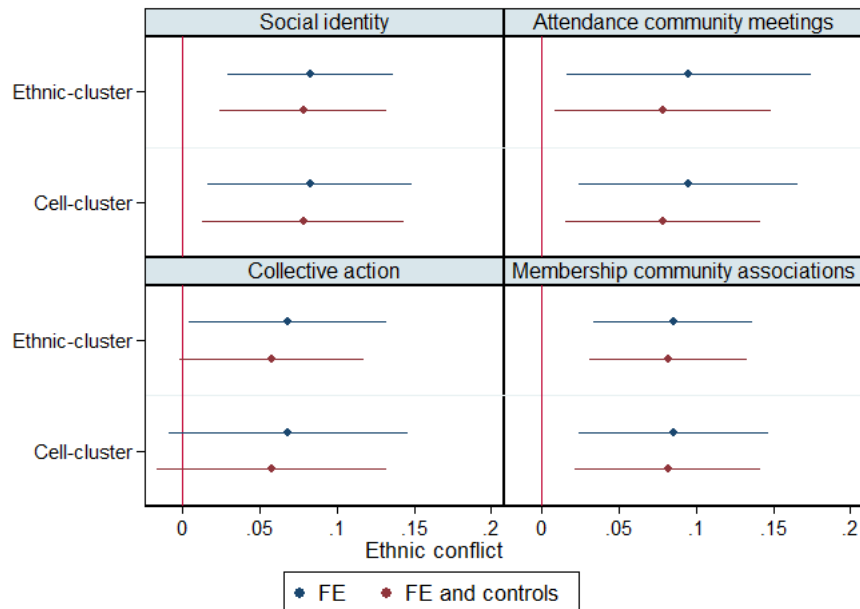


*Note:* The figure reports point estimates alongside their 90% confidence intervals for the coefficients of the conflict indicator. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Standard errors are clustered either at the cell- or ethnolinguistic-level.

# 4 Robustness tests mechanism analysis

## 4.1 *Distant conflict*

The model investigating the impact of remote conflict activity on social capital is:

$$y_{i,e,k,t} = \mu_{k,t} + \lambda_{e,t} + \beta C_{e,k,t} + \phi \mathcal{C}_{e,n(k),t} + \delta' x_{i,e,k,t} + u_{i,e,k,t} \qquad \text{(A2)}$$

Where $\mathcal{C}_{e,n(k),t}$ is the ethnic conflict indicator computed over the eight-cell neighbourhood of cell $k$. The meaning of the other variables is the same as in the paper. The coefficient of interest is $\phi$, which is expected to be positive.

Figure A3 checks robustness to alternative actor-ethnicity matching procedures, namely, that excluding imprecise matches and that dropping matches involving sub-groups of Afrobarometer languages. While $\phi$ is always positive, statistical significance is not always achieved. Estimates relative to collective action are virtually never statistically different than zero; point estimates for the other two outcomes, instead, tend to be significant at conventional levels with the exception of the regression of membership on community associations on ethnic conflict excluding sub-groups of Afrobarometer ethnolinguistic groups.

Figure A4 reports estimates of equation (A2) over two alternative samples, namely, that dropping ethnolinguistic groups not present in all Afrobarometer surveys and that using the Afrobarometer question on ethnicity as marker defining ethnolinguistic groups. The results mimic those in the baseline analysis, with the exception of the regression of membership in community associations on external conflict, whose estimates, albeit having the expected sign, are not significantly different than zero when the ethnic marker is employed.

Table A4 checks robustness to a binary recoding of the outcomes variables. While the estimates relative to the collective action outcome are small and insignificantly different than zero, the marginal impact of external conflict on attendance and membership in community organisations is positive and, when no additional Afrobarometer controls are included, significantly different than zero.
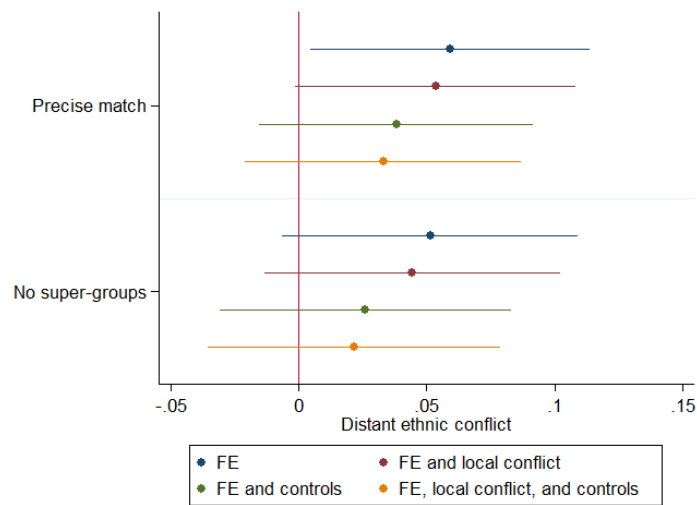
**Forthcoming: robustness tests using different spatial aggregation criteria for external conflict, namely, the 24- and 48-cell neighbourhoods.**

**Figure A3:** OLS regressions of Social cohesion variables on Distant ethnic conflict − Alternative matching rules
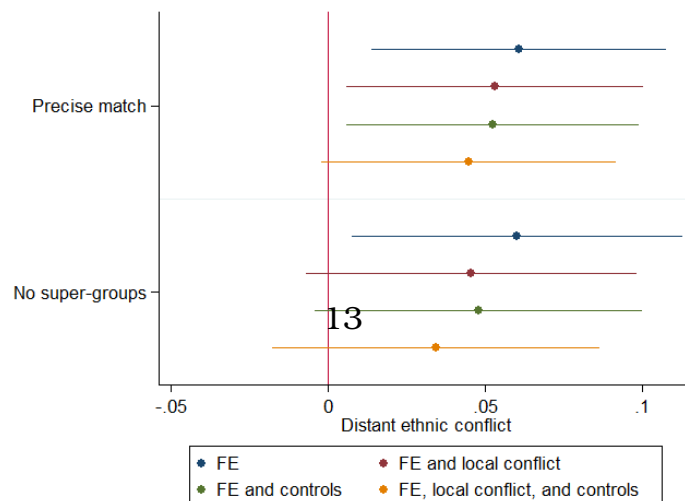
**(a)** Attendance community meetings
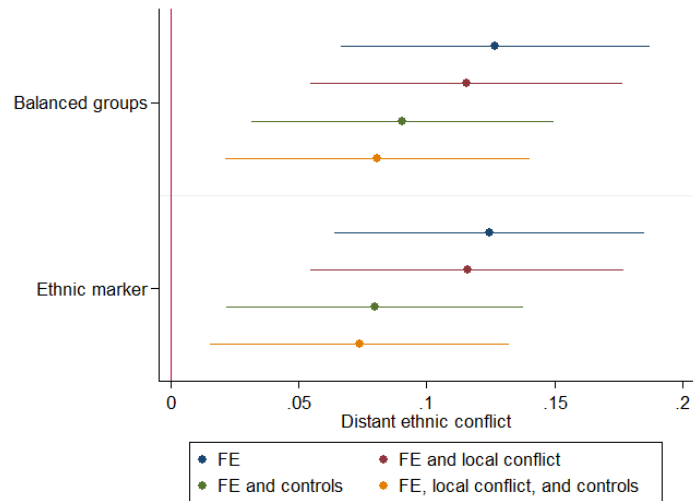


**(b)** Collective action
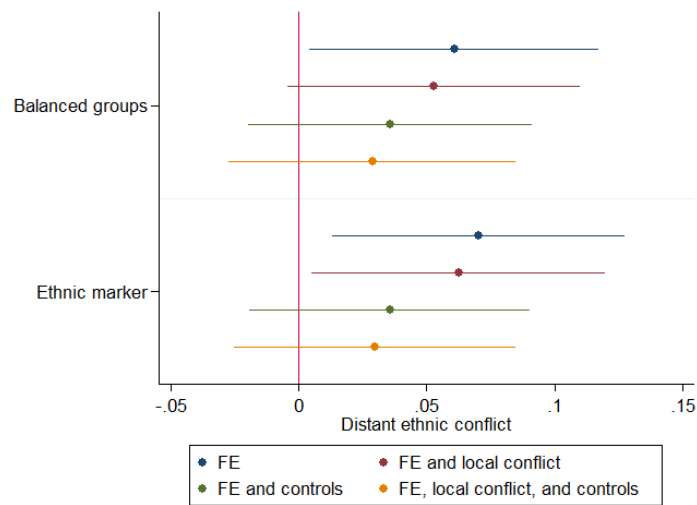


**(c)** Membership community associations



*Note:* The figure reports point estimates alongside their 90% confidence intervals for the coefficients of distant ethnic conflict. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Standard errors are clustered at the cell-ethnolinguistic-period level.

**Figure A4:** OLS regressions of Social cohesion variables on Distant ethnic conflict − Alternative sample restrictions
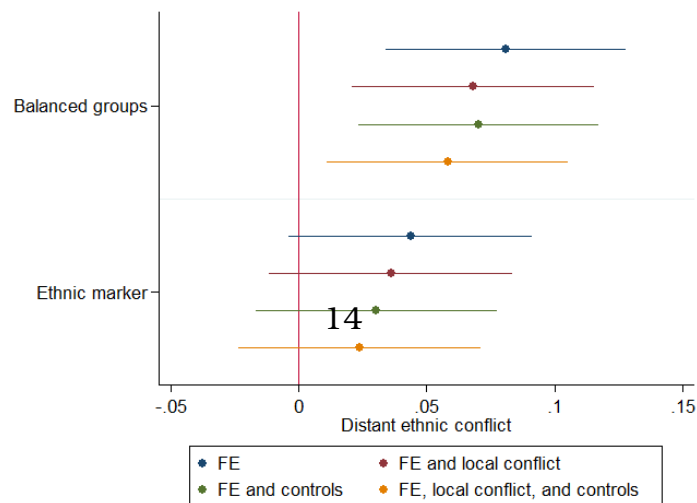
**(a)** Attendance community meetings



**(b)** Collective action



**(c)** Membership community associations



14

*Note:* The figure reports point estimates alongside their 90% confidence intervals for the coefficients of distant ethnic conflict. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Standard errors are clustered at the cell-ethnolinguistic-period level.

**Table A4:** OLS regressions of Social capital on Distant ethnic conflict –
Alternative dependent variable coding

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| ATTENDANCE COMMUNITY MEETINGS | | | | |
| Distant conflict | 0.031 | 0.025 | 0.019 | 0.013 |
|  | (0.014)** | (0.014)* | (0.013) | (0.014) |
| Local conflict |  | 0.049 |  | 0.045 |
|  |  | (0.015)*** |  | (0.015)*** |
| $R^2$ | 0.18 | 0.18 | 0.22 | 0.22 |
| $N$ | 110,718 | 110,718 | 108,389 | 108,389 |
| COLLECTIVE ACTION | | | | |
| Distant conflict | 0.013 | 0.007 | 0.005 | -0.001 |
|  | (0.013) | (0.013) | (0.013) | (0.013) |
| Local conflict |  | 0.044 |  | 0.042 |
|  |  | (0.015)*** |  | (0.015)*** |
| $R^2$ | 0.19 | 0.19 | 0.21 | 0.21 |
| $N$ | 110,464 | 110,464 | 108,151 | 108,151 |
| MEMBERSHIP COMMUNITY ASSOCIATIONS | | | | |
| Distant conflict | 0.026 | 0.022 | 0.021 | 0.017 |
|  | (0.013)** | (0.013)* | (0.013) | (0.013) |
| Local conflict |  | 0.026 |  | 0.025 |
|  |  | (0.015)* |  | (0.015)* |
| $R^2$ | 0.15 | 0.15 | 0.17 | 0.17 |
| $N$ | 110,335 | 110,335 | 108,029 | 108,029 |
| Fixed effects | Yes | Yes | Yes | Yes |
| Individual cov. | No | No | Yes | Yes |

The sample includes Afrobarometer rounds 3 to 6. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Distant conflict is spatially aggregated on the 8-cell neighbourhood of each observation. Standard errors are clustered at the cell-ethnolinguistic-period level and are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

## 4.2 *Ethnic vs non-ethnic conflict*

As reported in the paper, the model confronting ethnic and non-ethnic conflict is the following:

$$y_{i,e,k,t} = \mu_{k,t} + \lambda_{e,t} + \beta C_{e,k,t} + \psi Z_{-e,k,t} + \delta' x_{i,e,k,t} + u_{i,e,k,t} \tag{A3}$$

Where all the variables have the same meaning as in the paper. The coefficients of interest are $\beta$ and $\psi$, whereby we expect the former to be greater than the latter.
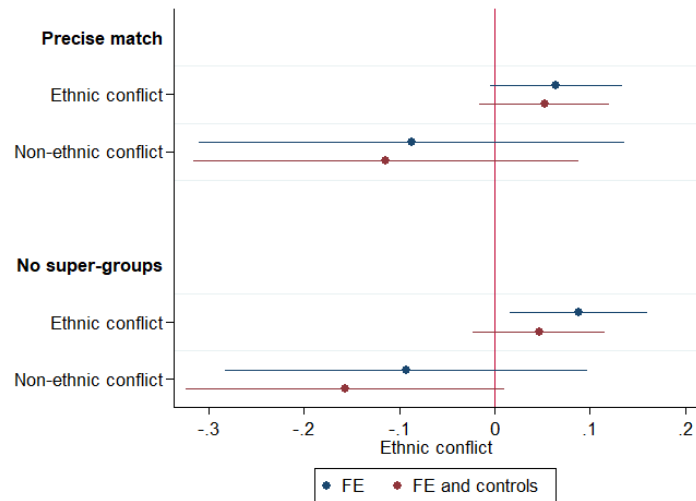
Figure A5 reports estimates from the models employing alternative matching procedures adopted in the definition of ethnic conflict. The point estimates for ethnic conflict ($\beta$) are always positive, albeit statistical significance is fully achieved only for the outcome on membership in community associations. The coefficient on non-ethnic conflict is instead always statistically indistinguishable from zero. More importantly, it is always considerably smaller than $\beta$ and in some specifications even turns negative.

Figure A6 investigates robustness to the alternative samples defined by the exclusion of transient ethnolinguistic groups and the use of the self-reported ethnic affiliation to define ethnolinguistic families, respectively. Estimates of $\beta$ are positive and virtually always statistically significant. On the other hand, the impact of non-ethnic conflict is more imprecisely estimated, generally lower, and statistically insignificant.
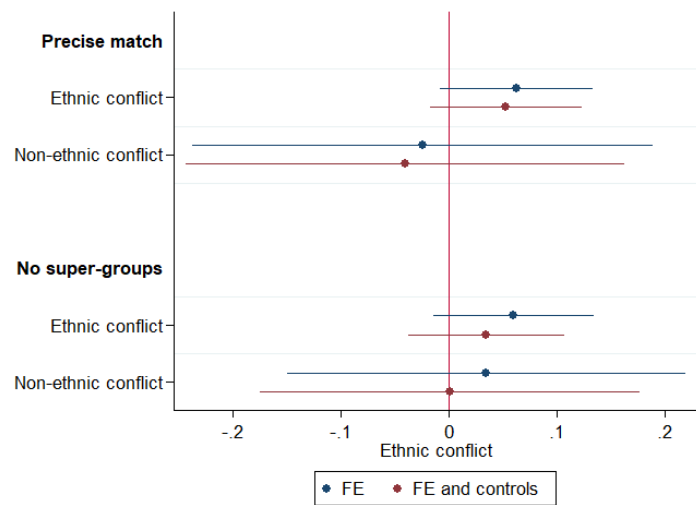
Finally, equation (A3) is estimated using the binary definition of the outcome variables. Results from this exercise are reported in Figure A7. All point estimates relative to ethnic conflict are positive and significant. The $\psi$ coefficient, instead, is always smaller than $\beta$ and never achieves statistical significance.

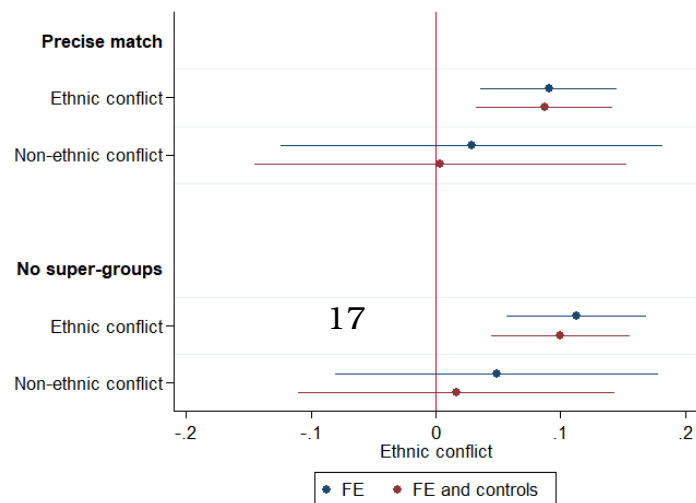**Figure A5:** OLS regressions of Social capital on Ethnic and non-ethnic conflict – Alternative matching rules

**(a)** Attendance community meetings
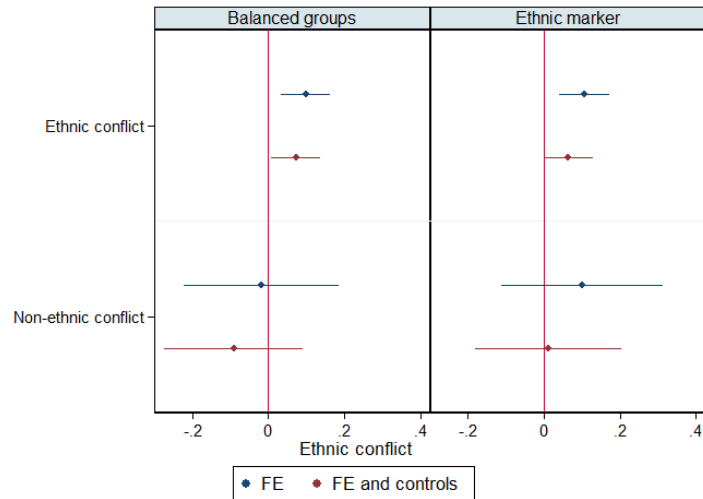


**(b)** Collective action
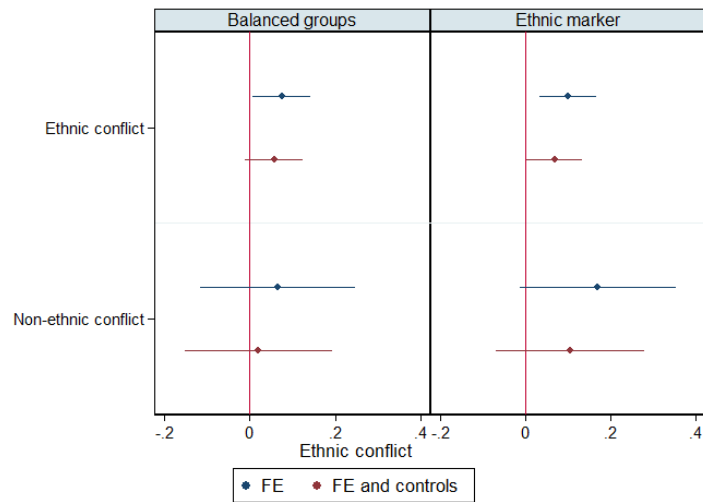


**(c)** Membership community associations



*Note:* The figure reports point estimates alongside their 90% confidence intervals for the coefficients of ethnic and non-ethnic conflict. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Standard errors are clustered at the cell-ethnolinguistic-period level.

**Figure A6:** OLS regressions of Social cohesion variables on Ethnic and non-ethnic conflict – Alternative sample restrictions
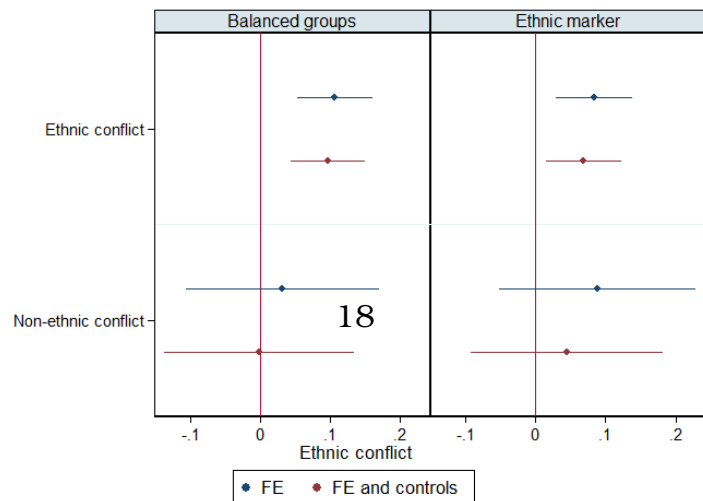
**(a)** Attendance community meetings



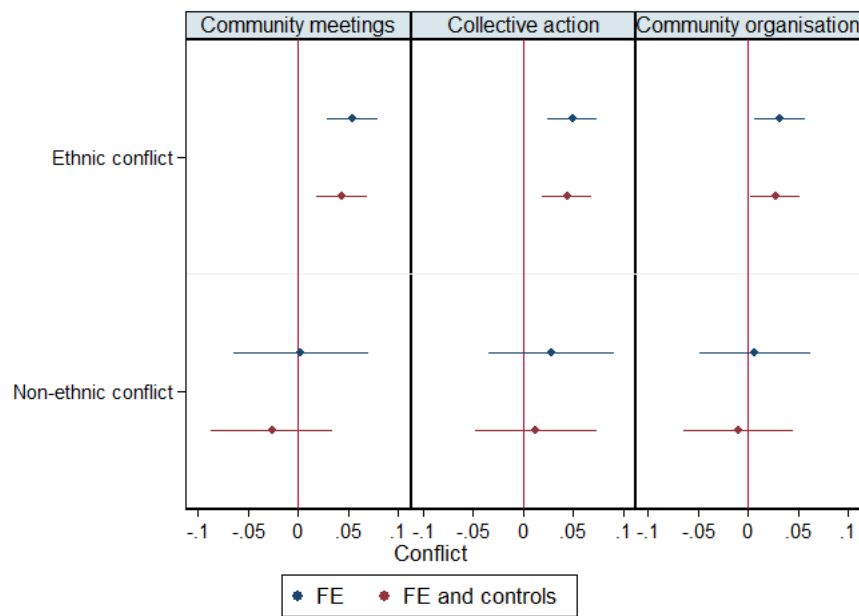**(b)** Collective action



**(c)** Membership community associations

18



*Note:* The figure reports point estimates alongside their 90% confidence intervals for the coefficients of ethnic and non-ethnic conflict. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Standard errors are clustered either at the cell-ethnolinguistic-period level.

**Figure A7:** OLS regressions of Social capital on Ethnic and non-ethnic conflict − Alternative dependent variable coding



*Note:* The figure reports point estimates alongside their 90% confidence intervals for the coefficients of ethnic and non-ethnic conflict. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Standard errors are clustered at the cell-ethnolinguistic-period level.

## 4.3 *Heterogeneous effects*

The model investigating heterogeneous effects with respect to the cell-period ethnic fractionalisation is:

$$y_{i,e,k,t} = \mu_{k,t} + \lambda_{e,t} + \beta C_{e,k,t} + \alpha(C_{e,k,t} \times ELF_{k,t}) + \delta' x_{i,e,k,t} + u_{i,e,k,t} \qquad \text{(A4)}$$

Where all the variables have the same meaning as in the paper. Interest lies in the joint effect of conflict and ethnolinguistic fractionalisation: the parameter $\alpha$, which we expect to be negative.

Table A5 checks robustness to the two already-mentioned alternative ethnicity-conflict actor matches. The coefficient of interest ($\alpha$) is always negative. All point estimates are statistically different than zero at least at the 5% significance threshold but for those retrieved from the regressions of membership in community associations when super-groups are excluded.

Figure A8 reports results from equation (A4) as estimated with the two already-defined alternative samples. When transient ethnolinguistic groups are excluded, magnitude and size of the coefficients closely resemble those of the baseline analysis. Results are less robust to the use of the Afrobarometer question on ethnicity to define ethnolinguistic aggregates. The estimated $\alpha$ is negative only for two outcomes and virtually never significant.

Finally, Table A6 checks robustness to the alternative binary coding of outcome variables. All regressions include the cell $\times$ period and ethnic $\times$ period fixed effects, with each column adding cumulatively the set of individual, socio-economic, and political variables. The joint impact of conflict and ethnolinguistic fractionalisation is always negative and, apart for the attendance at community meetings outcome, generally significant at the 5% threshold.

**Table A5:** OLS regressions of Social capital on Ethnic conflict – Alternative matching rules

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **ATTENDANCE COMMUNITY MEETINGS** | | | | |
| Ethnic prec. match | 0.302 | 0.298 | | |
|  | (0.089)*** | (0.089)*** | | |
| Ethnic-ELF int. prec. match | -0.392 | -0.385 | | |
|  | (0.155)** | (0.154)** | | |
| Ethnic no super-group | | | 0.281 | 0.280 |
|  | | | (0.090)*** | (0.089)*** |
| Ethnic-ELF int. no super-group | | | -0.314 | -0.328 |
|  | | | (0.153)** | (0.152)** |
| $R^2$ | 0.17 | 0.18 | 0.17 | 0.18 |
| $N$ | 110,335 | 108,029 | 110,335 | 108,029 |
| **COLLECTIVE ACTION** | | | | |
| Ethnic prec. match | 0.495 | 0.471 | | |
|  | (0.123)*** | (0.117)*** | | |
| Ethnic-ELF int. prec. match | -0.782 | -0.753 | | |
|  | (0.216)*** | (0.208)*** | | |
| Ethnic no super-group | | | 0.345 | 0.338 |
|  | | | (0.120)*** | (0.114)*** |
| Ethnic-ELF int. no super-group | | | -0.522 | -0.548 |
|  | | | (0.205)** | (0.193)*** |
| $R^2$ | 0.23 | 0.26 | 0.23 | 0.26 |
| $N$ | 110,464 | 108,151 | 110,464 | 108,151 |
| **MEMBERSHIP COMMUNITY ASSOCIATIONS** | | | | |
| Ethnic prec. match | 0.340 | 0.314 | | |
|  | (0.120)*** | (0.117)*** | | |
| Ethnic-ELF int. prec. match | -0.481 | -0.448 | | |
|  | (0.209)** | (0.203)** | | |
| Ethnic no super-group | | | 0.209 | 0.194 |
|  | | | (0.121)* | (0.115)* |
| Ethnic-ELF int. no super-group | | | -0.196 | -0.229 |
|  | | | (0.211) | (0.193) |
| $R^2$ | 0.22 | 0.27 | 0.22 | 0.27 |
| $N$ | 110,718 | 108,389 | 110,718 | 108,389 |
| Fixed effects | Yes | Yes | Yes | Yes |
| Personal cov. | No | Yes | No | Yes |

The sample includes Afrobarometer rounds 3 to 6. Fixed effects correspond to cell × period and ethnic × period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. The alternative conflict measures exclude actor-ethnicity matches that are either controversial or involve a sub-group of Afrobarometer language. Standard errors are clustered at the cell-ethnolinguistic-period level and are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

**Figure A8:** OLS regressions of Social cohesion variables on Ethnic conflict –
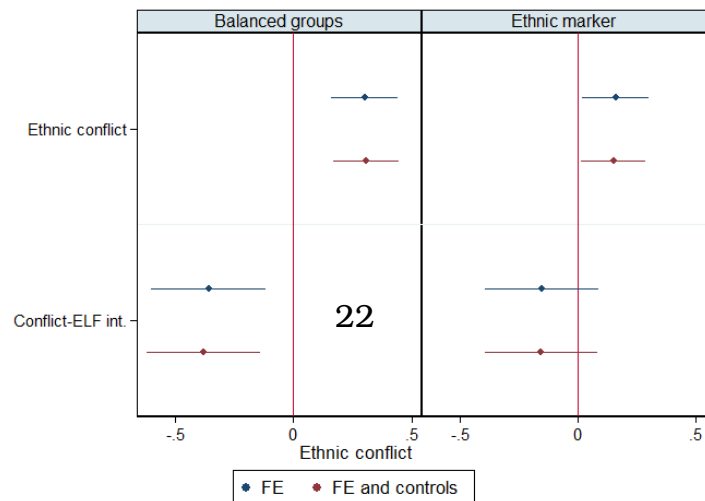Alternative sample restrictions
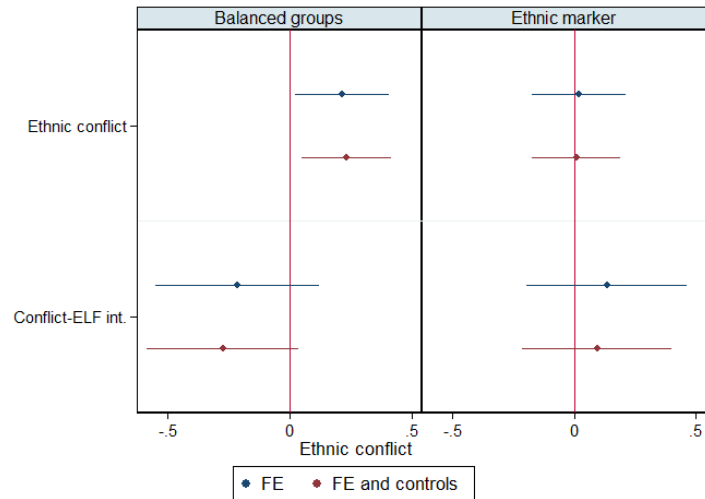
**(a)** Attendance community meetings



**(b)** Collective action
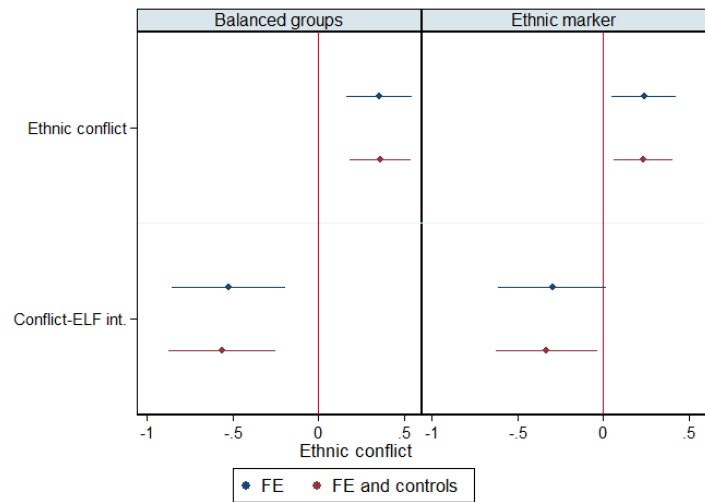


**(c)** Membership community associations



*Note:* The figure reports point estimates alongside their 90% confidence intervals for the coefficients of
ethnic conflict and its interaction with the ELF index. Fixed effects correspond to cell × period and ethnic
× period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator.
Standard errors are clustered either at the cell-ethnolinguistic-period level.

**Table A6:** OLS regressions of Social capital on Ethnic conflict – Alternative dependent variable coding

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| ATTENDANCE COMMUNITY MEETINGS | | | | |
| Ethnic conflict | 0.054 | 0.059 | 0.174 | 0.192 |
|  | (0.045) | (0.042) | (0.110) | (0.112)* |
| Conflict-ELF int. | -0.001 | -0.021 | -0.177 | -0.201 |
|  | (0.077) | (0.072) | (0.187) | (0.191) |
| $R^2$ | 0.18 | 0.22 | 0.27 | 0.28 |
| $N$ | 110,718 | 108,389 | 106,575 | 95,560 |
| COLLECTIVE ACTION | | | | |
| Ethnic conflict | 0.137 | 0.144 | 0.364 | 0.294 |
|  | (0.043)*** | (0.042)*** | (0.106)*** | (0.107)*** |
| Conflict-ELF int. | -0.166 | -0.185 | -0.549 | -0.421 |
|  | (0.073)** | (0.071)*** | (0.186)*** | (0.187)** |
| $R^2$ | 0.19 | 0.21 | 0.27 | 0.26 |
| $N$ | 110,464 | 108,151 | 106,374 | 95,411 |
| MEMBERSHIP COMMUNITY ASSOCIATIONS | | | | |
| Ethnic conflict | 0.081 | 0.084 | 0.284 | 0.273 |
|  | (0.040)** | (0.040)** | (0.086)*** | (0.089)*** |
| Conflict-ELF int. | -0.092 | -0.100 | -0.350 | -0.338 |
|  | (0.068) | (0.068) | (0.147)** | (0.152)** |
| $R^2$ | 0.15 | 0.17 | 0.19 | 0.20 |
| $N$ | 110,335 | 108,029 | 106,245 | 95,243 |
| Fixed effects | Yes | Yes | Yes | Yes |
| Personal cov. | No | Yes | Yes | Yes |
| Socio-economic cov. | No | No | Yes | Yes |
| Political cov. | No | No | No | Yes |

The sample includes Afrobarometer rounds 3 to 6. Fixed effects correspond to cell $\times$ period and ethnic $\times$ period dummies. Individual controls include: age and its square, gender, and an urban-rural indicator. Socio-economic controls include: education, employment status, an index of household wealth. Political controls include: trust towards the president, trust towards the local government, and an indicator function for voting in the last national election. Standard errors are clustered at the cell-ethnolinguistic-period level and are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

# References

- Asuni, J. B. (2009). Understanding the Armed Groups of the Niger Delta. *Council on Foreign Relations*, working paper

- Hazen, J. M. and Horner, J. (2007). *Small Arms, Armed Violence, and Insecurity in Nigeria: The Niger Delta in Perspective.* Geneva, Switzerland: Small Arms Survey

- Rohner, D., Thoenig, M. and Zilibotti, F. (2013). Seeds of distrust: conflict in Uganda. *Journal of Economic Growth*, 18(3), pp. 217-252